McMaster University
DigitalCommons@McMaster

Open Access Dissertations and Theses | Open Dissertations and Theses

4-1-2012

ESSAYS IN HEALTH ECONOMICS AND ROBUST ECONOMETRICS

Cong Li
McMaster University, lic28@mcmaster.ca

Follow this and additional works at: http://digitalcommons.mcmaster.ca/opendissertations

Part of the Econometrics Commons, and the Health Economics Commons

Recommended Citation
Li, Cong, "ESSAYS IN HEALTH ECONOMICS AND ROBUST ECONOMETRICS" (2012). *Open Access Dissertations and Theses.* Paper 6846.

This Thesis is brought to you for free and open access by the Open Dissertations and Theses at DigitalCommons@McMaster. It has been accepted for inclusion in Open Access Dissertations and Theses by an authorized administrator of DigitalCommons@McMaster. For more information, please contact scom@mcmaster.ca.

ESSAYS IN HEALTH ECONOMICS AND ROBUST ECONOMETRICS

By

CONG LI, B.Sc., M.A.

A Thesis submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree of Doctor of Philosophy

McMaster University

© Copyright by Cong Li, November 2011

DOCTOR OF PHILOSOPHY (2011) McMaster University
(Economics) Hamilton, Ontario

TITLE: Essays in Health Economics and Robust Econometrics

AUTHOR: Cong Li, B.Sc. (Hebei Normal University), M.A. (University of Guelph)

SUPERVISORS: Professor Philip DeCicca
Professor Jeremiah Hurley
Professor Lonnie Magee

NUMBER OF PAGES: xi, 127

ABSTRACT

Formal economic analysis of tobacco products dates back to the middle of the 20th century. At the beginning, most of the research was done by the tobacco industry itself. Later, research interests switched to the public health perspective following the publication of British and American reports on smoking and health. Many papers in the medical literature consistently demonstrate that smoking significantly damages health and that the cigarette tax is a popular policy tool to try to reduce smoking. My first two essays focus on the issue of the cigarette excise tax and smoking. In the first essay, we analyze a possible tax avoidance behaviour since much research finds that cigarette tax increases have not reduced smoking as much as predicted by prior work on the price sensitivity of cigarette smoking. The tax avoidance behavior we analyze is pack versus carton purchasing behaviour. The avoidance is related to the fact that a smoker can save much money by purchasing cartons in lieu of single packs, since carton purchasing is associated with a substantial quantity discount. Overall, we do not find any evidence of switching to carton purchasing in response to higher taxes, but find some evidence of pack purchasing instead. We then do subgroup analyses and find that smokers who intend to quit switch to packs in response to tax increases, consistent with the behavioral economic literature, while smokers who do not want to quit smoking systematically switch to cartons when taxes increase, consistent with the public health literature, which warns of this sort of behavioral response. Our results suggest that it may be very interesting to study whether

other tax avoidance behaviors also differ by smokers' intentions to quit and whether anti-smoking policies similarly have differential effects. In the second essay, we investigate whether the smoking participation of relatively older smokers is sensitive to cigarette tax increases. The consensus in the literature is that older smokers are not price-responsive. Medical research, however, suggests quitting smoking even at an older age can still generate substantial health benefit for the old smokers. We study this question by using the most recent large tax increases across US. We first find that the old smokers have greater desire to quit when tax increases. Moreover, we find that their actual smoking participation rate is also significantly reduced in response to tax increases. These results run contrary to most of the literature and may suggest that tax increases generate substantial health benefits for older smokers. The last essay includes both theory and application of the nonparametric kernel method. In this essay, we extend the model specification test proposed by Fan et al. (2006) to the extent that we also smooth the discrete dependent variable. The null distribution of the test statistic is derived and we also show that the test is consistent even when the null hypothesis fails to hold. Finally, a Monte Carlo simulation study shows that by smoothing the categorical dependent variable, our test enjoys substantial power gains.

