

NBER WORKING PAPER SERIES

SYNTHESIZING ECONOMETRIC EVIDENCE:  
THE CASE OF DEMAND ELASTICITY ESTIMATES

Philip DeCicca  
Donald S. Kenkel

Working Paper 20906  
<http://www.nber.org/papers/w20906>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 2015

We thank the Harvard Center for Risk Analysis for its financial support and the participants in HCRA's October 2013 "Methods for Research Synthesis: A Cross-Disciplinary Approach" workshop for their helpful comments. We also thank the editors of the Special Issue Lisa Robinson and James Hammitt, the Area Editor Warner North, and two anonymous referees. Finally, we thank Lawrence Jin for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Philip DeCicca and Donald S. Kenkel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Synthesizing Econometric Evidence: The Case of Demand Elasticity Estimates  
Philip DeCicca and Donald S. Kenkel  
NBER Working Paper No. 20906  
January 2015  
JEL No. H0,H2,I1,I18

**ABSTRACT**

Econometric estimates of the responsiveness of health-related consumer demand to higher prices are often key ingredients for policy analysis. Drawing on several examples, especially that of cigarette demand, we review the potential advantages and challenges of synthesizing econometric evidence on the price-responsiveness of consumer demand. We argue that the overarching goal of research synthesis in this context is to provide policy-relevant evidence for broad brush conclusions and propose three main criteria to select among research synthesis methods. We also contribute a new empirical exercise that puts the results of previous research synthesis to the test. In particular, we ask whether the “best” consensus estimates of the price-elasticity of smoking help predict trends in smoking from 1995 to 2010. The demographics of the smoking population in our baseline year predict a downward trend in smoking even if cigarette prices remained constant. Average cigarette prices, however, more than doubled in real terms by 2010. We find that the observed declines in smoking over this period are considerably smaller than smoking demographics combined with prior consensus elasticity estimates would predict. Our results suggest that these consensus estimates may have systematically overestimated the price responsiveness of cigarette demand.

Philip DeCicca  
Department of Economics  
422 Kenneth Taylor Hall  
McMaster University  
Hamilton, ON L8S 4M4  
CANADA  
and NBER  
decicca@mcmaster.ca

Donald S. Kenkel  
Department of Policy Analysis and Management  
College of Human Ecology  
Cornell University  
Martha Van Rensselaer Hall  
Ithaca, NY 14853-4401  
and NBER  
dsk10@cornell.edu

## 1. INTRODUCTION

Individual decisions to consume certain goods are an important source of risks to health. The leading behavioral risk factors in the US include tobacco use, poor diet and physical activity, alcohol consumption, and illicit use of drugs.<sup>(1)</sup> One policy approach to reduce these risks is to drive up these goods' prices by imposing special taxes or restricting supply. Consequently, econometric estimates of the responsiveness of consumer demand to higher prices are often key ingredients for risk policy analysis.

Indeed, econometric research helped establish excise taxes on cigarettes and alcoholic beverages as important public policies to reduce health risks. The 2000 Report of the Surgeon General reviews over 50 econometric studies and draws the strong conclusion that: "Substantial increases in the excise taxes on cigarettes would have a considerable impact on the prevalence of smoking and, in the long term, reduce the adverse health effects caused by tobacco."<sup>(2)</sup> The World Health Organization reaches similar conclusions about alcohol taxes: "One of the most effective strategies for reducing consumption of alcohol at the population level is through increasing alcohol prices, usually accomplished by raising alcohol taxes."<sup>(3)</sup> With increased policy interest in obesity, the price-responsiveness of consumer demand for calorie-dense foods, such as fast foods and soft drinks, become policy-relevant as new taxes are considered.<sup>(4,5,6)</sup> The price-responsiveness of the demand for illegal drugs helps predict the effectiveness of supply-side drug policies.<sup>(7)</sup>

Because of the relevance for public policy, researchers have also tried to synthesize the econometric evidence on the price-responsiveness of demand for health-related goods. Econometric studies typically report results in terms of an estimate or range of estimates of the price-elasticity of demand. The price-elasticity of demand equals the percentage change in demand over the percentage change in price. For example, if the price-elasticity of smoking is estimated to be -0.4, it means that a 10 percent increase in the price of cigarettes is predicted to reduce smoking by 4 percent. Usually there are many price-elasticity estimates available from econometric studies that use almost as many different sources of data, specifications, and other research design differences. One approach to synthesize the available evidence is to conduct a narrative review of existing estimates, which typically involves the review authors' judgments about the quality, reliability, and relevance of the studies. An example is Chaloupka and Warner's<sup>(8)</sup> oft-cited review of estimates of the price-elasticity of cigarette demand. More

recently, there have been several important meta-analyses of econometric research on price-elasticities of the demand for health-related goods. Gallet and List <sup>(9)</sup> report a meta-analysis of 523 price-elasticity estimates from 86 empirical studies of the demand for cigarettes. Wagenaar, Salois, and Komro <sup>(10)</sup> report a meta-analysis of 1003 price-elasticity estimates from 112 studies of the demand for alcoholic beverages. Gallet <sup>(11)</sup> reports a meta-analysis of 462 estimates from 42 studies of the price-elasticities of marijuana, cocaine, and heroin.

In this paper, we review the potential advantages and challenges of synthesizing econometric evidence on the price-responsiveness of consumer demand. This paper is part of a Harvard-organized symposium on Methods for Research Synthesis. Like the other Symposium papers, we do not report a new narrative review or meta-analysis. Instead, our goal is to contribute to an inter-disciplinary discussion about how to make research synthesis methods more useful across a range of public policy domains. The Symposium made it clear that the usefulness of different research synthesis methods depends on the nature of the research and how the synthesis results are used in the policy process. Based on our expertise we mainly draw on examples of research on consumer demand for health-related goods. Our analysis of current practice also contributes to academic research on the specific policy question of the effectiveness of higher cigarette prices to reduce smoking. However, most of the discussion is relevant for econometric research on other examples of consumer demand relevant to risk-reduction policies, such as higher gasoline taxes to reduce traffic fatalities.

In section 2 we propose criteria for selecting among research synthesis methods. In section 3 we apply the criteria to evaluate the first step of the research synthesis studies we review: the use of price-elasticity of consumer demand as a summary measure. In section 4 we compare two methods used to synthesize econometric consumer demand research: narrative review and systematic meta-analysis. In section 5 we contribute a new empirical exercise that puts the results of previous research synthesis to the test and asks whether the “best” consensus estimates of the price-elasticity of smoking help predict trends in smoking from 1995 – 2010. Section 6 concludes with a set of recommendations for improving methods used to synthesize econometric research on consumer demand elasticities to provide better guidance for public health risk policy analysis.

