Did Kicking the Habit Make Us Fat? The Impact of Smoking on the Likelihood of Being Overweight and Obese[*]

Peter M. Lance[a]

Mai N. Hubbard[b]

Gustavo Angeles[c]

[a]**Corresponding Author:** MEASURE Evaluation, Carolina Population Center, University of North Carolina at Chapel Hill. CB#8120, University Square, 123 West Franklin St., Chapel Hill, NC 27516-2524 USA. Phone: 1.919.843.3753. Fax: 1.919.966.2391 . Email: pmlance@email.unc.edu

[b]Department of Economics, University of North Carolina at Chapel Hill. CB#3305, Gardner Hall, UNC-Chapel Hill, Chapel Hill, NC. 27599-3305 USA. Email: mnoguchi@email.unc.edu

[c]Department of Maternal and Child Health, School of Public Health, CB#7445 University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7445 USA. Email: gangeles@email.unc.edu

Did Kicking the Habit Make Us Fat? The Impact of Smoking on the Likelihood of Being Overweight and Obese

Abstract

Two of the most pronounced health trends in the United States in the past several decades have been the precipitous decline in cigarette smoking and the striking increase in the proportions of the population that are overweight or obese. Nonetheless, relatively little research in the social sciences examines the potential link between the two developments, despite that obvious possibility given their dramatic, concurrent nature. Further, the prevailing theoretical explanations for the obesity epidemic in the economics literature, which quite sensibly focus essentially on caloric intake and energy expenditure, assign little role to the decline in smoking. However, the notion that declining cigarette smoking was a significant causal contributor to the obesity epidemic is consistent with these other explanations and can even reinforce them. We estimate the causal effect of smoking on the probabilities of being obese and overweight using a flexible semi-nonparametric instrumental variables, random effects strategy. We find that smoking significantly reduces the probability of these outcomes, and that failure to control for the endogeneity of smoking leads to underestimation of the impact of smoking on the probability of being obese or overweight. Our results contribute to a small but growing body of research that suggests that the largely successful campaign to reduce smoking in the United States may have played a significant role in the emergence of the obesity epidemic.

[JEL Classification:] I12, I18

[KEY WORDS:] Smoking; Obesity; Overweight.

1 Introduction

The past three decades have witnessed two dramatic behavioral shifts with potentially enormous implications for the health of the American people. First, smoking, in terms of both prevalence (the proportion of Americans who smoke) and intensity (e.g. overall per capita cigarette consumption) has declined dramatically, reversing trends that had prevailed essentially from the beginning of the 20th century. Second, an established, gradual upward trend in obesity prevalence by all appearances suddenly accelerated sometime around 1980: after rising from roughly 10 percent in the early 1960s to around 15 percent during the late 1970s, obesity prevalence went on to increase to roughly 30 percent by the turn of the 21st century. This manuscript considers the potential influence of the former development on the latter one.

Given the dramatic, essentially concurrent nature of these two health-related developments, there would seem at first glance to be a reasonable basis for suspecting a potential linkage between them. Moreover, a fairly substantial biomedical literature suggests physiological channels by which smoking should influence body weight. Nonetheless, the relationship between falling smoking and rising obesity has received surprisingly little research attention from health economists. The limited number of empirical inquiries to date offer little consensus regarding the possible nature or strength of the influence of cigarette smoking on body weight.

Smoking has received even less attention in theoretical explanations of the obesity epidemic. The most prominent theories offered by economists focus on caloric intake and energy expenditure. Thus, they appeal ultimately to changes to the American lifestyle that have been rather gradual in nature and hence seem unlikely to explain by themselves a sudden acceleration of the sort evident from around 1980 onward. Despite the dramatic, concurrent fall in cigarette consumption, the implicit assumption of these models seems to be that smoking could not plausibly explain an increase in obesity prevalence of the magnitude witnessed in recent decades.

In this manuscript, we suggest a way of reconciling these prevailing theoretical explanations of the obesity epidemic with the possibility that the decline in smoking in recent decades may indeed have been a critical contributing factor. We hypothesize that the comparatively high (and generally gradually rising) smoking rates in the decades preceding 1980 may have served to attenuate the overweight and obesity consequences of more gradual, incremental shifts in caloric intake and energy

expenditure to which these theories appeal. Our core argument is thus that rapidly declining smoking rates after 1980 may have generated a rather violent shift as Americans' body weights "caught up" to the level which would have prevailed had high and rising cigarette consumption not previously attenuated the consequences of slowly rising caloric intake and falling energy expenditure. Finally, we offer further empirical evidence regarding the link between smoking and body weight. Our empirical approach does not represent a dramatic methodological break with earlier work on the subject and will certainly not resolve the debate, but should help to expand the rather limited knowledge base regarding this crucial relationship.

The organization of this manuscript is as follows. The next section offers a review of trends in obesity prevalence and smoking in American society in the last several decades, reviews the empirical and theoretical literature regarding the potential link between declining smoking and rapidly rising obesity rates, and attempts to reconcile the various strands of theoretical thought on this subject. Section 3 presents empirical evidence regarding the link between smoking and body weight at the individual level. Section 4 concludes.

2 Smoking and Obesity: Theory and Evidence

2.1 Trends in Obesity: Empirical Patterns and Theoretical Explanations

Time series information regarding the population-level prevalence of overweight and obesity[1] in the United States is surprisingly spotty. For the most part our understanding is based on periodic snapshots formed from national health interview surveys which have occurred with varying frequency over time, though comprehensive monitoring of the prevalence of obesity and overweight has become far more routine in recent years. Nonetheless, a relatively coherent picture, presented in Figure 1,[2] makes clear that gradual increases in obesity prevalence from the early 1960s suddenly accelerated around 1980. Indeed, the presentation of the years on the horizontal axis of Figure 1,

---

[1]Following convention, we define obesity and overweight using Body Mass Index (BMI), which is simply a ratio of weight in kilograms to height in meters squared. The Centers for Disease Control and Prevention (CDC) defines merely overweight to be a BMI greater of 25 or more but less than 30, and obesity as a BMI of 30 or more. (When we refer to overweight, as opposed to *merely overweight*, we mean simply a BMI exceeding 25; such an individual could be either merely overweight or obese.)

[2]The data behind Figure 1 come from a variety of sources that appear to be widely relied upon by researchers. In particular, the obesity prevalence figures are based on the various National Health and Nutrition Examination Surveys (including the National Health Examination Survey, Cycle I for the early 1960s) and the 2007 OECD Health Data database.

with more compact spacing for the first two decades when observations were less frequent but the average annual increase in obesity was likely far more modest, probably visually *understates* the dramatic magnitude of the acceleration in obesity prevalence.

Although Figure 1 shows the prevalence of obesity for the overall population, the course for most population sub-groups (e.g. adults over 18) was similar. It is also worth noting that the proportion of Americans merely overweight (but not obese) remained roughly constant at around 30 to 35 percent between the early 1960s and the turn of the 21st century. Presumably the number of merely overweight Americans who became obese was matched by the number of normal weight Americans who became merely overweight.[3]

The precipitous rise in obesity prevalence and the overall proportion of Americans with a BMI of 25 or more has inspired a variety of theoretical explanations. Two approaches, respectively focusing particularly on energy expenditure and overall caloric intake, have attracted particular attention in the economics literature. Both emphasize the role of technological change, as the United States has evolved from an agrarian to a post-industrial consumer society. Philipson and Posner (1999) place special emphasis on the shift to more sedentary, less physically taxing jobs that have resulted in a decline in caloric expenditure. Cutler and Shapiro (2003), on the other hand, argue that declining food preparation costs, as a shift has occurred from time-consuming home cooking toward more centralized preparation, have allowed Americans to consume more food and hence increase their caloric intake. These are compelling explanations with solid foundations in established developments in American life in the latter half of the 20th century and arguably do capture the central forces behind the upward trend in the overall proportion of Americans with a BMI of 25 or more.

At the same time, Figure 1 suggests the likely presence of other, mediating forces in this process. In particular, both theoretical approaches are rooted in rather gradual, incremental changes to American life which certainly would seem to portend increasing body mass, but presumably in the same gradual fashion. However, as we have seen, the prevalence of obesity appears to have risen dramatically faster from 1980 onward. Against this, there has on balance been essentially no discernable correspondingly sudden, dramatic change in the evolution of the causal factors to

---

[3]The actual story is obviously more complex, as some Americans actually lost weight while others gained dramatically, but this statement probably captures the essence of the situation in some over-riding sense.

which these theories appeal.

In Figure 2 we illustrate this point with reference to the focus on changing caloric intake. Cutler and Shapiro emphasize the importance of changing nutritional patterns in driving the steadily rising body mass of Americans. The sub-figures presented in Figure 2 plot obesity prevalence against four nutritional indicators (per capita intake of sugars, fats and fruits and overall caloric intake) from 1960 to 2004.[4] To be sure, all four evolved in a fashion that provides generally consistent *prima facie* evidence for the food intake argument. Their comparatively smooth evolution over time does not, however, suggest an obvious explanation for the sudden acceleration in the upward trend in obesity prevalence evident from around 1980.

2.2   Trends in Smoking Prevalence

One health-related indicator that did shift abruptly around 1980 was tobacco consumption (particularly cigarette smoking, the main tobacco delivery mechanism in the 20th century). To begin with, annual per capita cigarette consumption rose dramatically from 35 to 2,645 from the turn of the 20th century to 1960, where Figure 3 picks up the story.[5] Per capita cigarette consumption then more or less rose gradually to around 2,800 by the late 1970s. Thereafter, it began to decline sharply through the 1980s and 1990s, reaching around 1,550 cigarettes per capita at the turn of the 21st century. This was reflected by a steep decline in cigarette smoking prevalence from 37.4 percent in 1970 to 20.9 percent in 2004.[6]

Plotting cigarette smoking and obesity prevalence together (as in Figure 4) reveals the essentially concurrent shifts to the trends in the two. Similar patterns are observed, in Figure 5, in some other OECD countries, including Australia, Finland and the Netherlands. In all three nations, the two indicators appear to have followed remarkably inter-related courses: declines in smoking are often closely matched by increases in the percentage obese. To be sure, however, there are examples where this kind of casual empiricism fails to reveal this pattern (e.g. Japan, in the Southeast

---

[4]The nutritional data in Figure 2 is drawn from the 2007 OECD Health Data database. While the data behind Figure 2 may be even shakier and more sporadic than that behind figure 1, clear patterns are nonetheless apparent.