ACKNOWLEDGEMENT

My first, and most earnest, acknowledgment must go to my supervisor, Professor Philip DeCicca. I appreciate all his contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating. The joy and enthusiasm he has for his research was contagious and motivational for me, even during tough times in the PhD pursuit. In every sense, none of this work would have been possible without him.

I would like to thank other members of my committee, Professor Jeremiah Hurley and Professor Lonnie Magee, for their helpful comments on my thesis and valuable advice during my whole PhD career. Their dedication, encouragement, support, and hard work constantly inspired me to better myself and aim higher. Special thanks to Professor Jeffrey Racine and Professor John Leach. Nearly six years ago, a conversation and lunch with them started me on the path I traveled at McMaster. Many thanks also go to many faculty and staff in McMaster, Professor Marc-Andre Letendre, Professor Alok Johri, Professor Sungjin Han, Professor Paul Contoyannis who have helped me a lot when I was taking courses during the first two years of my PhD studies, and our graduate secretary, Jan Martens who has helped me every day when I was there.

My most heartfelt gratitude is given to my wife Wei Pan. She has worked diligently and successfully to create a wonderful life for me in Hamilton. Her patience, support, encouragement, and companionship have turned my journey through graduate school into a pleasure. For all that, she has my everlasting love.

Last but not least, I would like to express my deepest gratitude to my parents, my father Ning Li and my mother Yuru Jiao. Thanks for bringing me to the world and creating the best living and studying environment for me. Your cultivation and education have benefited me and will benefit me the whole life. None of my achievement would be possible, without your unvarying love and support. Thank you.

PREFACE

The essays in Chapter 2 and Chapter 3 are co-authored with Professor Philip DeCicca. The essay in Chapter 4 is coauthored with Professor Jeffrey Racine.

Table of Contents

Abstract…….iii

Acknowledgement……..……v

Preface…….……..vii

Table of Contents……...…….viii

List of Tables……xi

Chapter 1 Introduction……...……...1

References…….…….8

Chapter 2 Do Higher Cigarette Taxes Lead Smokers to Lose (Self) Control? ……9

2.1 Introduction…….…….9

2.2 Background…….……...11

2.2.1 Why do smokers buy packs rather than cartons? …….……12

2.2.2 How might higher taxes affect the pack vs. carton choice? …….14

2.2.3 Some simple economics of the impact of taxes on the pack versus carton choice……15

2.3 Data……..……..16

2.3.1 Measures of carton or pack buying……...……17

2.3.2 Cigarette taxes……...……18

2.3.3 Other independent variables and useful variables…….19

2.3.4 Analysis sample……...…….19

2.4 Empirical strategy……..20

2.5 Estimates……...…….22

2.5.1 Main results……..……..22

2.5.2 Subgroup analyses…….25

2.5.3 Digging deeper into competing explanations……27

2.5.4 A simple simulation exercise……..……..30

2.6 Conclusions……33

References……..……..36

Tables……...……38

Chapter 3 Investigating the Price Sensitivity of Older Smokers…….……50

3.1 Introduction……..…….……50

3.2 Background……..……..52

3.2.1 Previous research on youth and young adult smoking behavior...…55

3.2.2 Previous research on older adults smoking behavior……...…….57

3.2.3 The impact of smoking cessation on health……..……59

3.3 Data……61

3.3.1 Measures of smoking participation and intentions to quit smoking……....62

3.3.2 Cigarette excise taxes……...……63

3.3.3 Other independent variables……..64

3.3.4 Analysis sample……65

3.4 Empirical strategy……..65

3.5 Estimates……68

3.5.1 Main results……...68

3.5.2 Heterogeneous tax effects……70

3.6 Conclusion……73

References……75

Tables……...78

Appendix……..90

Chapter 4 A Smooth Nonparametric Conditional Density Test for Categorical Responses……..104