## **2. CRITERIA FOR SELECTING AMONG RESEARCH SYNTHESIS METHODS**

The appropriate criteria for selecting among research synthesis methods depend upon how the results will be used. If the goal of the synthesis is to provide insights for academic research, methods such as meta-regression have scientific value because they explore sources of methodological diversity in estimates.<sup>(12)</sup> Our focus is on synthesizing econometric research on consumer demand for health-related goods to provide guidance for public policy. As illustrated by the quotes above from the Surgeon General Report and the WHO, policy makers typically use these econometric research results to make broad brush statements. The overarching concern, then, is to use research synthesis methods that provide the best guidance for such broad brush statements. For example, how should econometric research on cigarette and alcohol demand be synthesized to provide a strong basis for policy makers to decide if higher cigarette and alcohol taxes “have a considerable impact” or are “one of the most effective policies”?

The first criterion we propose is that the synthesis method should provide evidence about the unbiased, cause-and-effect relationship between price and consumer demand. For a single study, this criterion corresponds to whether the research uses a strong design in terms of internal validity. The econometric estimates of the price-elasticities of the demand for cigarettes, alcohol, and illicit drugs synthesized in the meta-analyses cited above are all based on nonexperimental data. In what has been termed the “credibility revolution in empirical economics” there is an increasing focus on the importance of research design in nonexperimental studies.<sup>(13)</sup> The internal validity of research designs lie on a continuum, ranging from the gold standard of randomized experiments, through credible quasi-experimental designs, to weaker designs that yield statistical associations. To use the analogy from randomized experiments, the credibility of a quasi-experimental econometric study hinges on finding a plausible control group. For example, econometric studies of youth cigarette demand that use U.S. cross-sectional data essentially compare the smoking rates in the treatment group of high-tax states with the smoking rates in the control group of low-tax states. Any difference in smoking rates is then attributed to these differences in tax rates. This approach yields biased estimates of the price-elasticity of youth smoking if there are other systematic differences between the treatment and control groups. To use terminology from econometrics, problems such as omitted variables bias, endogeneity bias, and simultaneous equations bias boil down to doubts about the internal validity of the estimate.

Ideally, there would be enough studies with credible quasi-experimental research designs that synthesis methods could simply exclude studies with weaker designs from the analysis. In practice, econometric research on consumer demand relies on research designs of varying strength along the continuum of internal validity. As we discuss in more detail below, this poses a challenge for research synthesis methods to glean some evidence from the weaker studies that yield statistical associations that might, or might not, reflect causation. An additional challenge is to communicate this methodological uncertainty when summarizing the evidence in a research synthesis provided to policy makers.

The second criterion we propose is that the synthesis method should provide evidence about the magnitude and practical significance of the relationship between price and consumer demand. In a single econometric study, the main focus is often on statistical significance: Did the relationship in the data sample occur by chance, or is it estimated precisely enough to reject the null hypothesis of no relationship with a reasonable degree of statistical confidence? McCloskey and Ziliak<sup>(14)</sup> argue that many empirical economists confuse statistical significance with practical significance. They urge empirical economists to pay attention to the scientific question: How large is the estimated effect in terms of the present conversation? The conversation among academic economists might sometimes mainly be about the null hypothesis. For example, evidence that the demand for addictive goods such as cigarettes or heroin responds at all to higher prices sheds light on the relevance of the still-controversial rational addiction model.<sup>(15)</sup> But for the conversation among policy makers about how to discourage smoking or heroin use, what matters is the magnitude of the impact of policy-relevant (e.g., feasible) changes in prices on consumer demand. Regardless of the underlying academic conversation, the research synthesis should contribute to the policy conversation.

We see the standard criterion of precision mainly as it relates to our proposed criterion about magnitude and practical significance. In a single econometric study, the usual practice is to present both a point estimate and the associated statistical confidence interval around the point estimate. However, when the results are summarized it is not uncommon to provide a single price-elasticity estimate. Depending upon the study's statistical precision, this practice might obscure the fact that the study's estimate might not contribute that much to the policy conversation. For example, the 95 percent confidence interval around an imprecisely estimated elasticity might include very small effect sizes, in which case the study's results do not rule of

the possibility that a price increase might do very little to reduce unhealthy consumption. Research synthesis methods clearly need to take into account the statistical precision of the estimates that form the research base, in addition to the more fundamental concerns about the internal validity of the research design and the practical significance of the estimates.

The third criterion we propose is that the synthesis method should provide evidence about the external validity of the estimated relationship between price and consumer demand. External validity is a standard criterion for research design and asks: Can the results be extrapolated or generalized? As Heckman<sup>(16)</sup> emphasizes, internally valid estimates of cause-and-effect relationships in a given environment are not enough to provide guidance for policy: “policy evaluation is conducted with an eye toward the future and toward informing decisions about new policies and applications of old policies to new environments.” For example, DeCicca, Kenkel and Mathios<sup>(17)</sup> caution that it would be problematic to use their estimates that modest hikes in cigarette taxes had little impact on youth smoking to predict the impact of much larger tax hikes. As another example, it would also be problematic to use estimates of the price-responsiveness of youth smoking in the U.S. to predict the impact of higher taxes in a much different environment such as a low-income country. Approaches used in economic health policy evaluation vary in sophistication and range from simple extrapolations to dynamic population simulation models (the approach used in section 5) to structural econometric models<sup>(16, 18)</sup>. The structural econometric approach emphasizes external validity and uses economic theory in the form of an explicit model “to extrapolate observed responses to environmental changes to predict responses to other not-yet-observed changes.”<sup>(19)</sup> Regardless of the policy evaluation approach, research synthesis provides more useful guidance when it sheds light on the relevance of the results for a range of policy changes (e.g. tax hike sizes) and environments (e.g. countries).

### **3. THE USEFULNESS OF THE PRICE-ELASTICITY OF DEMAND AS A SUMMARY MEASURE**

Typically, the first step in research synthesis is to express the results of different studies in a common metric in order to obtain an “effect size.” Econometric research on the size of the effect of prices on consumer demand is typically summarized by the price-elasticity of demand. To set the stage for a discussion of the usefulness of this common metric, in this section we first review some basic economic concepts. We then apply our criteria from section 2 to discuss the advantages and disadvantages of the price-elasticity as a summary measure.

A large body of empirical econometric research provides estimates of consumer demand curves for a wide range of goods and services. Conceptually, the demand curve shows the quantity of a good consumers purchase at different prices, holding other demand determinants constant. The slope of this curve reflects the price-responsiveness of demand. Econometric studies use multivariate linear regression and its extensions to estimate the determinants of consumer demand, including the slope of demand with respect to price. The price-elasticity normalizes this slope in percentage terms. The total price elasticity captures consumer responses on both the extensive and intensive margins. When the data are rich enough, econometric studies can separately estimate price-responsiveness at both margins. For example, many empirical health economics studies estimate the price-elasticity of smoking participation (the extensive margin) and conditional on being a smoker, the price-elasticity of the demand for the number of cigarettes smoked per day (the intensive margin). For addictive goods it is also important to distinguish the price-elasticities of initiation versus cessation. DeCicca, Kenkel and Mathios<sup>(20)</sup> emphasize that smoking participation at a point in time reflects a series of past decisions about smoking initiation and cessation and show that the elasticity of smoking participation is a weighted average of the elasticity of smoking initiation and the elasticity of smoking cessation.