[5]The source for these statistics and Figure 3 is the Tax Burden on Tobacco, which primarily presents overall annual cigarette consumption per capita. However, the time series for other, perhaps more specific measures such as smoking per adult over age 18 follow a very similar pattern. Figure 3 presents per capita cigarette consumption for the years for which we observe obesity prevalence in Figures 1 and 2.

[6]This data comes from the Centers for Disease Control and Prevention. For further information on smoking prevalence, readers should refer to: http://www.cdc.gov/tobacco/data_statistics/index.htm

corner of Figure 5).

The cross-national trends presented in Figure 5 compare obesity rates and tobacco consumption, which admittedly might not be the main story in other OECD countries, where rates of obesity are generally lower than in the United States. Indeed, the proportion of society merely overweight has remained relatively stable in the United States in recent decades while the real action, so to speak, has involved the obesity rate. The reverse is true in many other OECD countries, where the epidemic has been primarily in growth of the proportion merely overweight. For the four nations presented in Figure 5, we thus plot tobacco consumption against the prevalence of overweight in Figure 6. A pattern between the two similar to that in Figure 5 emerges.

2.3   Previous Literature on the Link Between Smoking and Body Mass

Despite the policy urgency of the obesity epidemic and strong *prima facie* aggregate time series evidence for a possible link between smoking and body mass, the health economics literature has produced relatively few empirical studies of the causal impact of the former on the latter, and still less consensus regarding the magnitude and direction of this relationship. In perhaps the first word on the subject in health economics of which we are aware, Chou et al. (2004) explored the relationship as part of a broad investigation of the determinants of obesity. Applying numerous estimation methods to a sample from the Behavioral Risk Factors Surveillance System (BRFSS), the authors pursued a reduced-form model of obesity that related it to several key potential determinants (such as prices, education, and occupation type) and recovered a positive relationship between obesity and cigarette prices. This implies that smoking indeed lowers the probability of obesity, assuming the theoretically anticipated relationship between smoking and cigarette prices.[7]

Chou et al. (2004) was followed by two studies which together offer little endorsement of its key finding. First, Gruber and Frakes (2006) outlined and then addressed in their own estimations some possible weaknesses[8] of Chou et al. (2004). They found a negative relationship between obesity and cigarette taxes. This should, by Gruber and Frakes ready admission, be interpreted with care for two reasons: their sample appears unusual in certain respects and, more importantly, a positive

---

[7]Which is probably a safe thing to do, in light of the large literature on the price sensitivity of cigarette demand; see Chaloupka and Warner (2000) for a thorough review.

[8]These include a potential endogeneity problem arising from Chou et al.'s reliance on cigarette prices rather than taxes and the reliance of Chou et al. on quadratic time trends.

causal relationship between smoking and obesity runs counter to the implications of essentially the entire biomedical literature on the subject. Second, Rashad (2006) explored the relationship using a more structural approach and a sample from the National Health and Nutrition Examination Survey (NHANES), and recovered a positive but insignificant effect of smoking on obesity. This branch of the literature thus provides three compelling papers that reach very different conclusions about the same essential causal relationship.

Perhaps unsurprisingly, this relationship has received somewhat more attention in the mainstream public health literature, but with essentially no more degree of consensus emerging. Albanes et al. (1987), a comparatively early contribution, found smoking to be associated with lower body mass. A more recent contribution is provided by Flegal (2007), who essentially attempts to estimate the trajectory of obesity prevalence had smoking not declined, and finds it hardly altered from what actually did occur. While the public health branch of this literature is fairly large, a reasonable criticism that could be applied to virtually all work in it is that it fails to address the potential endogeneity of cigarette smoking.

A slightly different vein of inquiry considers the effect of quitting smoking on body mass. The US Surgeon General, for instance, found that quitters gained four pounds on average (US Department of Health and Human Services, 1990)[9] while many other studies (e.g. Pisinger and Jorgensen 2007 for a recent example) place the figure closer to 10 pounds. However, some (e.g. Caan et al., 1996, Froom et al., 1999, Mizuoue et al., 1998) have found that these gains level off with of the passage of time from the cessation event; a few even reported that some of the gained weight was eventually lost. Recognizing the selectivity implications of quitting, Eisenberg (2006) applied an instrumental variables estimation approach to data from smoking cessation trials and found a substantial body mass increase five years after cessation: quitters and non-quitters differed in weight by an average of 21.4 pounds, a figure that is about two to four times the weight gain estimated by most previous studies. Thus, while this literature has produced a certain degree of consensus that smoking cessation is associated with weight gain, the extent or sustainability of those gains remains far from resolved.

The biomedical literature has, by contrast, provided a generally greater degree of consensus

[9]See, for instance, Williamson et al. 1991 for another example of a study along these lines

regarding the relationship between smoking and body mass or, more specifically, between nicotine[10] and the mediating factors behind body mass.[11] Perhaps most importantly, nicotine consumption is associated with increased energy expenditure (Hofstetter et al., 1986; Dallosso and James 1984; Collins et al., 1994; Moffatt and Owens 1991, etc.). It may also act as an appetite suppressant (Jessen et al., 2005; Perkins et al., 1991). While there have unsurprisingly been some studies that differ in subtle but important ways from this conventional wisdom [12] and there remain areas where understanding is limited,[13] it seems fair to conclude that the balance of biomedical evidence suggests that nicotine acts to speed metabolism and suppress appetite. Both have clear implications for body mass.

The empirical literature regarding this potentially critical causal relationship thus provides a mixed range of results, but a few points are worth emphasizing. First, taken as a whole the results have been somewhat lopsided in the sense that nearly all of the findings of a significant relationship carry the implication that smoking lowers body mass. Few studies find a significant relationship in the opposite direction, even implicitly. Second, outside of the limited body of work by economists and that work in the biomedical tradition which can be viewed as essentially experimental in nature, few studies address the endogeneity of smoking. Finally, though many studies have, to be sure, recovered no significant relationship, those that have are sufficiently numerous that a significant link between smoking and obesity cannot be ignored as a scientific possibility. In many instances the link has been found to be quite strong, implying that declining smoking rates may have made a potentially important contribution to the obesity epidemic.

2.4 Smoking and the Prevailing Theoretical Explanations of Obesity

Though the prevailing explanations in the economics literature for the increasing proportions of Americans obese and overweight would seem to focus rather exclusively on long-term technological and lifestyle-related change that affected caloric intake and energy expenditure, cigarette smoking may plausibly have served as a sort of critical mediating factor in the health production process

---

[10] As Dr. Jeffrey Wigand famously explained on *Sixty Minutes*, from a biochemical standpoint cigarettes are, first and foremost, nicotine delivery devices.

[11] Chiolero et al. (2008) provide an excellent and timely overall review.

[12] For example, Perkins et al. 1992 found that those administered nicotine actually ate *more* than those who received a placebo.

[13] For instance, Ciolero et al. (2008) point out that the long term metabolic effects of nicotine remain poorly understood.

underlying the essential relationships to which these theories appeal. Specifically, the increase in smoking levels through the late Seventies may have attenuated the gradual increases in body mass that would have occurred, given gradually rising caloric intake and falling energy expenditure, in its absence. The sudden removal of such a moderating factor could have then caused a sudden increase in body mass as it "caught up" with its fundamentally appropriate nicotine-free level.

This idea is fairly simple and we begin with a generalized graphical representation in Figure 7 of theoretically hypothetical evolving trends in body mass and lifestyle at the population level. Suppose that lifestyle (L) is measured in such a fashion that it is thus rising over time.[14] Moreover, assume that this upward trend would suggest that indicators associated with body mass (e.g. average BMI, prevalence of obese and overweight, etc.) should likewise be rising over time. To fix ideas, in Figure 7 the representative indicator of body mass is the percent obese (O).[15] The interpretation of this figure is thus that there is a time-trend to lifestyles L that we take as given[16] that is driving a steady increase in the obesity rate O. Thus far, this is fairly straightforward. For instance, as net caloric intake (representing, roughly, gross caloric intake less daily energy expenditure) rises so too should body mass, *ceteris paribus.*

Now suppose for simplicity that there are only six potential levels of cigarette smoking consumption, represented by S_j for $j = 1, ..6$, where S_1 and S_6 represent, respectively, the highest and lowest cigarette smoking intensity levels. Holding constant the trend in L introduced above, Figure 8 presents the obesity prevalence-lifestyle relationship for each level of cigarette smoking by the line denoted by $O(S = S_j)$ for $j = 1, .., 6$.[17] Put simply, for any given level of lifestyle L, lower smoking implies lower obesity. Mathematically, a pattern such as that in Figure 8 could be represented by the following linear relationship between obesity (O), cigarette smoking consumption (represented generically by S), and lifestyle (L):

$$O = \beta_0 + \beta_1 \cdot L + \beta_2 \cdot L \cdot S + \beta_3 \cdot S$$

where $\beta_1 > 0$, $\beta_2 < 0$, $\beta_3 < 0$. Once again, this is a fairly straightforward expression: taken as a

---

[14]Applying this to the American experience in recent decades, one could view L as hence representing, in some rough fashion, net caloric intake per day.

[15]Without any loss of generality, it could be re-cast in terms of other outcomes, such as the percent overweight.

[16]Explaining the forces shaping the trend in L over time is beyond the scope of this manuscript. Refer to Philipson and Posner 1999, Cutler and Shapiro 2003 and others for excellent discussions of the forces shaping L.

[17]Henceforth, for visual simplicity, the L axis is omitted.

whole, the social science and biomedical literature suggests that, on balance, it is likely that *ceteris paribus* smoking lowers body mass. Very few studies find the opposite; nearly all that conclude that this inverse relationship does not exist simply find no statistically significant relationship at all. The debate, it is reasonable to say, is really about the strength of the relationship between the two.

The consequences of a change in smoking consumption levels over some interval of time is shown in Figure 9. If we allow smoking consumption levels to decrease from S_1 to S_6, there is a movement away from obesity prevalence trajectory $O(S = S_1)$ toward that given by $O(S = S_6)$. This would imply a sudden acceleration in the increase in the obesity rate as it not only rose due to the ongoing secular trend in L but, more importantly, "caught up" with the obesity trend line it should be on at the new lower smoking level S_6. An analogous possibility can emerge as cigarette smoking rises, as shown in Figure 10: increasing cigarette consumption results in an attenuation of the increases in obesity prevalence below what otherwise would have prevailed, given evolving lifestyles L.