4.1 Introduction……..104

4.2 A Smooth Nonparametric Conditional Density Test……...108

4.2.1 Parametric bootstrap Test……110

4.3 Monte Carlo Results……111

4.4 Conclusion……...113

References……..115

Tables……116

Appendix……118

Chapter 5 Conclusion……124

List of Tables

Chapter 2

Table 1. Percentage of smokers who usually buy cigarettes by the carton……38
Table 2. Monthly state real per-pack cigarette taxes denominated in February 2003 dollars……..39
Table 3. Summary statistics……...41
Table 4. Comparisons of real per-pack price ratio and passthrough rate ratio…..43
Table 5. Estimated effect of real cigarette tax on carton purchase……44
Table 6. Estimated effect of real cigarette tax on carton purchase by smoking status……45
Table 7. Estimated effect of real cigarette tax on carton purchase by income groups……...46
Table 8. Estimated effect of real cigarette tax on quitting intentions and purchasing behavior……47
Table 9. Simulation results of cigarette tax effects on carton purchasing……49

Chapter 3

Table 1. Descriptive statistics for intention to quit and participation by year…...78
Table 2. Monthly state real per-pack cigarette taxes denominated in February 2003 dollars……..79
Table 3. Summary statistics for years from 2003 to 2007……81
Table 4. Estimates of the impact of cigarette excise taxes on smoking participation……82
Table 5. Estimates of the impact of cigarette excise taxes on smoking participation, by personal income……..84
Table 6. Estimates of the impact of cigarette excise taxes on smoking participation, by education……86
Table 7. Estimates of the impact of cigarette excise taxes on smoking participation, by gender……..88

Chapter 4

Table 1. Size of the tests……..116
Table 2. Power of the tests……...117

Chapter 1

Introduction

As one of the oldest crops, tobacco has long history in North America. Traditionally, it was used medicinally to induce calmness, for pain relief and so on. However, following the arrival of the Europeans, tobacco became a recreational drug and increasingly popular in the commercial realm. As a result, a huge amount of tobacco products was produced and it fostered the economy for the southern United States where tobacco growth flourished.

Formal economic analysis of the tobacco industry dates back to the middle of the 20th century. At that time research was done by the industry itself, and focused on marketing and profits. Later, following the publication of seminal British and American reports on smoking and health, the focus of economic research on

smoking has shifted toward the analysis of the potential public health damage associated with tobacco. As reported by the World Health Organization (WHO), tobacco use is the leading cause of preventable diseases (heart attacks, strokes, chronic obstructive pulmonary disease (COPD), emphysema, and lung cancer) and deaths; in particular, WHO estimates that in recent years, it causes 5.4 million deaths per year.

Among tobacco products, cigarettes account for 95 percent of total sales. Cigarette smoking not only affects smokers themselves, but may also affect nearby non-smokers. In order to combat the rapid growth of tobacco products and related harm to health, many governments have implemented a variety of anti-smoking policies to discourage smoking and protect non-smokers from exposure to second-hand smoke. Such policies mainly include increasing cigarette prices via higher taxes, restriction or bans on cigarette advertising, development and dissemination information on the health consequences of smoking, counter-advertising, limiting smoking in public places, and legal restriction on youth access to tobacco products.

Among anti-smoking policies, substantial tax-induced price increases are often suggested in legislative proposals and are hotly debated in the economic literature. Since price is likely to be endogenous and a cigarette excise tax comprises an important component of cigarette price, many studies rely on the spatial variation of taxes to explore their impact on smoking behaviour. The first two chapters of my thesis address this issue by examining issues related to the

impact of cigarette excise taxes on smokers' behaviors.

In the first essay, we investigate the impact of cigarette tax increases on smokers' purchasing behaviours. In particular, we examine whether smokers switch to carton purchasing, in lieu of more standard single pack purchasing, in response to tax increases. Cigarette taxation has garnered interest by public health officials, because a higher cigarette tax may induce smokers to quit smoking. However, researchers in this field have found that the cigarette tax is not working as well as anticipated (e.g., many smokers do not quit smoking even when the cigarette tax is very high). Therefore, tobacco researchers have started to explore the reasons for this apparent lack of price sensitivity. Studies in the public health literature suggest that a crucial reason is that smokers can avoid high cigarette taxes. For example, they can cross borders into the neighbouring areas which offer lower prices or they can purchase discount brand cigarette instead of premium brands. Our paper is placed into the literature on tax avoidance behavior by focusing on pack versus carton purchasing behaviour. Cigarettes are commonly sold in two ways, in single pack of 20 cigarettes, but also in cartons which include 10 packs of 20 cigarettes each, for a total of 200 cigarettes. In general, the per-pack price associated with carton purchasing is much lower than that of a single pack purchase, so smokers can avoid the high prices by purchasing cartons. Although carton purchasing involves this quantity discount, which can be substantial, roughly two-thirds of smokers forgo this discount and buy single packs. As suggested in the behavioural economics literature, smokers may view