In many applications, a key advantage of the price-elasticity of demand is that it satisfies our second criterion and provides evidence about the magnitude and practical significance of the relationship between price and consumer demand. The normalization in percentage terms is often useful to judge policy relevance. For example, if the estimated price elasticity is small in absolute value, say  $-0.01$ , it means that a 100 percent increase in price would only reduce demand by one percent. It should usually be straight-forward for a policy maker to judge if a price increase of 100 percent or larger is policy-relevant in terms of feasibility, unintended consequences, and so on. In principle, the policy maker should also be able to judge whether a reduction in demand by one percent is large relative to what might be achieved by other policies, but of course this requires evidence about the other policies in question. Whether a price-elasticity is large enough to be policy-relevant will differ across policy contexts, but in general the magnitude of the price-elasticity will often be useful to help policy makers draw broad brush guidance from econometric research.



A disadvantage to using the price-elasticity in research synthesis is that a single “effect size” cannot in general summarize the results of an econometric study of consumer demand. The demand curve or function is a schedule that shows the quantity demanded at various different prices. Neither the marginal effect of price nor the price-elasticity is necessarily constant along the demand curve, except for the mutually exclusive special cases of a linear demand curve or a constant elasticity demand curve.<sup>1</sup> Using a single price-elasticity to summarize the demand curve is analogous to using the effect of a specific drug treatment dosage to summarize the entire dose-response curve. In the context of drug treatments, the solution is to conduct additional randomized controlled trials (RCTs) that provide evidence on the relative effects of different dosages compared to each other, as well as when compared to the control group. In that context, the RCTs generally consider a fairly limited range of alternative dosages.

A single price-elasticity estimate often provides very limited information about alternative “policy dosages” because the range of policies can be quite broad. For example, the cigarette tax rates chosen by different states within the US currently vary by more than an order of magnitude, from \$0.17 per pack in Missouri to \$4.35 per pack in New York.<sup>(21)</sup> Because policy makers often look at other countries’ experiences, it should be noted that policies can vary even more widely across countries. The challenge for research synthesis is thus not only to summarize evidence about the price-consumption relationship in the observed data, but also to consider policy makers’ interest in extrapolating beyond the range of policies observed in any given country. This illustrates the importance of our third criterion that the research synthesis needs to consider the external validity of the price-elasticity estimate: Is the estimate generalizable to the context of the new public policy being considered?

Another challenge is that in practice policy-makers may overlook important distinctions between different types of price-elasticity estimates included in a common research synthesis. As noted above, price-responses along the extensive and intensive margins of consumer demand correspond to two types of elasticities: a participation elasticity of demand and a conditional demand elasticity of demand. For addictive goods it is also important to distinguish short-run versus long-run elasticities, and initiation versus cessation elasticities. The different elasticity

---

<sup>1</sup> The cases are mutually exclusive because of the normalization involved in the elasticity formula. Although the slope showing the marginal effect of price is constant along a linear demand curve, the price-elasticity is different at each point on a linear demand curve. Conversely, the slope or marginal effect of price is different at each point along a constant elasticity demand curve.

measures represent very different concepts and summarize very different types of consumer responses. For example: the short-run conditional elasticity of cigarette demand reflects the impact of higher prices on current smokers' decisions to smoke fewer cigarettes per day; while the long-run smoking participation elasticity reflects the accumulated impact of higher prices over several years on youth decisions to start smoking and current smokers' decisions to quit smoking. These differences again raise the issue of external validity, because it can be inappropriate to generalize the results for one type of elasticity that reflects one type of consumer behavior to other types of elasticities that reflect different consumer behaviors.

The use, or perhaps mis-use, of research synthesis results in public health discussions about smoking policies illustrates the potential pitfalls of overlooking key distinctions between the different elasticity concepts. Policies that decrease smoking participation, either by discouraging initiation or by encouraging cessation, substantially reduce the major health risks due to smoking. Policies that encourage continuing smokers to smoke fewer cigarettes per day yield much smaller health risk reductions, particularly if smokers who cut down on the number of cigarettes smoke each cigarette more intensely<sup>(22)</sup>. When considering cigarette taxation as a public health policy, discussions should focus on the price-elasticity of smoking participation. However, the research syntheses mainly summarize estimates of the total price-elasticity of the demand for cigarettes, which reflects both the participation elasticity and the conditional demand elasticity of cigarettes smoked per day by continuing smokers. The syntheses reflect the limitations of the econometric research reviewed. For example, in the meta-analysis by Gallet and List,<sup>(9)</sup> over 400 of the 523 price-elasticity estimates summarized are from studies that use data aggregated at the national or state/province level. Because they use aggregate data these studies cannot separate how much of the total price-elasticity is due to changes in participation versus changes in the number of cigarettes smoker per day by continuing smokers. In the narrative review in the 2000 Report of the Surgeon General,<sup>(2)</sup> almost half of the 55 reviewed studies use aggregate data, and most of the studies that report participation elasticities are limited to youth. At the time of these research synthesis studies there were only a very limited number of estimates of the impact of higher prices on adult smoking participation elasticity and virtually no estimates on the elasticity of cessation. The research base of evidence on whether higher cigarette prices reduce smoking participation is much thinner than it might appear from public health policy discussions of this evidence.

#### 4. COMPARING METHODS TO SYNTHESIZE ECONOMETRIC EVIDENCE ON PRICE-ELASTICITIES

In this section we compare two methods commonly used to synthesize econometric research: a narrative review of existing estimates; or a systematic meta-analysis of estimates. We focus on their relative strengths and weaknesses in terms of our proposed criteria for selecting synthesis methods.

In principle, either a narrative review or a meta-analysis can address our first criterion that the synthesis method should provide evidence about the unbiased cause-and-effect price-consumption relationship. Because of the sometimes subtle research design issues involved, however, the method of narrative review has advantages in meeting this criterion. As noted above, econometric research on the effects of price on health-related consumption relies almost exclusively on non-experimental data. The identification of cause-and-effect relationships in non-experimental data is arguably the central challenge of applied econometrics. Meyer<sup>(23)</sup> emphasizes: “If one cannot experimentally control the variation one is using [to identify a key parameter of interest], one should understand its source.” Modern applied econometric research often involves lengthy discussions of a variety of sources of evidence to shed light on whether the source of variation provides “clean identification” corresponding to a credible quasi-experiment.<sup>(13)</sup> While narrative reviews can summarize each study’s discussion of the source of identification, the nature of these discussions makes it harder to incorporate in a systematic meta-analysis.