These two theoretical stories are broadly consistent with events in the post-WWII United States: despite a gradual, fairly steady trend toward lifestyles that would portend rising body mass, the obesity rate rose comparatively slowly until roughly 1980. In the context of Figure 10, the obesity consequences of steadily more sedentary, higher calorie intake lifestyles were partly attenuated by high and generally gradually rising cigarette consumption. Then, around 1980, cigarette consumption began to fall dramatically, leading to a rapid increase in the obesity rate as it "caught up" with the levels that would have prevailed in the absence of the high, gradually increasing cigarette consumption levels witnessed before that time (as shown in Figure 9).

This theoretical approach to the link between smoking and obesity does not necessarily rely on a particularly strong relationship between the two. Rather, in principal it requires simply some significant inverse relationship of reasonable magnitude. It is plausible that even a modest link between the two could allow the gap between the obesity prevalence O that should prevail, given lifestyles L and low cigarette consumption S_6, and those that actually would prevail under high cigarette consumption S_1, to grow large with sufficient time. This would open the possibility of a very steep adjustment in obesity prevalence O in response to a precipitous fall in smoking from S_1 to S_6.

In reality, the link between smoking and obesity at the population level would reflect not only

direct exposure by cigarette smokers but also secondhand exposure by non-smokers. A large literature finds secondhand smoke exposure to be deleterious across a variety of channels of health (see Glymour et al. 2008 for a very recent example) due essentially to the same biomedical mechanisms behind the direct effects that smokers experience. It seems reasonable that a similar sort of externality should apply in the case of body mass.[18] In the period after 1980, overall cigarette smoke exposure, so defined, arguably actually fell even more steeply than cigarette consumption itself. For instance, recent decades have witnessed increasingly stringent restrictions on smoking in the workplace and public spaces that likely served to reduce secondhand smoke exposure even more quickly than might have been expected before those regulations were instituted. Finally, an obvious implication of this logic is that the growth in the obesity rate should eventually begin to slacken as further reductions in cigarette smoking become more modest.[19].

The argument presented in this section is not novel, in any sense. In effect, it simply restates the assertion of one camp: that the decline in smoking contributed to the increase in overweight and obesity experienced over the past three decades. However, it appeals to a subtle interplay between gradually evolving lifestyles and smoking rates. By appealing to the notion that the rapid increase in obesity and overweight prevalence that occurred after 1980 reflected a sharp adjustment between fundamental obesity trajectories implied by different cigarette consumption levels, this approach is somewhat logically akin to certain models of speculative crises (e.g. Salant and Henderson 1978, Krugman 1979) that focus on an ever-growing gap between a commodity or currency's actual price and fundamental value as the impetus for a sudden, abrupt adjustment in price.

3   The Causal Link Between Smoking and Obesity and Overweight: An Empirical Investigation

The story presented in the last section relies on a preliminary assumption: that there is indeed a significant casual link between smoking and obesity. As the review of empirical work on the subject makes clear, this is at present far from established (particularly in the health economics literature),

---

[18]Extending this argument, the rise in childhood obesity evident in recent decades may be due in part to a reduction in secondhand smoke exposure, particularly in the home.

[19]As seems likely to be the case as the group of remaining smokers is likely more committed to the habit and, frankly, smoking levels become so low that the scope for further dramatic reductions is mathematically reduced

though the results to date do hint at essentially a one-sided bet regarding the possible direction of this link. Of course, simply establishing a significant inverse relationship between smoking and obesity is not sufficient to substantiate fully the theory presented in the last section.[20] Nonetheless, for present purposes, the focus will be on the simply estimating the direct causal effect of cigarette smoking on the probability of being overweight and obese.

3.1 Empirical Model

We seek to estimate the causal impact of smoking on the probabilities of being overweight and obese (through separate estimations focused solely on the probability of being overweight and that of being obese). An obvious challenge is the likely endogeneity of smoking as a predictor of either the probability of being overweight or of being obese. To address this possibility we jointly estimate equations determining the probabilities of smoking and of being overweight or obese in a fashion that permits us to integrate out with respect to confounding unobservables.

Let S_{it} be an indicator of smoking such that

$$S_{it} = \begin{cases} 1, & \text{if individual } i \text{ currently smokes at time } t; \\ 0, & \text{otherwise.} \end{cases}$$

Let O be an indicator that an individual's BMI is greater than or equal to some threshold (25 in the case of estimations focused on the probability of being overweight, 30 in the case of estimations focused on obesity). Specifically, let O_{it} be defined as follows:

$$O_{it} = \begin{cases} 1, & \text{if individual } i \text{ achieves or exceeds the threshold BMI at time } t; \\ 0, & \text{otherwise.} \end{cases}$$

Finally, let $\mathbf{X}_{it}$ be a series of socioeconomic and demographic characteristics of individual i at time t and $\mathbf{Z}_{it}$ be a vector of variables that influence $\Pr(S_{it} = 1)$ but not $\Pr(O_{it} = 1)$ (i.e., $\mathbf{Z}_{it}$ are instruments for smoking).

---

[20]It is merely *necessary*: sufficiency would require establishing its magnitude in the fullest sense (including by way of the interplay with the sort of lifestyle-related factors to which the theories of Cutler and Shapiro 2003 and Philipson and Posner 1999 appeal), a task beyond the scope of this manuscript.

Behind the discrete outcomes O and S are the latent variable processes

$$O_{it}^* = \beta_0 + \beta_1 \cdot \mathbf{X}_{it} + \beta_2 \cdot S_{it} + \mu_{it}^O + \epsilon_{it}^O$$

and

$$S_{it}^* = \delta_0 + \delta_1 \cdot \mathbf{X}_{it} + \delta_2 \cdot \mathbf{Z}_{it} + \mu_{it}^S + \epsilon_{it}^S$$

where the ϵs and μs are unobserved. We assume that the ϵs are independently and identically distributed but the μs are correlated (and hence create a channel of confounding unobservables which might render S an endogenous determinant of O). We thus introduce an error component structure, with one part independent across equations (the ϵs) and another correlated across them (the μs).[21]

Assuming that in each instance the ϵs are the difference of two Type-I Extreme Value distributed random variables gives rise to logit choice probabilities for each of the discrete outcomes (O_{it} and S_{it}) and the following contribution to the likelihood function for individual i at time t:[22]

$$L_{it}\left(\mathbf{X}_{it}, \mathbf{Z}_{it} | \mu_{it}^O, \mu_{it}^S\right) = \left(\frac{\exp\left(\beta_0 + \beta_1 \cdot \mathbf{X}_{it} + \beta_2 \cdot S_{it} + \mu_{it}^O\right)}{1 + \exp\left(\beta_0 + \beta_1 \cdot \mathbf{X}_{it} + \beta_2 \cdot S_{it} + \mu_{it}^O\right)}\right)^{O_{it}}$$

$$\cdot \left(\frac{1}{1 + \exp\left(\beta_0 + \beta_1 \cdot \mathbf{X}_{it} + \beta_2 \cdot S_{it} + \mu_{it}^O\right)}\right)^{1-O_{it}} \cdot \left(\frac{\exp\left(\delta_0 + \delta_1 \cdot \mathbf{X}_{it} + \delta_2 \cdot \mathbf{Z}_{it} + \mu_{it}^S\right)}{1 + \exp\left(\delta_0 + \delta_1 \cdot \mathbf{X}_{it} + \delta_2 \cdot \mathbf{Z}_{it} + \mu_{it}^S\right)}\right)^{S_{it}}$$

$$\cdot \left(\frac{1}{1 + \exp\left(\delta_0 + \delta_1 \cdot \mathbf{X}_{it} + \delta_2 \cdot \mathbf{Z}_{it} + \mu_{it}^S\right)}\right)^{1-S_{it}}$$

---

[21] An alternative approach would be to assume that the errors ξ_{it}^O and ξ_{it}^S, where

$$\xi_{it}^O = \mu_{it}^0 + \epsilon_{it}^O$$

and

$$\xi_{it}^S = \mu_{it}^S + \epsilon_{it}^S,$$

follow a joint normal distribution (in other words, eschew the error components structure and assume that the stochastic portion of the indirect utility functions follow a joint normal distribution). This would give rise to the bi-variate probit model.

[22] Once one recognizes the correlation of the μs, the likelihood function can no longer be regarded as separable for purposes of obtaining consistent estimates of the effect of smoking on either the probability of being overweight or obese. However, once can still estimate O and S equations in isolation. Both will give some idea of the efficiency consequences of joint estimation, while the former provides a potentially biased and inconsistent estimate of the causal effect smoking on O which can serve as a baseline against which to consider the consistent estimates obtained through our joint estimation strategy.

Regardless of the assumption made regarding the distribution of the errors ϵ, one would still confront the conditional likelihood $L_{it}(\mathbf{X}_{it}, \mathbf{Z}_{it}|\mu_{it}^O, \mu_{it}^S)$ under the error components approach. The challenge is then to recover consistent estimates of the conditional distribution of $\Pr(O_{it} = 1|\mathbf{X}_{it}, S_{it})$ given the presence of the confounding unobservables μ_{it}^O and μ_{it}^S (or, in other words, to recover the unconditional likelihood $L_{it}(\mathbf{X}_{it}, \mathbf{Z}_{it})$).

The usual estimation strategy involves imposing some parametric assumption regarding the joint distribution of the errors μ (joint normality is a typical choice) and estimating by means such as Monte Carlo integration. The drawback to this approach (as well as to the bivariate probit model discussed in an earlier footnote) is the adoption of a potentially strong and typically un-testable assumption regarding the distribution of a critical component of the errors (the confounding elements μ). We take an alternative approach by adopting a flexible, non-parametric discrete approach joint distribution of the μs motivated by Heckman and Singer (1984).[23] This basic strategy has been analyzed in the context of multi-equation systems (where the challenge is to model the joint distribution of unobservables) by Mroz (1999). In particular, Mroz (1999) found that the Heckman and Singer (1984) strategy performed far better than more parametric approaches to the joint distribution of the unobservables.

Under this approach, individual i's unconditional contribution to the likelihood function at time t would be

$$L_{it} = \sum_{k=1}^{K} \pi_k \cdot L_{it}\left(\mathbf{X}_{it}, \mathbf{Z}_{it}|\mu_{it}^O = \mu_k^O, \mu_{it}^S = \mu_k^S\right)$$

where

$$\pi_k = \Pr\left(\mu^O = \mu_k^O, \mu^S = \mu_k^S\right)$$

and we assume that there are K points of support to the discrete joint distribution of the unob-servables μ. A standard and necessary normalization is

$$\mu_1^O = 0$$

---

[23]Heckman and Singer (1984) advocated estimation of a discrete approximation to the distribution of a persistent unobservable in a hazard model (rather than a more parametric assumption regarding the distribution of that unob-served heterogeneity, which they find can lead to poor estimation performance) in order to address a potential source of confounding that would otherwise lead to a biased estimate of the time-shape of the hazard model.

and

$$\mu_1^S = 0$$

The probabilities are parameterized through the logistic function as follows:

$$\pi_k = \frac{exp\left(\theta_k\right)}{1 + \sum_{j=2}^{K} exp\left(\theta_j\right)}$$

for $k = 2, ..., K$ and

$$\pi_1 = \frac{1}{1 + \sum_{j=2}^{K} exp\left(\theta_j\right)}$$

The θs are thus estimated along with the other parameters of the model.