the single-pack purchasing as a self-control device either to limit the amount smoked or to aid in smoking cessation. Put differently, smokers claim they buy packs because they want to quit smoking. So, if higher prices induce a greater desire to quit, we might see more single pack purchasing when taxes increase. Given the different implied predictions of public health literature and behavior economic literature, the impact of cigarette tax increases on carton versus single pack purchasing is an empirical question. We empirically investigate the impact of higher cigarette taxes on smokers' choice of carton vs. pack using the data from Tobacco Use Supplements of Current Population Survey (TUS-CPS). In addition to testing the overall effect, we also analyze the purchasing behavior changes in response to tax increases by subgroups. These data can help us understand potential heterogeneity in the effect of taxation and may provide important implications for future studies and policy considerations.

In the second essay, we examine another important implication of cigarette excise taxes—smoking participation. As we have discussed, cigarette taxation has been viewed as an important policy tool to discourage smoking, so many studies in the literature analyze the price-responsiveness of smokers. Many studies find evidence that the smoking behavior of youth and young adults are sensitive to tax increases, while relatively few studies find such pattern for older smokers. In this chapter, we exploit large cigarette tax increases between 2003 and 2007 to reevaluate the price sensitivity of older smokers. As discussed in DeCicca et al. (2008), price participation elasticities can be explained as weighted average of

initiation and cessation elasticities. With respect to youth, smoking initiation dominates smoking participation, while for older adults, the group that we study, it is likely that any reduction in smoking participation represents quitting behavior since very few people start smoking even after the age of twenty-five. We begin by examining the impact of cigarette taxes on smokers' intentions to quit smoking and then study the impact of tax increases on actual smoking participation. A common difficulty existing in this literature is that state level cigarette taxes may be endogenous to state level attitudes towards smoking. In response to this identification difficulty, we estimate a two-way fixed model using repeated cross-sectional data from the TUS-CPS to control for the time-invariant heterogeneity across states. As a robustness check, we also employ a measure of state anti-smoking sentiment (SASS) in the model[1]. The results of this chapter may have important policy implications, since medical studies suggest that quitting smoking even at older ages can still generate substantial health benefits.

The last essay of my thesis includes the theory and application of the nonparametric econometric method. Compared with the parametric models, the nonparametric method relaxes some crucial assumptions such as the distribution of the data and the specification of the functional form in regression models, so it is flexible and more robust econometric method. Its newly developed ability of handing both continuous and categorical data makes it a useful method for a broad range of empirical questions. This chapter contributes to the literature by

[1] SASS is borrowed from DeCicca et al. (2008)

generalizing this method. In the field of nonparametric econometrics, the optimal data-driven bandwidth selection (i.e., least squares cross-validation method) is of crucial importance in nonparametric estimation and the testing of associated econometric models. The least squares cross validation method has many advantages such as automatically removing irrelevant covariates. When there exist discrete variables, it avoids sample splitting and often leads to substantial efficiency gains versus the conventional nonparametric frequency method. In this chapter, we generalize the nonparametric conditional density specification test proposed by Fan, Li, and Min (2006) to the extent that we also smooth the discrete dependent variable. We establish the asymptotic null distribution of the test statistic with data-driven selected stochastic smoothing parameters. We also show that the test is consistent when the null hypothesis fails to hold. Monte Carlo simulations are employed to evaluate the finite-sample performance of this method. With respect to application, this method can be applied to the parametric binary choice models, multinomial response models, ordered models and count models.