In the practice of research synthesis of econometric estimates of cigarette price-elasticities, neither narrative reviews nor meta-analysis meet our first criterion very well. For example, neither the prominent narrative review by Chaloupka and Warner<sup>(8)</sup> nor the meta-analysis by Gallet and List<sup>(9)</sup> discuss the identification of the price-consumption relationship in the sense of whether the studies reviewed use a credible research design. Both studies discuss a related identification issue: the problem that prices and quantities are simultaneously determined by the forces of supply and demand. For example, Gallet and List explore whether studies that use the method of two-stage least squares to address this problem yield systematically different price-elasticity estimates. However, this does not address the question of whether the underlying source of variation in prices, typically variation in taxes over time or across jurisdictions, is a credible source of identification. For example, policy endogeneity suggests that taxes and

consumption may be correlated for reasons other than the cause-and-effect demand relationship between prices and consumption. Economic theory and empirical evidence predict that failing to control for policy endogeneity problem leads to bias in the price-elasticity estimate, although of course the true value will never be known. The failure to discuss the difficulty of identifying cause-and-effect relationships in price-elasticity studies might be explained by the fact that the reviews, and in particular the studies reviewed, were completed during the early days of, and sometimes pre-date, the “credibility revolution” in applied econometrics. However, the gap is still apparent in more recent narrative reviews of cigarette price-elasticities.<sup>(24, 25, 26)</sup>

A weakness of a narrative review compared to a meta-analysis is the potential for reviewer bias to creep in, which in turn implies that the results of the research synthesis might not provide unbiased evidence. In contrast, the econometric meta-analyses aim to provide a systematic summary of estimates, and are at least partly motivated by the goal to avoid potential pitfalls in narrative reviews, such as “the subjective decision of the reviewer to discount certain studies or characteristics...”<sup>(9)</sup> In the academic research debate on the price-elasticity of youth smoking, the simple fact that different narrative reviews of the same body of research reach different conclusions is worrisome. For example, while previous narrative reviews led to a claimed consensus that youth smoking is much more price-responsive than adult smoking, DeCicca, Kenkel and Mathios<sup>(17)</sup> characterize the empirical evidence as “mixed.” The controversy has not been resolved over time. A recent narrative review by Bader et al<sup>(25)</sup> identifies 21 studies of the impact of higher prices on smoking initiation and concludes from these studies that “There was strong evidence that raising cigarette prices through increased taxes is a more effective tobacco control policy measure for reducing smoking behavior among youth...” In contrast, the recent narrative review by Guindon<sup>(26)</sup> concludes: “...this review concludes that the evidence is too limited to make any conclusive statements about the impact of tobacco prices or taxes on smoking onset.”

We reach similar conclusions about the potential and practice of research synthesis methods in terms of our second proposed criterion about the magnitude and practical, policy-relevant significance of the evidence. In principle, either a narrative review or a meta-analysis can meet this criterion. In current practice as applied to research on cigarette demand price-elasticities, the methods used to date have a mixed record. Although we do not see precision as a criterion per se, the precision of the estimates underlying either a narrative review or a meta-analysis has

important implications for whether the evidence provides useful guidance about magnitude and practical significance. Here the method of meta-analysis might seem to have advantages. However it is important to keep in mind that the goal of a meta-analysis of price-elasticity estimates is different than the goal of meta-analysis in many clinical and public health applications. In many applications of meta-analysis pooling together results from studies with low statistical power improves precision.<sup>(27)</sup> In contrast, a lack of statistical power and the need for more statistically precise estimates are not the main motivation for many meta-analyses of demand elasticity estimates. In their meta-analysis of cigarette demand studies Gallet and List<sup>(9)</sup> suggest that due to differences in econometric techniques and data “it is not surprising to find significant differences in elasticities across studies....while the mean price elasticity across the 86 studies is -0.48...the standard deviation is quite large (0.43), as is the range of estimates (-3.12 to 1.41).” Taken at face value, this wide range of estimates provides little guidance for policy: the extreme estimates imply that a 10 percent price increase might reduce cigarette demand by over 30 percent or increase cigarette demand by as much as 14 percent. The narrative reviews conclude that a much narrower range is justified, but this conclusion rests on the subjective judgements of the reviewers. To date, the existing meta-analyses and narrative reviews do not provide hard evidence about the likely precision of their summary cigarette price-elasticity estimates.

The practical relevance of a price-elasticity estimate also depends on whether the policy maker has other effective tools. Narrative reviews, such as the review by Chaloupka and Warner,<sup>(8)</sup> often discuss other policies, although they do not necessarily provide head-to-head comparisons of policy alternatives. Meta-analyses of price-elasticities do not systematically discuss policy alternatives.

Narrative reviews and meta-analyses often fail to explicitly discuss the external validity of the estimates for the context relevant to current policy. At the same time, the syntheses often provide information that sheds valuable light on external validity, i.e. on the generalizability of the results reviewed. Narrative reviews and meta-analyses typically report the time period and country studied. When used in policy discussions, the conventional wisdom about generalizability is that is most appropriate to consider recent estimates from the same country, or at least countries that are similar in terms of income levels and so on. Research syntheses also

often report estimates by population sub-groups, such as youth and young adults, which again are potentially useful when generalizing the results.

## **5. EVALUATING A META-ANALYSIS OF DEMAND ELASTICITIES**

In this section, we put the results of previous research synthesis to the test and ask: Do the “best” consensus estimates of the price-elasticity of smoking demand help predict what happened to smoking in the US from 1995 – 2010? Over this time period, after adjusting for inflation the average cigarette price more than doubled from \$2.57 per pack to \$5.55 (in constant 2010 dollars). The 1998 Master Settlement Agreement with the tobacco industry, several federal tax hikes and over 100 state tax hikes contributed to the price increase. This sharp increase in prices provides a policy experiment to evaluate whether the 2000 Surgeon General Report drew the correct lesson from previous research synthesis that higher prices have a “considerable impact on the prevalence of smoking.” The exercise is relevant to the three criteria we propose in section 2 to select among research synthesis methods. First, if previous research synthesis yields an unbiased estimate of the cause-and-effect relationship of higher prices on smoking, the estimate should be useful for the prediction exercise. Second, the exercise also sheds light on the magnitude and practical significance of the relationship between cigarette prices and smoking. Third, the exercise directly addresses the question of the external validity of the consensus estimate and whether it is generalizable to predict the impact of the price hikes that occurred from 1995 – 2010.