In practice K is determined by an iterative process. First, estimate the model with 1 point of support. Then add points of support one at a time and re-estimate until the likelihood function fails to improve from the last estimation.[24] Under this approach, the data is thus allowed to determine fully the joint distribution (i.e., the number of mass points, as well as their locations and probabilities).

One issue that has not been addressed yet is the longitudinal nature of the PSID: there are individuals in the data set who appear as observations more than once over time. Suppose individual i is observed T_i times, where

$$1 \leq T_i \leq 4$$

Then, an individual's total contribution to the likelihood function is given by

$$L_i = \prod_{t=1}^{T_i} \sum_{k=1}^{K} \pi_k \cdot L_{it}\left(\mathbf{X}_{it}, \mathbf{Z}_{it} | \mu_{it}^O = \mu_k^O, \mu_{it}^S = \mu_k^S\right)$$

To account for the likely non-independence of an individual's observations over time, we compute Huber-White-type standard errors assuming clustering at the individual level. A richer approach to the modeling of the error terms might have yielded some, probably limited information at the margin, but at the cost of either unsatisfactory behavioral assumptions or potentially considerably greater computation complexity, possibly reducing the accuracy (from, for instance, a mean-squared

---

[24]Specifically, we pursue the rough rule of adding points of support until the likelihood fails to improve by at least the number parameters added to the model through the addition of another point of support (which would be three in this model). We did continue estimation beyond that point, if only to insure the stability of the results under this stopping criteria.

error standpoint) of estimates of the parameter of interest.[25]

Finally, our primary parameter of interest will be the marginal effect of smoking on the probabilities of being obese or overweight:

$$\Pr\left(O_{it} = 1 | S_{it} = 1, \mathbf{X}_{it}, \mathbf{Z}_{it}\right) - \Pr\left(O_{it} = 1 | S_{it} = 0, \mathbf{X}_{it}, \mathbf{Z}_{it}\right)$$

We compute this for each observation in our sample and then report the average of these marginal effects across them. Standard errors for this marginal effect are recovered via the Delta method.[26]

3.2 Data Source and Sample

Our estimation sample is drawn from the Panel Study of Income Dynamics (PSID), a longitudinal, nationally representative survey of American families and individuals (men, women, and children). The study began in 1968, when 4,800 families were selected for interview. Following the baseline panel in 1968, the PSID was conducted annually until 1997. Follow-ups have been collected biennially since then. The PSID routinely gathers detailed information on respondents' family structure, employment and marital histories, and income and wealth. Information on the height and weight of the heads of PSID families and their spouses has been collected since 1999, allowing their BMIs to be calculated for four years: 1999, 2001, 2003, and 2005. We classify respondents as overweight or obese if their BMIs achieved the commonly accepted thresholds of 25 or more and 30 or more, respectively. We define a respondent as a smoker if they report currently smoking a positive number of cigarettes per day as of the time of interview.

Cigarette prices were obtained from *The Tax Burden on Tobacco* (Orzechowski and Walker, 2005). We match each respondent with the average price of a pack of cigarettes (including generic

---

[25]For instance, another possible way of addressing this would have been to specify the overall contribution of individual i as

$$L_i = \sum_{k=1}^{K} \pi_k \cdot \prod_{t=1}^{T_i} L_{it}\left(\mathbf{X}_{it}, \mathbf{Z}_{it} | \mu_{it}^O = \mu_k^O, \mu_{it}^S = \mu_k^S\right)$$

This is tantamount to specifying the confounding unobservable as fixed over time, an approach that we reject since it seems behaviorally unreasonable: individuals in our sample transitioned into or out of being overweight/obese or currently smoking, suggesting that the confounding unobservables may not have been constant over time. Another approach would have been to model fixed and time-varying elements of the unobservables. A potential difficulty with this is the limited number of observations (at most 4) per individual, rendering it potentially challenging in practice to identify separately time-varying and fixed unobservables (in principal, they are identified by the longitudinality of our sample).

[26]Identification tests involving the retail price of smoking, a continuous variable, are obtained by simulating a one percent increase in it, with the standard errors similarly obtained via the Delta method.

brands) in their state of residence at the time of interview. These prices are deflated to 1982-1984 prices using the Consumer Price Index (CPI) from the Bureau of Labor Statistics. Other tobacco control-related policy variables were also merged into the sample by state and year. These included indicators of smoking prohibitions in government or private sector workplaces, and restaurants using information readily available from CDC data on clean indoor air regulations.

Descriptive statistics for the sample are reported in Table 2. It consists of 30,584 individual-year observations. Approximately 25 percent of the sample is obese, a figure somewhat lower than the rate (of 30 percent) found by the 2003-4 National Health and Nutrition Examination Survey (NHANES). Thus this sample is, like that used by Gruber and Frakes (2006), somewhat atypical, but there is no reason to believe that this will necessarily preclude estimates that are at least internally consistent.[27] 62 percent are overweight. Around 22 percent of the sample reports currently smoking. Finally, blacks are over-represented (at 27 percent), around 75 percent of the sample has less than a college education, and the average age of the sample is around 41 years.

3.3 Results

We begin with the effect of smoking on the probability of being obese, estimation results for which are in Tables 2 and 3. In a pattern that will be repeated for modeling of the probability of being overweight, Table 2 provides the parameter estimates for logit models of smoking and obesity. Results in the left-hand columns are drawn from separately estimated models of smoking and obesity (thus offering no control for the potential endogeneity of smoking in the obesity equation) and those on the right from a jointly estimated model of the two (that thus controls for the endogeneity of smoking in the obesity equation using the Heckman and Singer 1984 discrete approximation approach described earlier). At the bottom of Table 2 are two estimates of the marginal effect of smoking on the probability of being obese: on the left, that from the obesity model estimated in isolation and, on the right, that from the joint-estimation model controlling for confounding unobserved heterogeneity. Finally, Table 3 provides the estimated mass points and probability weights for the distribution of the confounding unobservables.

Beginning with the single equation estimates with no controls for the possible endogeneity of

---

[27]Moreover, we would expect there to be some difference in most indicators computed from the two samples, since the PSID is a longitudinal sample originally representative of its target population in 1968, while NHANES is a fresh probability sample broadly representative of the American population at the time that it was gathered.

smoking, many of the explanatory variables in both the smoking and obesity equation proved significant, a finding that is perhaps not particularly surprising given that many of these variables often prove significant in models of obesity and smoking. Such naive (ie without recognition of the endogeneity of smoking) estimation of the obesity equation yielded a marginal effect of smoking on the probability of being obese of -0.083, which was significant at the 1 percent level.

Turning to the results from joint estimation of the obesity and smoking equations (in the last three columns of Table 2), the same general pattern of statistical significance to the estimated parameters of the logit models persisted, if on a somewhat more tenuous basis. The instrument in this model, cigarette prices, was a significant predictor of the probability of smoking with a marginal effect of cigarette prices on the probability of smoking that had a t-statistic of -2.589.[28] Crucially, the marginal effect of smoking on the probability of obesity more than doubled in magnitude under joint estimation with control for confounding unobserved heterogeneity, to a statistically significant -.174 (though the standard error suggested less precision in estimation than had been the case with the obesity equation estimated in isolation, with no control for the endogeneity of smoking). Controlling for the endogeneity of smoking thus led to a far stronger marginal effect of smoking on the probability of obesity. This is an intuitively reasonable finding. For instance, those who smoke may on balance have generally lower health preferences, leading to underestimation of the effect of smoking on the probability of obesity from simple single equation models.

The estimated heterogeneity distribution parameters in Table 3. Estimation was stopped at 6 mass points (as per the stopping criteria described earlier). In virtually every variation on this model considered (e.g. the various identification tests) the likelihood improvement criterion led to

---

[28]This is a test of one important criteria for an instrument: "first-stage" explanatory power (in this context, the ability of cigarette prices to predict the probability of smoking). The other is "second-stage" exogeneity (ie whether cigarette prices are not a significant predictor of the probability of obesity, once one has controlled for the endogeneity of smoking in the obesity equation), which we tested by jointly estimating the smoking and obesity equations with the unobserved heterogeneity controls, which provide identification by non-linearity as a basis for over-identification, and the cigarette price variable in both the smoking and obesity equations. Cigarette prices proved to be insignificant in the obesity equation (for instance, the t-statistic for the marginal effect of cigarette prices on the probability of being obese is -0.305). Given Gruber and Frakes (2006) concern with using prices rather than taxes, we re-estimated the full set of models with cigarette taxes. The same essential pattern of results emerged, though estimation proved slightly more challenging (an experience that we attribute to the fact that taxes generally had less "first-stage" power than prices). It is not clear to us that taxes should be less endogenous than prices: taxes in part reflect voter sentiment, and as such may be correlated with unobservables in the smoking demand equation. At the same time, price variation may reflect many supply-side related shifters, such as state-level retail operating costs. In any case, cigarette prices have an established track record as instruments for smoking in many contexts (see, to cite one recent example, Auld 2005). We did not use the other tobacco control policy variables as instruments because they did not consistently pass over-identification tests of "second stage" exogeneity, perhaps due to the degree to which they reflect voter sentiment.

stopping with either 5 or 6 mass points. The mass points themselves reveal the presence of extremely strong unobservables from an explanatory power standpoint. For instance, the second mass point indicates that roughly 13.5 percent of the sample belongs to an unobserved subgroup who essentially are never obese but do have a somewhat greater than normal tendency to smoke. These large mass points also suggest the source of ongoing difficulties estimating this model with STATA's mvprobit command, which would not converge beyond the third significant decimal place for parameter estimates, regardless of the number of draws employed.[29] From the estimated distribution of the unobserved heterogeneity presented in Table 3, it is clear that joint-normality was a poor fit to the data. Though this is a pronounced distribution to the unobserved heterogeneity, it is one with which are fairly confident since its' basic contours emerged very quickly as mass points were added and persisted across the various specifications of this model that we considered.