The thesis is organized as follows. In Chapter 2, we study pack versus carton purchasing behaviors in response to tax increases for all smokers and smokers in different subgroups. Chapter 3 employs a two-way fixed model to estimate the impact of cigarette excise tax increases on older smokers' intentions to quit smoking and actual smoking participation. Chapter 4 focuses on the issue of nonparametric model specification test in that we generalize the conditional

density test proposed by Fan et al. (2006) to the extent of smoothing categorical dependent variable as well. The last chapter is a brief conclusion.

References

WHO Report on the global tobacco epidemic. World Health Organization (2008). "Tobacco is the single most preventable cause of death in the world today".

DeCicca, P., D. Kenkel, and A. Mathios (2008), "Cigarette Taxes and The Transition from Youth to Adult Smoking: Smoking initiation, cessation, and participation", *Journal of Health Economics* 27: 904-917.

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

Fan, Y., Q. Li and I. Min (2006), "A nonparametric bootstrap test of conditional distributions", *Econometric Theory* 22: 587-613.

Chapter 2

Do Higher Cigarette Taxes Lead Smokers to Lose (Self) Control?

2.1 Introduction

Despite a substantial quantity discount, most smokers purchase cigarettes in single packs, rather than by the carton. According to surveys, many smokers view this as a self-control mechanism, intended either to limit the amount smoked each day or to aid in smoking cessation. Relatively little research investigates smokers' choice between buying cigarettes by the carton or in single packs. In this paper, we investigate the impact of higher cigarette excise taxes on carton versus pack purchasing behavior. While research in public health suggests higher taxes induce

smokers to purchase cartons, to at least partially avoid the higher tax-induced price by realizing the quantity discount, such price increases might result in greater pack purchasing if they induce a greater desire to quit. We first examine the economics of the carton versus pack purchase, demonstrating the choice is a function of relative prices per pack and relative tax pass-through rates. We then empirically investigate the impact of higher cigarette excise taxes on smokers' choice of purchasing cigarettes by cartons versus single packs using data from the U.S. Tobacco Use Supplements of the Current Population Surveys (TUS-CPS) from the 2003 and 2006-2007 waves. Overall, we find no systematic evidence that higher cigarette taxes lead smokers to a greater likelihood of carton purchasing. Instead, our main estimates imply smokers substitute towards pack purchasing when taxes increase, and this does not appear to be driven by resource constraints as there is no discernible pattern in the relevant estimate when allowed to vary by household income. Similarly, there is no apparent pattern across smoker type with respect to the whether or not one smokes everyday and whether or not one is a heavy or moderate smoker. However, consistent with the notion that smokers buy packs as a self-control device, we find that smokers who have no intention of quitting do indeed substitute towards carton purchase while those who express a prior interest in quitting are more likely to purchase single packs in response to higher cigarette taxes. Since higher cigarette taxes may affect the composition of these two groups (i.e., those motivated to quit and those who do not want to quit), we employ simple simulation techniques designed to correct for

such tax-induced differences. We continue to find that those who have expressed an interest in quitting are systematically more likely to switch to packs when cigarette excise taxes increase, while those who do not want to quit switch to cartons.

In the next section we first describe the mechanisms by which higher cigarette taxes might lead to substitution between pack and carton purchasing. In general, there are two loose hypotheses. As alluded to above, the public health hypothesis that smokers will switch to cartons as a "high price avoidance" strategy and the behavioral hypothesis that suggests smokers will switch to pack purchasing to the extent that higher cigarette taxes induce a greater desire to quit smoking. In addition, we briefly describe the economics of carton versus pack purchasing and its implications for our analysis. Section 3 describes our data, including careful description of the construction of key variables and our analysis sample. Section 4 describes our empirical strategy, which involves the estimation of two-way fixed effect models, which allow us to relate changes in cigarette taxes to changes in carton versus pack purchase behaviour. More specifically, we model the choice of carton purchase as a function of state cigarette excises taxes. Section 5 presents our main estimates as well as some important extensions, while Section 6 concludes the paper.