To conduct the exercise, we need to describe the counter-factual of what smoking would have been, in the counter-factual world where cigarette prices stayed constant at the 1995 level. Fortunately, a series of studies by Mendez and colleagues essentially describe this counter-factual. Mendez, Warner and Courant<sup>(28)</sup> develop a population dynamic model to predict smoking prevalence rates in the US. The model uses estimates of birth rates, mortality rates, and age-specific smoking initiation and cessation rates to simulate the number of adults smoking in the US over time. The results of their simulation model show that “the demographics of smoking imply that prevalence will inexorably continue to decline over the next several decades, even without any intensified efforts aimed at tobacco control.” The model is calibrated with actual data through 1995, when adult smoking prevalence in the US was 24.7 percent. The model predicted 2005 smoking prevalence (as measured in the National Health Interview

Survey) exactly at 20.9 percent; the model predicted 2010 smoking prevalence at 19.9 percent, 0.6 percentage points higher than the NHIS2010 estimate.<sup>(29)</sup>

The key for our exercise is that the predictions from Mendez et al do not incorporate any adjustments to capture the effects of higher cigarette prices or other tobacco control policies. The simulation model assumes the continuation of the age-specific rates of smoking initiation and cessation as of 1995. There are no assumptions about any time trends in individuals' smoking behaviors; the dynamics over time reflect only changes in the demographic mix of the smoking population. As such, the predictions from the simulation model describe what would have happened to US smoking rates in the counter-factual world where cigarette prices stayed constant.

We have built a dynamic population model that closely replicates Mendez et al's predictions for smoking prevalence. We then extend the approach to predict the average number of cigarettes smoked per day by current smokers and to predict total cigarette consumption. We use data from the 1992 – 1995 cycles of the Tobacco Use Supplements to the Current Population Survey (TUS-CPS) on the average number of cigarettes smoked per day by different birth cohorts of adult smokers. For our baseline predictions we do not incorporate any adjustments to capture the effects of higher prices. Instead, we assume that each birth cohort of smokers continues to smoke the same number of cigarettes per day as they did in 1995. The population-average number of cigarettes per day changes over time solely due to population dynamics, as the heavier-smoking older birth cohorts die or quit smoking. We combine our predictions of smoking prevalence and the number of cigarettes smoked per day to predict total cigarette consumption. We adjust the self-reported TUS-CPS data on cigarettes smoked per day to match aggregate data on total cigarette consumption.

To capture the effects of higher cigarette prices, we follow the general strategy used by Harris<sup>(30)</sup> to predict future smoking trends and use a simple constant elasticity model of smoking::

$$\ln(Q) = \alpha + \beta T + \delta \ln(P)$$

where  $\ln(Q)$  is the natural logarithm of smoking ( $Q$ ),  $T$  is a time trend measured in years since 1995, and  $\ln(P)$  is the natural logarithm of the inflation-adjusted cigarette price ( $P$ ). Given this form of the demand equation, the coefficient  $\delta$  is the price-elasticity of smoking demand. We

develop three versions of the constant elasticity model of smoking demand to predict: smoking prevalence, the number of cigarettes smoked per day, and total cigarette consumption.

We build into the constant elasticity prediction models estimates of  $\delta$  based on existing research syntheses of smoking price-elasticities. As discussed above, the broadest demand elasticity measure is the elasticity of total cigarette consumption, which captures consumer responses on both the extensive and intensive margins. The narrative review by Chaloupka and Warner<sup>(8)</sup> and the meta-analysis by Gallet and List<sup>(9)</sup> both suggest a synthesis estimate of the total price-elasticity of around -0.4. Although neither study provides a confidence interval for their synthesis estimate, based on the range of estimates discussed we explore a range for  $\delta$  from -0.2 to -0.6. The research syntheses provide less direct evidence on the elasticity of smoking participation or prevalence and the elasticity of cigarettes smoked per day conditional upon participation.<sup>2</sup> However, the participation elasticity plus the conditional demand elasticity must sum to the elasticity of total cigarette consumption. Chaloupka and Warner<sup>(8)</sup> suggest that the price-elasticity of smoking participation is about half the total price-elasticity. We therefore assume that the elasticity of smoking prevalence and the elasticity of conditional demand range from -0.1 to -0.3 (consistent with their sum ranging from -0.2 to -0.6).

For each version of the constant-elasticity prediction model, we calibrate  $\alpha$  to fit the 1995 data, and we calibrate the time trend  $\beta$  to fit the predictions from our population dynamic model. In other words, our calibrated time trend captures the demographic factors built into the population dynamic model. Although the model is more complicated, a linear time trend fits the predictions reasonably well. We re-emphasize that our approach does not assume any time trend in individual smoking behaviors; the time trend is solely due to the changing demographics of the smoking population.

The Figures below shows the results of the exercise. As in Warner and Mendez,<sup>(29)</sup> if we hold P constant at its 1995 value Figure 1 shows that we slightly over-estimate the actual prevalence of smoking in 2010. Over most of the period the predictions are very close to observed smoking prevalence. Using the actual history of cigarette prices and the consensus price elasticity of smoking participation ( $\delta = -0.2$ ), smoking prevalence is predicted to fall to

---

<sup>2</sup> In econometric studies of smoking, the term “smoking participation” is used to refer to whether an individual smokers or not. The term “smoking prevalence rate” refers to the number of current smokers as a fraction of the population. Thus, the smoking participation elasticity can be used to predict changes in the smoking prevalence rate.



about 17 percent in 2010, compared to the observed rate of 19.3 percent. Put differently, while observed smoking prevalence fell by 5.4 percentage points (from 24.7 percent to 19.3 percent), taking into account the price increase we predict it should have fallen by about an additional 2 percentage points. The observed rate falls outside the range of predictions based on the plausible range of price elasticity estimates of -0.1 to -0.3.

Figure 2 shows that if we hold P constant at its 1995 value, we over-estimate the average number of cigarettes smoked per day. Based on the demographics of the smoking population our model predicts that the average number of cigarettes smoked per day falls from about 20 cigarettes in 1995 to about 17 cigarettes in 2010. The observed average was about 15 cigarettes. Using the actual history of cigarette prices, the model that incorporates the consensus price elasticity of conditional demand ( $\delta = -0.2$ ) predicts the average observed in 2010. However, in most years prior to 2010 the observed averages are closer to the prediction from the model that incorporates a price elasticity estimate of  $\delta = -0.1$ .

Figure 3 shows that if we hold P constant at its 1995 value, we over-estimate total cigarette consumption in 2010. The demographics predict that total cigarette consumption falls from about 475,000 million cigarettes to about 350,000 million cigarettes. The demographics predict observed cigarette consumption pretty well from 1995 through 2007. Using the actual history of cigarette prices, the model that incorporates a total elasticity of demand of -0.2 predicts observed consumption in 2010. However, for the years before 2009 this model under-estimates total cigarette consumption.