We now turn to the results from estimation of the effect of smoking on the probability of being overweight (i.e. BMI $\geq$ 25), presented in Tables 4 and 5. In general, estimation of these models proved slightly easier than those concerned with obesity. For the most part this reflected the more subtle estimated joint distribution of the confounding unobservables. The single equation logit results (i.e. separate estimation) in Table 4 reveal a statistically significant marginal effect of smoking on the probability of being overweight of -.112. With joint estimation of the smoking and overweight equations with control for confounding unobserved heterogeneity, that figure once again grew in magnitude, to a statistically significant -.178.[30] In other words, essentially the same pattern emerged as did with obesity: single equation estimation of the overweight equation with no control for the endogeneity of smoking appears to have led to underestimation of the effect of smoking on the probability of being overweight. Table 5 provides the estimated joint distribution of the unobserved heterogeneity, which in this instance had 5 points of support. Once again, this distribution had some pronounced mass points, though nowhere near to the degree seen in the case of obesity (perhaps this is unsurprising, since a much larger proportion of the sample was overweight). Moreover, the probabilities for these mass points were generally far more evenly

---

[29]Nonetheless, the estimated marginal effects under mvprobit were consistent with those reported herein.

[30]In terms of the validity of the price instrument, the marginal effect of cigarette prices on the probability of smoking had a t-statistic of -1.998. An over-identification test similar to that described in an earlier footnote concerning the obesity results yielded a marginal effect of cigarette prices on the probability of being overweight with a t-statistic of -.2. Finally, substitution of cigarette taxes for cigarette prices yielded essentially the same results, though "first stage" explanatory power was again lower (i.e. taxes were less powerful predictors of the probability of smoking than prices).

balanced than had been the case with the obesity equation.[31]

It would be fair to say that we are slightly more confident in the overweight than obesity results: key relationships were generally estimated with somewhat more precision, the heterogeneity distribution was less pronounced (rendering estimation computationally more straightforward), etc. However, both sets of models yielded a remarkably similar and persistent pattern of results:

- Even in the absence of endogeneity controls, smoking had a significant, negative effect on the probability of being obese and overweight;

- That effect roughly doubled with endogeneity controls involving joint estimation under a flexible instrumental variables, random effects approach;

- Cigarette prices exhibited "first-stage" explanatory power and "second-stage" exogeneity, which, combined with the consistently stable pattern of results obtained, suggests adequate identification.

Of course, the last point is always debatable. It would have been preferable if one of the tobacco control policy variables had exhibited both "first-stage" power and "second-stage" exogeneity, allowing for richer evidence of identification. Unfortunately, as is always the case with these sorts of models, the chips fell where they did. Nonetheless, the results as a whole are at the least strongly suggestive of a possibly significant link between cigarette smoking and the probabilities of obesity and overweight.

4   Conclusion

The empirical results presented in this manuscript should serve to stoke further the debate regarding this critical relationship. They suggest that smoking does indeed lower the probability of being overweight or obese, and that failure to control for the potential endogeneity of smoking can lead to substantial underestimation of causal effect of smoking on these outcomes. In the context of the health economics literature, in some sense the game is tied at this point, with Chou et al.

---

[31]We suspected at first that the mass point probabilities were simply failing to move off of their initial values and responded with every computational remedy used in the past of which we are aware. However, the same basic pattern to the estimated probabilities emerged repeatedly, and hence we conclude that this is simply the result. The generally more evenly balanced probabilities and less pronounced mass point values may simply reflect the fact that a far larger proportion of the sample has a BMI of 25 or more than a BMI of 30 or more.

(2004) and the present manuscript finding a significant inverse relationship between smoking and body mass (or key threshold indicators of it) and Rashad (2006) and Gruber and Frakes (2006) recovering no significant link. Further research into this question is clearly warranted.

Even in the absence of other, contradictory voices, the empirical work in this manuscript would certainly not resolve the question of a link between declining cigarette smoking and the surging proportions of Americans who are overweight and obese. To begin with, there is always scope for methodological improvement. For instance, our preferred instrument (cigarette prices) may have passed over-identification tests with the present sample and specifications, but this is admittedly not an uncontroversial instrument (e.g. Gruber and Frakes 2006) regardless of present performance. In any case, richer identification is certainly always preferable and as we develop extensions to the present work we will try to broaden the instrument set. The behavioral pathway considered in this manuscript is also a rather simple one. Future work may more explicitly consider dynamic pathways (allowing examination not only of more complex pathways, but also exploitation of some of the identification sources such models can provide), as well as the possibility of reverse causality (the subject of its own growing literature; see Cawley et al. 2004 a,b).

Nonetheless, the background discussion in this manuscript certainly renders the basic results presented herein plausible on some basic level. For instance, they are in line with many of others on the basic subject from the public health and biomedical literature. Moreover, an approach to the obesity and overweight epidemics that assigns a significant role to declining smoking rates would not necessarily displace the present mainstream approach, which has focused on caloric intake and expenditure. Indeed, introducing cigarette smoking to the story can render these explanations more compelling by identifying a mediating factor in the health production process that explains why generally linear trends to caloric intake and energy expenditure have resulted in such a non-linear course for obesity prevalence.

References

[1] Albanes D, Jones D, Micozzi M, Mattson M. (1987) Associations between smoking and body weight in the US population: analysis of NHANES 11. American Journal of Public Health 1987; 77; 439-444.

[2] Auld M. Smoking, Drinking, and Income. Journal of Human Resources 2005; 40; 2; 505-518.

[3] Caan B, Coates A, Schaefer C, Finkler L, Sternfeld B, Corbett K. Women Gain Weight 1 Year after Smoking Cessation while Dietary Intake Temporarily Increases. Journal of the American Dietic Association 1996; 96; 11; 1150-1155.

[4] Cawley J, Markowitz S, Tauras J. Lighting Up and Slimming Down: the Effects of Body Weight and Cigarette Prices on Adolescent Smoking Initiation. Journal of Health Economics 2004a; 23; 293-311.

[5] Cawley, J, Markowitz S, Tauras J. Body Weight, Cigarette Prices, Youth Access Laws and Adolescent Smoking. Working Paper 2004b.

[6] Chaloupka F, Warner K 2000. The Economics of Smoking. In: Culyer A, Newhouse J (Eds), Handbook of Health Economics, vol. 1B. Elsevier, North Holland, 2000. p. 1540-1627.

[7] Chiolero A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. American Journal of Clinical Nutrition 2008; 87; 4; 801-809.

[8] Chou S, Grossman M, Saffer H. An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System. Journal of Health Economics 2004; 23; 565-87.

[9] Collins L, Cornelius M, Vogel R, Walker J, Stamford B. (1994) Effect of caffeine and/or cigarette smoking on resting energy expenditure. International Journal of Obesity and Related Metabolic Disorders 1994; 18; 8; 5516.

[10] Cutler D, Glaeser E, Shapiro J. Why Have Americans Become More Obese? NBER working paper No. 9446; 2003.

[11] Dallosso H, James W. The role of smoking in the regulation of energy balance. International Journal of Obesity 1984; 8; 4; 36575.

[12] Eisenberg D. Estimating the Effect of Smoking Cessation on Weight Gain: An Instrumental Variables Approach. Health Services Research 2006; 41; 6; 2255-2266.

[13] Flegal K. The Effects of Changes in Smoking Prevalence on Obesity Prevalence in the United States. American Journal of Public Health 2007; 97; 8; 1510-1514.

[14] Froom P, Kristal-Boneh E, Melamed S, Gofer D, Benbassat J, Ribak J. Smoking cessation and body mass index of occupationally active men: the Israeli CORDIS Study. American Journal of Public Health, American Public Health Association 1999; 89; 5; 718722.

[15] Glymour M, DeFries T, Kawachi I, Avendano M. Spousal Smoking and Incidence of First Stroke: The Health and Retirement Study. American Journal of Preventive Medicine 2008; 35; 3; 245-248.

[16] Gruber J, Frakes M. Does Falling Smoking Lead to Rising Obesity? Journal of Health Economics 2006; 25; 183-97.

[17] Heckman J, Singer B. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. Econometrica 1984; 52; 271-320.

[18] Hofstetter A, Schutz Y, Jequier E, Wahren J. Increased 24-hour energy expenditure in cigarette smokers. New England Journal of Medicine 1986; 314; 2; 7982.

[19] Jessen A, Buemann B, Toubro S, Skovgaard I, Astrup A. The appetite-suppressant effect of nicotine is enhanced by caffeine. Diabetes, Obesity and Metabolism 2005; 7; 4; 32733.

[20] Krugman P. A Model of Balance-of-Payments Crises . Journal of Money, Credit and Banking 1979; 11; 3; 311-325.

[21] Mizoue T, Ueda R, Tokui N, Hino Y, Yoshimura T. Body mass decrease after initial gain following smoking cessation. International Journal of Epidemiology 1998; 27; 6; 984989.

[22] Moffatt R, Owens S. Cessation from cigarette smoking: changes in body weight, body composition, resting metabolism, and energy consumption. Metabolism 1991; 40; 5; 46570.

[23] Molarius A, Seidell J, Kuulasmaa K, Dobson A, Sans S. Smoking and relative body weight: an international perspective from the WHO MONICA Project. Journal of Epidemiology and Community Health 1997; 51; 3; 25260.

[24] Mroz T. Discrete Factor Approximations for Use in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome. Journal of Econometrics 1999; 92; 233-274.

[25] Perkins K, Epstein L, Stiller R, Fernstrom M, Sexton J, Jacob R, Solberg R. Acute effects of nicotine on hunger and caloric intake in smokers and nonsmokers. Psychopharmacology 1991; 103; 1; 1039.

[26] Perkins K, Epstein L, Sexton J, Solberg-Kassel R, Stiller R, Jacob R. Effects of nicotine on hunger and eating in male and female smokers. Psychopharmacology 1992; 106; 1; 539.

[27] Philipson T, Posner R. The Long-Run Growth in Obesity as a Function of Technological Change. NBER working paper No. 7423 1999.

[28] Pisinger C, Jorgenson T. Waist circumference and weight following smoking cessation in a general population: The Inter99 study. Preventive Medicine 2007; 44; 4; 290295.

[29] Rashad I. Structural Estimation of Caloric Intake, Exercise, Smoking, and Obesity. NBER working paper No. 11957 2006.

[30] Salant S, Henderson D. Market Anticipation of Government Policy and the Price of Gold. Journal of Political Economy 1978; 86; 627-648.

[31] US Department of Health and Human Services. The Health Benefits of Smoking Cessation: A Report of the Surgeon General. Rockville, MD: Office on Smoking and Health. DHHS Publication CDC 90-8416; 1990.