2.2 Background

In the U.S., cigarettes are sold via two primary vehicles—in single packs of twenty cigarettes and in cartons which include ten packs, for a total of two-

hundred cigarettes. In general, smokers can realize a substantially lower per-pack price by purchasing cigarettes by the carton, rather than the single pack. In essence, they are offered a quantity discount. This discount is decidedly non-trivial: DeCicca, Kenkel and Liu (2010), for example, show that in 2006-2007 the average price per-pack paid by those who purchase single packs is just under $4, while the corresponding average for those who purchase cigarettes by the carton is about $2.85, a difference of about $1.15 per pack, or roughly forty percent.[2] Similarly, using the 2002 wave of the California Tobacco Survey (CTS), White et al. (2005) find that those who purchased their cigarettes by the carton saved $1.01 per pack on average. These estimates imply that a pack-a-day smoker might save as much as $400 per year by purchasing their cigarettes via cartons. Since these authors are not able to control for cigarette brand, and since carton purchasers may be more likely to buy less-expensive brands, these figures likely represent upper bounds of the true average price differential. Nevertheless, it is clear that smokers can reap substantial savings by purchasing their cigarettes in cartons instead of single packs.

2.2.1 Why do smokers buy packs rather than cartons?

Despite a large difference in relative per-pack prices, most smokers buy their cigarettes via packs, rather than cartons. Indeed, roughly two-thirds of smokers buy their cigarettes in the form of single packs, while only one third purchase them via cartons (DeCicca, Kenkel and Liu, 2010). Similarly, White et

[2] These authors also find a similar relative price difference using similar data from 2003

al. (2005) find that only twenty-eight percent purchase their cigarettes in carton form. Absent absolute income or liquidity constraints, such behaviour is puzzling. Cigarettes stay "fresh" for relatively long periods of time. While estimates vary from two to six months if unopened, all but the very lightest smokers would easily finish smoking ten packs long before their putative expiration date.[3] Recent work implies that smokers might purchase single packs rather than cartons as a commitment device designed to address what might be called a self-control problem (O'Donoghue and Rabin, 2000; Thaler and Shefrin, 1981).[4] Indeed, several studies develop a conceptual basis for what Wertenbroch (1998) labels "quantity rationing". There are many subtleties across different models, but the overriding idea, in the context of cigarette smoking, is that smokers understand that they have some kind of self-control problem that induces them to consume more than they plan to consume over some period and, as a result, they take steps to limit their consumption of the good in question by imposing various transactions costs on the consumption of that good. While certainly intuitively appealing, there is descriptive evidence that smokers do indeed ration quantities by purchasing packs, as opposed to cartons. For example, using data from the Survey on Smoking (SOS), Khwaja, Silverman and Sloan (2007) report that most smokers report using some kind of commitment device to limit their smoking.[5] In

[3] Interestingly, though perhaps not surprisingly, web searches related to the shelf life of cigarettes suggests that many smokers preserve them by keeping them in the freezer or refrigerator.
[4] Studies in economics also discuss this outside the context of the pack versus carton buying decision (cf., Gruber and Koszegi, 2001 and Gruber and Mullainathan, 2005).
[5] The Survey on Smoking was collected by the consulting firm Battelle Group from October 2004 to January 2005 and consists of adult respondents aged 50-70 years old

particular, over one-quarter of smokers say that they purchase single packs to keep their smoking behaviour in check.[6] If roughly-two thirds of smokers buy their cigarettes in single packs, this figure suggests that nearly forty percent of pack smokers claim that their purchase of packs is at least somewhat a commitment device.[7] Beyond pack versus carton purchase, the SOS asked individuals an open-ended question on what other commitment devices they employ; more prevalent responses included "keeping hands busy", "chewing gum" and other similar activities.