Our results confirm and extend the analysis by Warner and Mendez.<sup>(29)</sup> Based on the changing demographics of the smoking population we predict downward trends from 1995 – 2010: in the prevalence of smoking, in the average number of cigarettes smoked per day by smokers, and in total cigarette consumption. The observed decline in smoking prevalence from 1995 to 2010 corresponds very closely to what would be predicted for a counter-factual world with no price increase. The observed declines in cigarettes smoked per day and total cigarette consumption are larger than the predicted declines in the counter-factual world, especially for the last few years of the prediction exercise. When we use the actual history of cigarette prices and incorporate price-elasticity estimates in the range suggested by existing research syntheses, our models predict that smoking prevalence and total cigarette consumption should have fallen by substantially more than what was observed from 1995 to 2010.

In addition to the demographic dynamics, there have been multiple new tobacco control efforts since 1995 including: increased availability of smoking cessation products; widespread bans on smoking in worksites, restaurants, and other public places; state tobacco control programs including telephone quitlines and mass media anti-smoking campaigns; and national mass media anti-smoking campaigns. Quantifying the impacts of these policies is beyond the scope of our empirical exercise. Addressing this limitation would reinforce our main findings: including the influence of other tobacco control policies would cause our models to even more substantially under-predict smoking prevalence and total cigarette consumption/ over-predict the drop in smoking that should have been attributable to the effect of the observed price increase.

In sum, the results of our quantitative exercise are suggestive evidence that previous efforts to synthesize econometric research on the price-smoking relationship do not perform well. It is difficult to reconcile the actual histories of smoking and cigarette prices with the consensus estimates of the price-elasticity of smoking participation. The exercise is somewhat more supportive of the consensus estimates of the price-elasticity of the demand for cigarettes smoked per day and the price-elasticity of total consumption. Above in section 4 we suggest that the research syntheses behind the consensus price-elasticity estimates fail to fully meet reasonable criteria, so the poor performance is not surprising. Many of the existing price-elasticity estimates might be systematically biased away from the true cause-and-effect relationship between prices and smoking. The failure to account for internal validity of the estimates could explain the poor performance of the research syntheses. Alternatively, even if the estimates being synthesized were internally valid for the earlier time period, they might lack external validity when extrapolated to the later period 1995 – 2010. For example, Warner<sup>(30)</sup> suggests that more of today's smokers are hard-core smokers who are unable or unwilling to quit.

## **6. CONCLUSIONS AND RECOMMENDATIONS**

Econometric estimates of the price-responsiveness of consumer demand for health-related goods are often used to make broad brush conclusions about the effectiveness of public policies to reduce health risks. In this context, we propose three criteria for research synthesis methods: the synthesis method should provide evidence about the unbiased cause-and-effect relationship between price and consumer demand; the method should provide evidence about the magnitude and the practical significance of that relationship; and the

method should provide evidence about the external validity of the estimated relationship. We argue that to date, synthesis of econometric research on consumer demand for health-related goods fails in several ways to meet our proposed criteria. Limitations include: the focus on a summary price-elasticity measure rather than tracing out the price-demand relationship; lack of consideration of research design issues needed for credible estimates of the cause-and-effect relationship; inadequate consideration of the magnitude and practical or policy significance of the estimated relationships; and lack of consideration of factors that determine whether the estimated relationship will provide useful guidance to predict the impact of new policies or of old policies in new environments.

We also contribute a new evaluation of how well existing syntheses of the price elasticity of cigarette demand perform. We extend a relatively simple dynamic population model of cigarette demand to incorporate the impact of corresponding consensus elasticities. Our findings imply that the consensus estimates from previous syntheses do not perform well in explaining the observed smoking trends over time. Perhaps most striking, our results imply that the consensus price-elasticity of participation of  $-0.2$  is too large in absolute value; our model, which incorporates this estimate, consistently predicts a lower smoking prevalence than is observed over a fifteen year period. There are several possible reasons for the poor performance of the consensus-based predictions. On one hand, the culprit could be that prior research syntheses relied on estimates that lacked internal validity. This seems very plausible because many estimates are from studies that pre-date the “credibility revolution” that has emerged in empirical economics. More recent studies that apply modern econometric methods to data with a time component (panel data or repeated cross sectional data) over periods with multiple tax hikes that provide credible quasi-experiments should improve internal validity. That said, even if the studies used by prior syntheses were reasonably internally valid, it might be the case that the consensus estimates they generate do not achieve external validity, perhaps due to the changing composition of smokers over time.

Although we have emphasized challenges and limitations of research syntheses of demand elasticity estimates, we believe more work or research synthesis methods to address these limitations will be productive and important. We therefore conclude this paper with a set of recommendations for future research synthesis estimates.

Our critique of the use of price elasticity as a measure of price sensitivity implies that any method of research synthesis such as a meta-analysis should control for research decisions such as which demographic group is studied, what type of elasticity is estimated (e.g. short-versus long-run, or participation versus conditional demand), the time period of the data, and so on. Including such control variables in a meta-analysis is a first step that recognizes that these research decisions might have important effects on the price-elasticity estimates. However, taking this first step does not “fix” the estimates or address the more difficult questions about which estimates provide the most useful guidance for policy.

While we recognize authors will continue to report price-elasticities, we suggest that research synthesis should also use the marginal effects of various price increases implied by their coefficient estimates.<sup>3</sup> While this sort of representation has its own limitations (e.g., potential non-linearity in the impact of price), we believe it more accurately conveys price sensitivity and allows for more valid cross-group comparisons as well as comparisons over time. By mapping out the estimated price-consumption relationship over the range of prices observed in the data, this approach better meets our criterion that research synthesis should provide policy-relevant evidence about the magnitude and practical significance of the econometric results, and begins to address the issue of external validity.

Research synthesis should take into account the credibility of the research design of studies that yield estimates of the effect of higher prices on consumer demand for health-related goods. If the implicit control group in a quasi-experimental study is flawed, the concern is not precision, but bias. Moreover, the number of studies that reach similar estimates does not help reduce bias, if the studies rely on similarly flawed research designs. One approach to resolve these problems is to rely on strict eligibility criteria for inclusion in the meta-analysis. Just as non-experimental studies are not eligible to be included in a meta-analysis of clinical trials, econometric studies that use flawed research designs could be excluded from a meta-analysis of price elasticity estimates. At the least, the meta-analysis should test for systematic differences in the estimates from studies that rely on different research designs.

A more ambitious agenda for future research is to explore the potential of a structural utility-theoretic approach to the meta-analysis of demand elasticities. Environmental

---

<sup>3</sup> Often, such results are presented as the impact of an X cent increase in cigarette price on smoking behavior.

economists have developed and applied this approach to the problem of benefits transfer, i.e. the use of pre-existing measures of benefits to conduct a cost-benefit analysis of a new but similar policy. The approach, also known as preference calibration, uses a fully specified structural theoretical model of consumers' preferences over the non-market good being valued (e.g. air quality) versus other goods and services. Tightly linking pre-existing benefit measures to the structural framework identifies the model parameters that describe consumer preferences, for example the parameters of the consumer utility function that define the marginal rate of substitution between some aspect of environmental quality and other goods and services. Applications of this approach include structural meta-analyses of estimates of the value of water quality improvements<sup>(32)</sup>, the value of statistical life<sup>(33)</sup> and the value of preventing beach erosion<sup>(34)</sup>.