[32] Ward K, Klesges R, Vander Weg M 2001. Cessation of smoking and body weight. In: Bjrntop P (Ed.) International Textbook of Obesity. Wiley & Sons Ltd: Chichester (UK); 2001. p.32336.
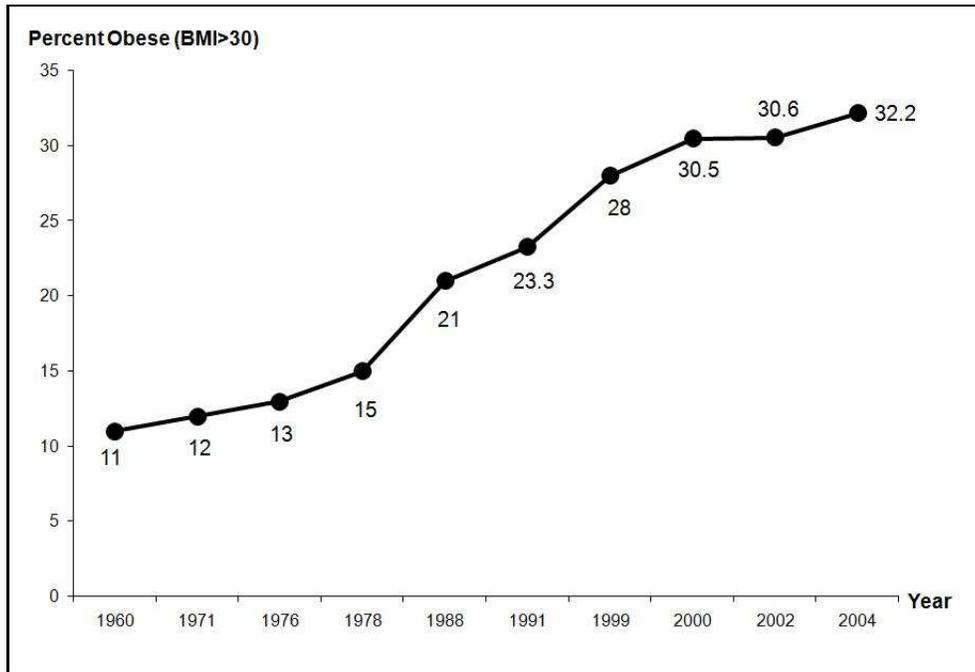
Figure 1: The Evolution of Obesity Prevalence in the United States, 1960-2004

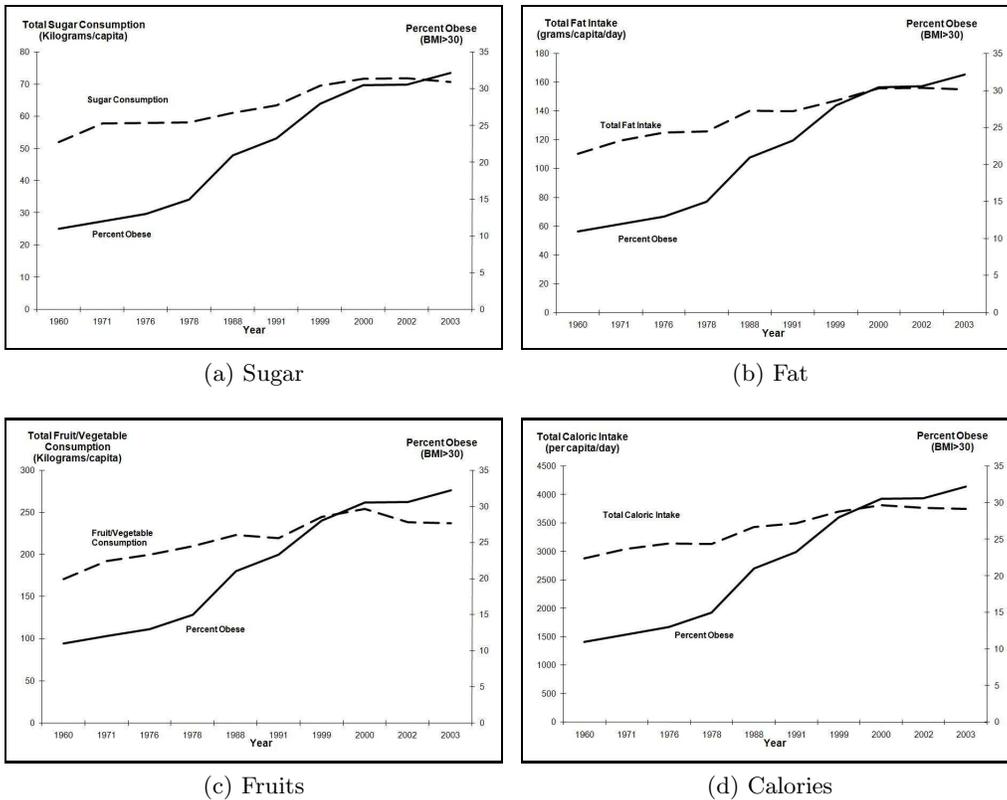(a) Sugar

(b) Fat

(c) Fruits

(d) Calories

Figure 2: Trends in Obesity and Various Nutritional Indicators in the United States, 1960-2004
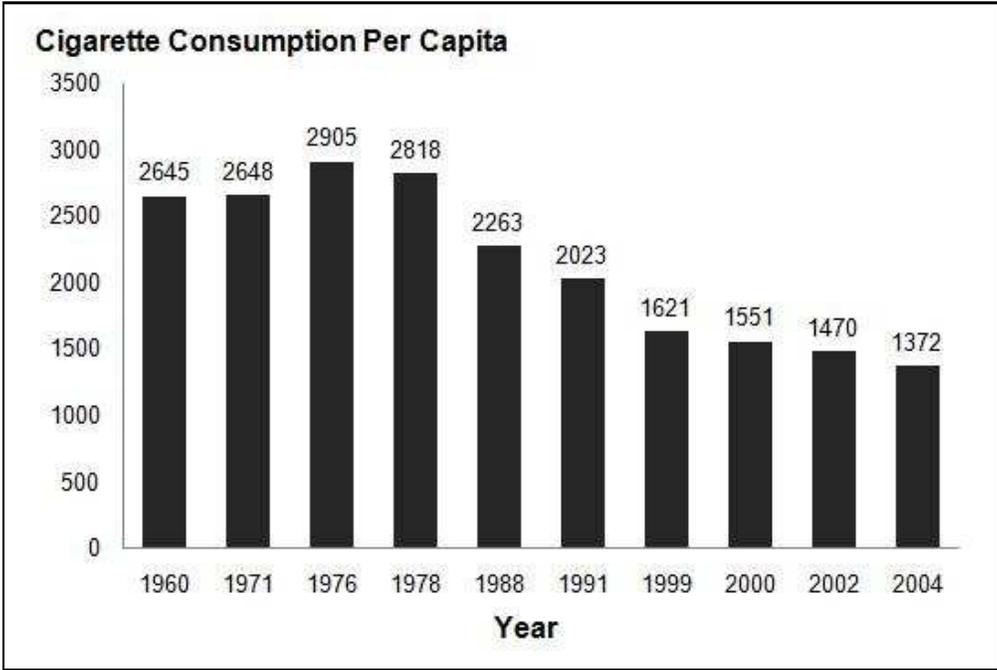
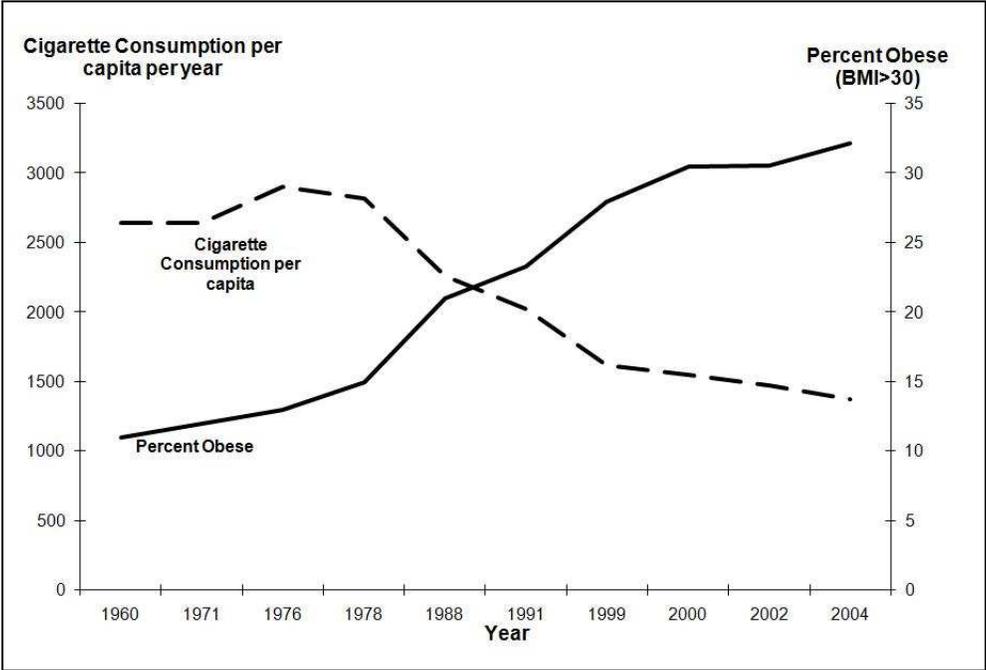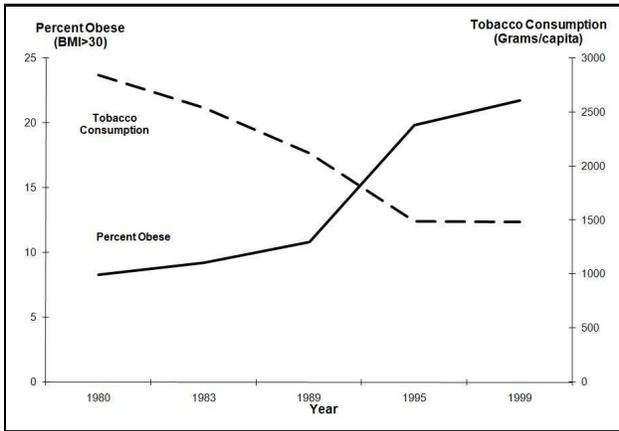Figure 3: The Evolution of Cigarette Consumption in the United States, 1960-2004
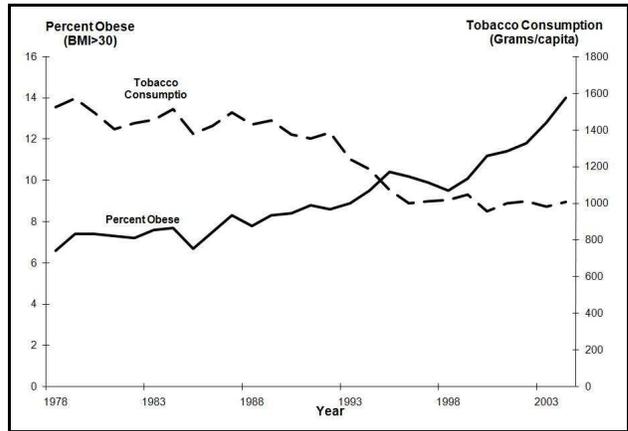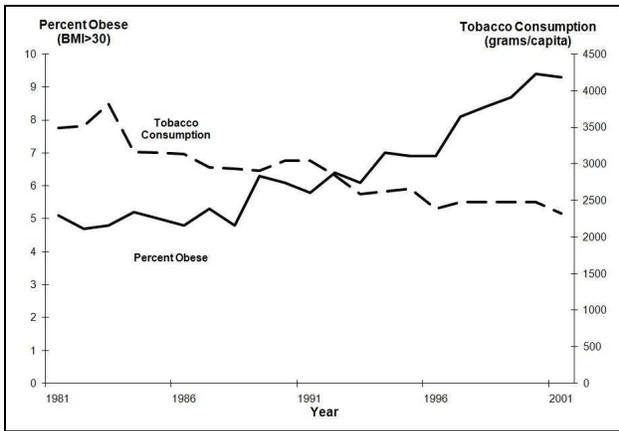