The application of the structural approach to the meta-analysis of demand elasticities faces somewhat different challenges. While meta-analysis for environmental benefits transfer focuses on identifying preference parameters, Chetty<sup>(35)</sup> shows that preference parameters are not always needed for policy analysis. In particular, he shows that high-level elasticities are sufficient statistics for analysis of policies such as tax hikes. Nevertheless, a more structural approach could help synthesize existing price-elasticity estimates. A structural meta-analysis of smoking elasticities could begin by specifying a version of the rational addiction model<sup>(36)</sup> that includes the decisions to start and quit smoking<sup>(20)</sup>. The structural model will impose restrictions on how different price-elasticities fit together, including the relationship between short-run and long-run elasticities and the relationships between the elasticities of participation, initiation, and cessation. The study by Gallet and List<sup>(9)</sup> includes some of these considerations so it begins to approximate what has been termed a weak structural utility-theoretic approach<sup>(37)</sup> to the meta-analysis of cigarette price elasticities. A more structural approach to the meta-analysis of estimates from studies with strong internal validity could address all of our criteria for research synthesis methods.

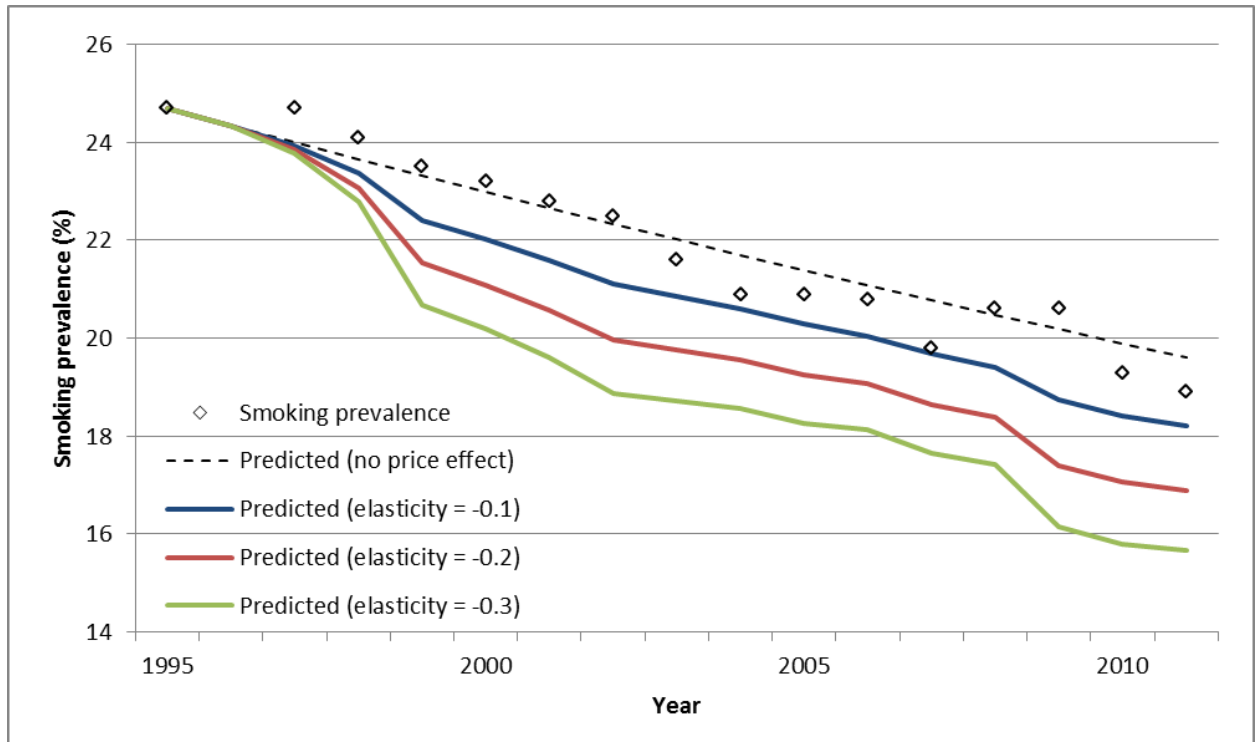
## REFERENCES

1. Mokdad AH, Marks J S., Stroup, D F, Gerberding J L. Actual causes of death in the United States: 2000. *JAMA* 2004; 291 (10): 1238-1245.
2. U.S. Department of Health and Human Services. Reducing Tobacco Use: A Report of the Surgeon General. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. 2000
3. World Health Organization. Global Status Report on Alcohol and Health. WHO Press, Geneva, Switzerland. 2011
4. Chou, S-Y, Grossman M, Saffer, H. An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics* 2004; 23 (3): 565 - 87.
5. Fletcher J, Friswold D E, Teft N. Are soft drink taxes an effective mechanism for reducing obesity? *Journal of Policy Analysis and Management* 2011; 30 (3): 655-662.
6. Nghiem, N, Wilson N, Genc M, and Blakely T. Understanding price elasticities to inform public health research and intervention studies: Key issues. *American Journal of Public Health* 2013; 103 (11): 1954-1961.
7. Manski C F, Pepper J V, Petrie C V, editors. Informing America's Policy on Illegal Drugs. National Academy Press, Washington DC. 2001.
8. Chaloupka F J, Warner K E. The economics of smoking. In Anthony Culyer and Joseph P. Newhouse, editors, *Handbook of Health Economics (Volume 1B)*, pp. 1539 - 1627). Amsterdam: North Holland. 2000.
9. Gallet C A, List J A. Cigarette demand: a meta-analysis of elasticities. *Health Economics* 2003; 12 (10): 821 - 836.
10. Wagenaar A C, Salois M J, Komro, K A. Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. *Addiction* 2009; 104 (2): 179 - 190.
11. Gallet C A. Can price get the monkey off our back? A meta-analysis of illicit drug demand." *Health Economics* 2013; 23 (1): 55-68.
12. Thompson S G, Higgins J P T. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 2002; 21: 1559 – 1573.
13. Angrist J D, Pischke J-S. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives* 2010; 24 (2): 3 – 30.
14. McCloskey D N, Ziliak S T. The standard error of regressions. *Journal of Economic Literature* 1996; 34: 97-114.
15. Becker G S, Murphy K M. A theory of rational addiction. *Journal of Political Economy* 1988; 96 (4): 675 700.
16. Heckman J J. Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature* 2010; 48: 356 – 398.
17. DeCicca P, Kenkel D S, Mathios A D. Putting out the fires: Will higher taxes reduce the onset of youth smoking? *Journal of Political Economy* 2002; 110 (1): 144-169.
18. Heckman J J, Vytlacil E. Structural equations, treatment effects and econometric policy evaluation." *Econometrica* 2005; 73 (3): 669 – 738.