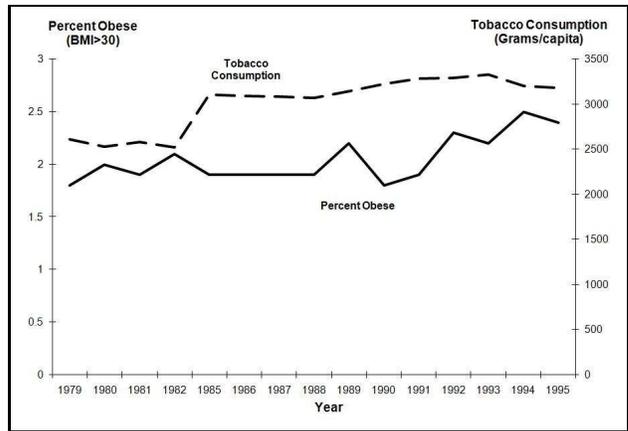Figure 4: The Evolution of Obesity and Smoking Prevalence in the United States, 1960-2004
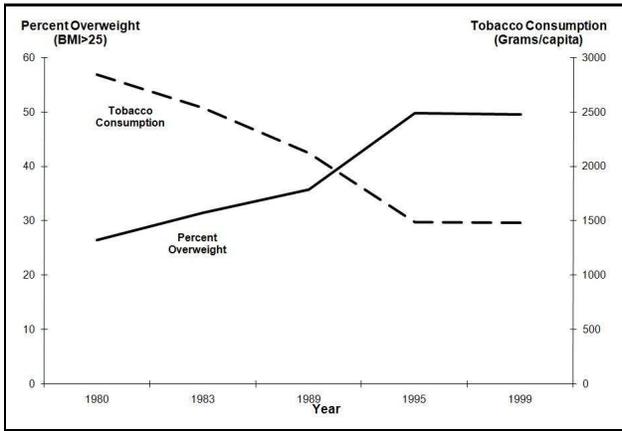
(a) Australia

(b) Finland

(c) Netherlands

(d) Japan

Figure 5: Smoking and Obesity Trends: Select Other OECD Nations

(a) Australia
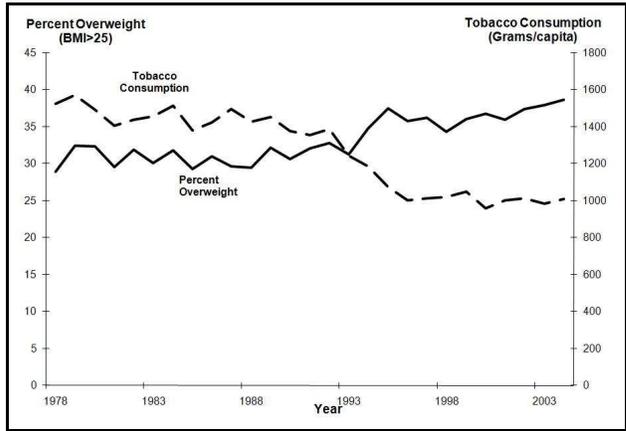
(b) Finland
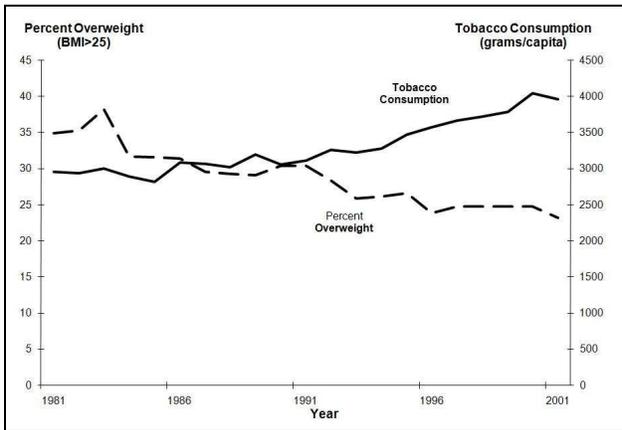
(c) Netherlands
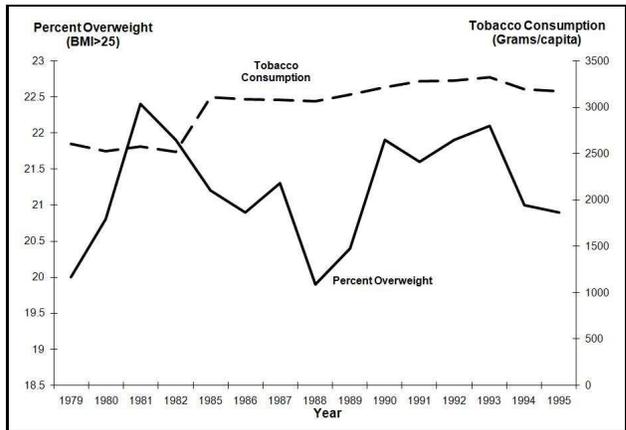
(d) Japan

Figure 6: Smoking and Overweight Trends: Select Other OECD Nations

Figure 7: The Evolution of Obesity Prevalence as Lifestyle Gradually Evolves (Holding Smoking Behavior Constant)
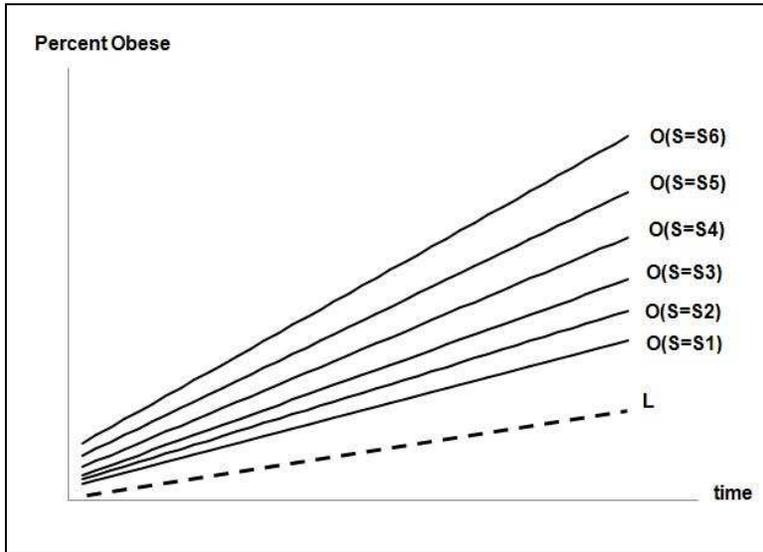
Figure 8: The Evolution of Obesity Prevalence for Various Smoking Levels
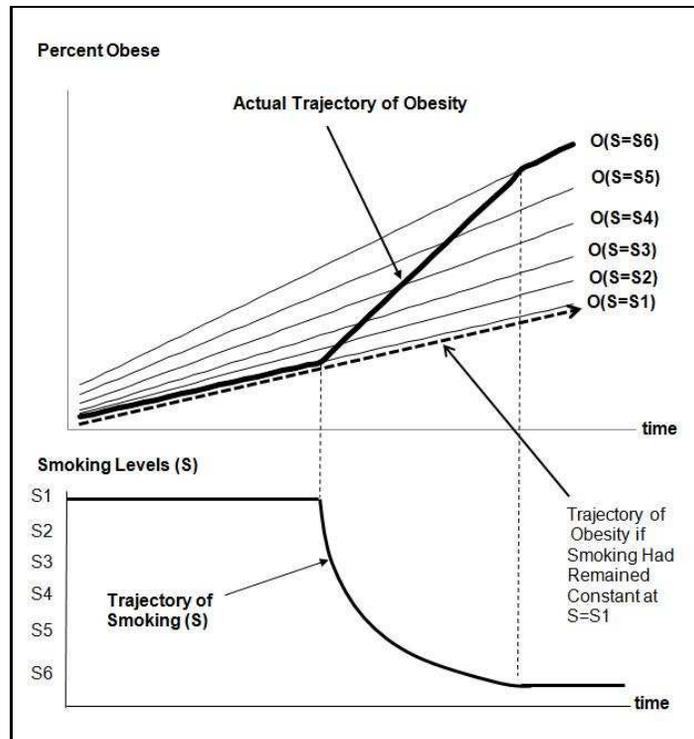
Figure 9: The Evolution of Obesity Prevalence as Smoking Levels FALL
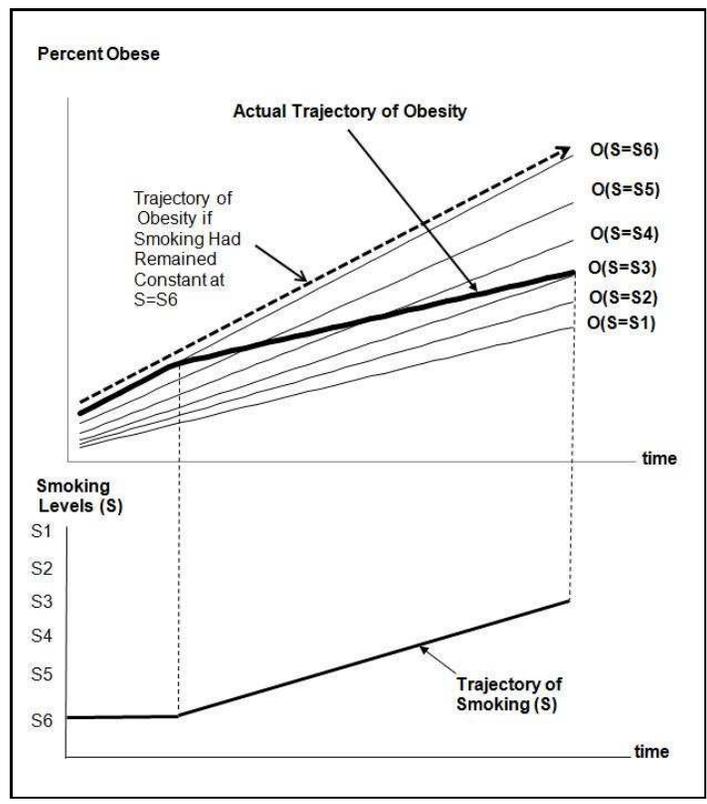
Figure 10: The Evolution of Obesity Prevalence as Smoking Levels RISE