19. Nevo A, Whinston M D. Taking the dogma out of econometrics: Structural modeling and credible inference.” *Journal of Economic Perspectives* 2010; 24 (2): 69-82.
20. DeCicca, P, Kenkel D S, Mathios A D. Cigarette taxes and the transition from youth to adult smoking: Smoking initiation, cessation, and participation. *Journal of Health Economics* 2008; 27 (4): 904-917.
21. Federation of Tax Administrators (2013). [Http://www.taxadmin.org](http://www.taxadmin.org).
22. Adda J, Cornaglia F). Taxes, cigarette consumption, and smoking intensity. *American Economic Review* 2006; 96 (4): 1013-1028.
23. Meyer B D. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics* 1995; 13 (2): 151 – 161.
24. World Health Organization. IARC Handbooks of Cancer Prevention, Tobacco Control, Vol. 14: Effectiveness of Tax and Price Policies for Tobacco Control. Lyon, France 2011.
25. Bader P, Boisclair D, Ferrence R. Effects of tobacco taxation and pricing on smoking behavior in high risk populations: A knowledge synthesis. *International Journal of Environmental Research and Public Health* 2011; 8: 4118 – 4139.
26. Guindon, G E. The impact of tobacco prices on smoking onset: A methodological review. *Tobacco Control* 2013; 23 (e5): 97-111.
27. Petitti, D B. *Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. New York: Oxford University Press. 1994
28. Mendez D, Warner K E, Courant P N. Has smoking cessation ceased? Expected trends in the prevalence of smoking in the United States. *American Journal of Epidemiology* 1998; 148 (3): 249 – 258.
29. Warner K E, Mendez D. Accuracy and importance of projections from a dynamic simulation model of smoking prevalence in the United States. *American Journal of Public Health* 2012; 102 (11): 2045 - 2048.
30. Harris J E. Prepared Statement Before the Senate Democratic Task Force on Tobacco. Sen. Kent Conrad (D-ND), Chairman, Washington DC, October 21, 1997.
31. Warner K E. Charting the science of the future: where tobacco-control research must go. *American Journal of Preventive Medicine* 2007; 33 (6): S314-S317.
32. Smith V K, van Houtven G L, Pattanayak S K. Benefit transfer via preference calibration: ‘Prudential Algebra’ for policy. *Land Economics* 2002; 78 (1): 132 – 152.
33. Smith V K, Pattanayak S K, van Houtven G L. Structural benefit transfer; An example using VSL estimates. *Ecological Economics* 2006; 60: 361 – 371.
34. Van Houtven G L, Poulos C. Valuing welfare impacts of beach erosion: An application of the structural benefit transfer method.” *American Journal of Agricultural Economics* 2009; 91 (5): 1343-1350.
35. Becker G, Grossman M, Murphy K. Economic analysis of cigarette addiction. *American Economic Review* 1994; 4: 396-418
36. Chetty, R. Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. *Annual Review of Economics* 2009; 1: 451–488.
37. Bergstrom J C, Taylor L. Using meta-analysis for benefits transfer: Theory and practice. *Ecological Economics* 2006; 60: 351 – 360.

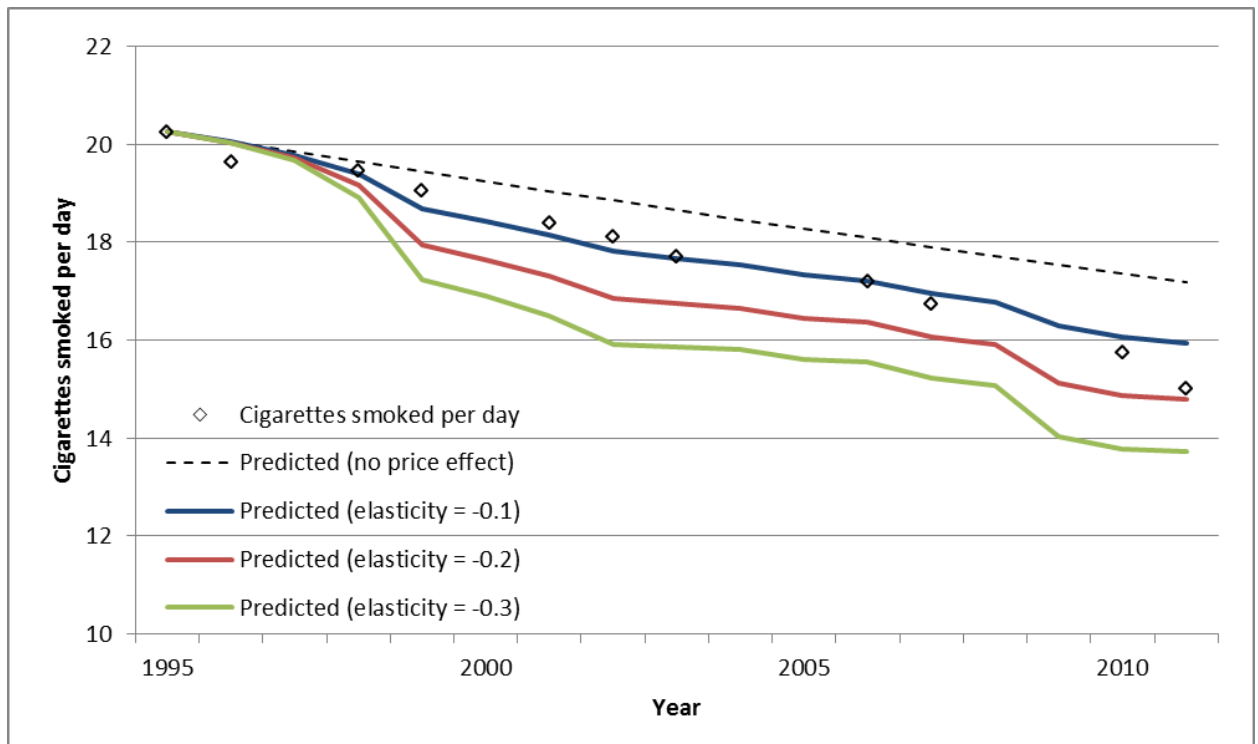
**FIGURE 1: Using Consensus Elasticity Estimates to Predict Trends in Smoking**

**Participation**





**FIGURE 2: Using Consensus Elasticity Estimates to Predict Trends in Cigarettes Smoked Per Day**



**FIGURE 3: Using Consensus Elasticity Estimates to Predict Trends in Total Cigarette Consumption**