Table 1: Summary Statistics

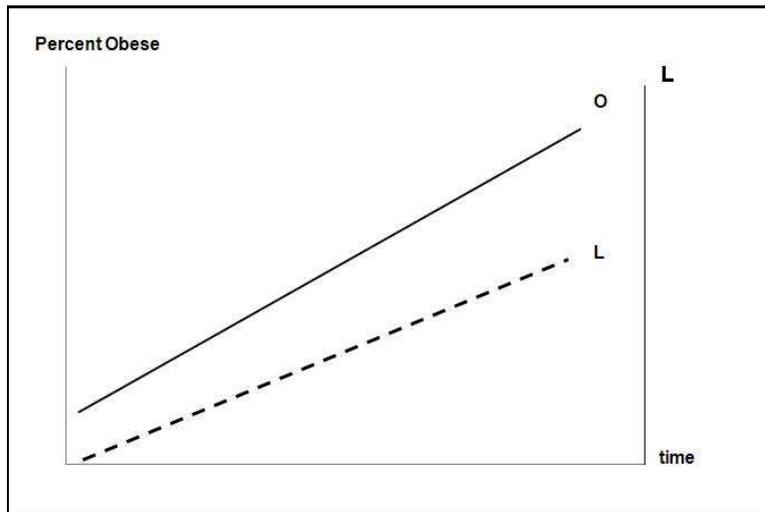| Variable	Mean	Standard Deviation
Obese (%)	24.310	42.896
Overweight (%)	61.676	48.618
Smoke (%)	22.097	41.490
Male (%)	50.477	49.998
Married (%)	69.729	45.943
Age	41.550	11.345
Black (%)	27.831	44.817
Hispanic (%)	4.761	21.293
Education:		
Less than 12 years (%)	14.593	35.304
12 years (%)	33.308	47.132
>12 and < 16 (%)	26.494	44.131
$\geq$ 16 years (%)	25.604	43.645
Annual Income (in 10,000 of $s)	3.658	5.156
Year:		
1999 (%)	24.597	43.067
2001 (%)	25.611	43.649
2003 (%)	25.311	43.479
2005 (%)	24.480	42.997
Cig Price (in $s)	1.946	0.317
Smoking Regulations		
Government (%)	25.026	43.317
Private (%)	44.418	49.688
Restaurant (%)	55.891	49.652
Number of observations	30,584	

Table 2: Effect of Smoking on Obesity Risk: Coefficient Estimates

Variable	Separate Estimation			Joint Estimation With Unobserved Heterogeneity Correction		
Equation:Obese						
Constant	-2.343	(0.229)	***	-3.507	(0.958)	***
Male	0.177	(0.045)	***	0.802	(0.253)	***
Marital	-0.142	(0.044)	***	-0.505	(0.208)	**
Age	5.403	(1.084)	***	15.99	(4.058)	***
Age Squared (/100)	-5.323	(1.210)	***	-16.21	(4.497)	***
Black	0.618	(0.048)	***	1.642	(0.381)	***
Hispanic	0.063	(0.102)		0.049	(0.355)	
High School	-0.142	(0.065)	**	-0.515	(0.273)	*
Some College	-0.202	(0.069)	***	-0.764	(0.294)	***
College Plus	-0.745	(0.081)	***	-1.821	(0.346)	***
Wage	-0.200	(0.089)	**	-0.916	(0.301)	***
2001	0.134	(0.026)	***	0.352	(0.087)	***
2003	0.270	(0.028)	***	0.667	(0.116)	***
2005	0.383	(0.030)	***	0.992	(0.161)	***
Smoke	-0.507	(0.053)	***	-3.571	(1.989)	*
Equation: Smoke						
Constant	-0.899	(0.294)	***	-1.734	(0.782)	**
Male	0.359	(0.049)	***	0.873	(0.254)	***
Marital	-0.681	(0.045)	***	-1.423	(0.311)	***
Age	8.093	(1.306)	***	18.81	(3.807)	***
Age Squared (/100)	-10.50	(1.514)	***	-24.10	(4.588)	***
Black	-0.387	(0.055)	***	-0.906	(0.205)	***
Hispanic	-1.371	(0.127)	***	-3.168	(0.668)	***
High School	-0.693	(0.062)	***	-1.214	(0.229)	***
Some College	-1.103	(0.069)	***	-2.128	(0.388)	***
College Plus	-1.931	(0.087)	***	-4.271	(0.809)	***
Wage	-0.590	(0.108)	***	-2.008	(0.503)	***
2001	0.050	(0.028)	*	0.152	(0.061)	**
2003	0.085	(0.036)	**	0.218	(0.077)	***
2005	0.030	(0.037)		0.091	(0.070)	
Cig Price	-0.140	(0.085)	*	-0.448	(0.188)	**
Government	0.046	(0.056)		0.120	(0.107)	
Private	-0.150	(0.061)	**	-0.329	(0.123)	***
Restaurant	0.052	(0.063)		0.186	(0.123)	
Marginal effect of Smoking on Pr(Obese)	-.083	(0.008)	***	-.174	(0.047)	***

Note: Standard errors in parentheses.

***, **, and *: significance at 1, 5, and 10 percent levels, respectively.

Table 3: Heterogeneity Parameters from Regression of Smoking on Obesity Risk

Point of Support	Probability Weight	Obesity Equation	Smoking Equation	
1	0.220	0.000	0.000	
2	0.135	-280.603	3.979	***
		(fixed)	(0.715)	
3	0.422	-358.460	-0.800	
		(fixed)	(1.314)	
4	0.066	270.880	1.454	**
		(fixed)	(0.718)	
5	0.048	3.311	23.477	***
		(2.183)	(4.872)	
6	0.109	0.520	-16.246	
		(2.885)	(288.5)	

Note: Standard errors in parentheses.

***, **, and *: significance at 1, 5, and 10 percent levels, respectively.

Table 4: Effect of Smoking on Overweight Risk: Coefficient Estimates

Variable	Separate Estimation			Joint Estimation With Unobserved Heterogeneity Correction		
Equation:Overweight						
Constant	-1.798	(0.207)	***	-9.984	(1.942)	***
Male	0.959	(0.040)	***	5.258	(1.182)	***
Marital	-0.051	(0.042)		-0.536	(0.237)	**
Age	7.445	(0.969)	***	34.986	(5.683)	***
Age squared (/100)	-6.659	(1.092)	***	-30.915	(5.459)	***
Black	0.757	(0.048)	***	4.493	(1.035)	***
Hispanic	0.452	(0.091)	***	2.073	(0.635)	***
High School	-0.123	(0.063)	**	-1.440	(0.433)	***
Some College	-0.236	(0.066)	***	-2.104	(0.551)	***
College Plus	-0.696	(0.069)	***	-4.051	(0.800)	***
Wage	0.059	(0.047)		-0.178	(0.106)	*
2001	0.120	(0.022)	***	0.482	(0.122)	***
2003	0.214	(0.025)	***	0.835	(0.192)	***
2005	0.322	(0.027)	***	1.368	(0.259)	***
Smoke	-0.517	(0.046)	***	-4.298	(0.993)	***
Equation: Smoke						
Constant	-0.899	(0.295)	***	-1.966	(0.429)	***
Male	0.359	(0.048)	***	0.475	(0.070)	***
Marital	-0.681	(0.031)	***	-0.806	(0.062)	***
Age	8.093	(1.317)	***	11.013	(1.748)	***
Age squared (/100)	-10.504	(1.528)	***	-14.090	(2.048)	***
Black	-0.387	(0.055)	***	-0.505	(0.076)	***
Hispanic	-1.371	(0.127)	***	-1.751	(0.179)	***
High School	-0.693	(0.063)	***	-0.792	(0.078)	***
Some College	-1.103	(0.069)	***	-1.282	(0.089)	***
College Plus	-1.931	(0.086)	***	-2.449	(0.137)	***
Wage	-0.590	(0.109)	***	-1.021	(0.181)	***
2001	0.050	(0.027)	*	0.088	(0.037)	**
2003	0.085	(0.036)	**	0.132	(0.047)	**
2005	0.030	(0.036)		0.078	(0.047)	*
Cig Price	-0.140	(0.084)	*	-0.227	(0.114)	**
Government	0.046	(0.056)		0.063	(0.077)	
Private	-0.150	(0.060)	**	-0.221	(0.080)	***
Restaurant	0.052	(0.062)		0.110	(0.081)	
Marginal effect of Smoking on Pr(Overweight)	-0.112	(0.010)	***	-0.178	(0.042)	***

Note: Standard errors in parentheses

***, **, and *: significance at 1, 5, and 10 percent levels, respectively.

Table 5: Heterogeneity Parameters from Regression of Smoking on Overweight Risk

Point of Support	Probability Weight	Overweight Equation		Smoking Equation	
1	0.248	0.000		0.000	
2	0.257	4.161 (0.957)	***	-0.013 (0.384)	
3	0.225	-48.338 (13.508)	***	1.259 (0.232)	***
4	0.234	23.481 (7.617)	***	1.481 (0.256)	***
5	0.036	6.383 (1.354)	***	16.102 (2.199)	***

Note: Standard errors in parentheses.
***, **, and *: significance at 1, 5, and 10 percent levels, respectively.