2.2.2 How might higher taxes affect the pack vs. carton choice?

Beyond the issue of why so many smokers purchase their cigarettes in single packs rather than cartons, there has been some speculation as to the impact of higher cigarette taxes on the pack versus carton choice. In other words, might higher taxes induce smokers to switch either from packs to cartons or vice versa? Research in public health suggests that higher taxes might induce smokers to purchase via cartons, as a "high price avoidance" strategy.[8] In other words, higher tax-induced prices may lead smokers to realize the quantity discount that they have foregone by buying single packs, rather than cartons. Indeed, the possibility is supported by the survey responses of smokers, themselves. For

[6] The relevant question asked smokers about the extent to which they agreed with the following statement, "To limit my smoking, I buy packs rather than cartons". Meaningful replies included: "disagree strongly", "disagree", "disagree somewhat", "agree somewhat", "agree" and "agree strongly"

[7] This figure is nearly identical to White et al. (2005) who report that nearly forty percent of single pack buyers report that they do not buy cartons because they "fear that they may smoke too much".

[8] Other such strategies include: switching to lower-priced brands (cf., Cummings et al., 1997; Hyland et al., 2005), increasing likelihood of cigarette purchase in lower-price jurisdictions (cf., DeCicca, Kenkel and Liu, 2010 and Merriman and Lovenheim) as well as tax-induced behavioural responses in how cigarettes are smoked (cf., Evans and Farrelly, 1998 and Adda and Cornaglia, 2006).

example, using data from the 2002 California Tobacco Survey (CTS), White et al. (2005) find that over one-quarter of smokers say that they have purchased their cigarettes via carton, rather than single pack, in order to reduce their expenditures on cigarettes. That said, these same authors find that many pack smokers are hesitant to switch to cartons since they thought they might smoke "too much", which is consistent with our earlier discussion of why most smokers tend to pass up the per-pack quantity discount associated with carton purchasing. Indeed, to the extent that higher cigarette taxes induce a greater desire among smokers to quit smoking, and to the extent that these individuals utilize commitment devices like single pack, instead of carton, purchasing, they might induce some carton smokers to switch to pack purchasing. While we cannot test this proposition directly, its intuition will be a key part of our empirical strategy.

2.2.3 Some simple economics of the impact of taxes on the pack versus carton choice

Another possible mechanism that might influence the pack versus carton choice is the impact of higher cigarette taxes on relative prices; in particular, the relative per-pack price of cartons versus single packs. To explore this, consider the following expression for the tax-inclusive relative price of cigarettes sold in cartons to those sold in individual packs:

$$\frac{P_C + aT}{P_P + bT} \qquad (1)$$

Here, the numerator is the tax-inclusive price per-pack of cartons and the denominator is the tax-inclusive price of single packs. Individually, P_C is the per-pack price of cigarettes sold by the carton and P_P is the price of single packs. T represents the per-pack cigarette tax, which is characteristic of cigarette taxes in the U.S., while the parameters a and b represent the pass-through rates of cartons and single packs, respectively.

If we differentiate this relative price expression with respect to T, set it equal to zero, and rearrange, we get the following equilibrium condition:

$$\frac{P_C}{P_P} = \frac{a}{b} \qquad (2)$$

This says that the impact of a tax increase on the relative price of cartons versus packs depends on the relative sizes of the prices to the pass-through rates. In particular, if the ratio of pass-through rates is larger than the price ratio then a tax increase (T) will increase the price of cartons relative to packs, and vice versa. Of course, if the two ratios are equal, a tax increase will have no impact on relative prices. We calculate these quantities and report them in Section 5 before we discuss our main estimates, but the calculations suggest that taxes have a very slight effect on relative prices.

2.3 Data

We use data from the Tobacco Use Supplements which are sponsored by the National Cancer Institute and administered as part of the Current Population Survey (TUS-CPS). The TUS-CPS is a key source of national and state level data

on smoking and other tobacco use in the United States because it uses a large, nationally representative sample that contains information on about 240,000 individuals in later waves. The TUS-CPS is repeated cross-sectional data and comprises seven cycles from 1992 to 2007. All cycles of TUS-CPS include demographic, socioeconomic, and detailed smoking information which are very useful for tobacco related research purposes. Starting with the 2003 wave, it provides novel information on price-search behaviours such as cross border and pack versus carton purchasing. Therefore, we focus on the 2003 cycle, which was conducted in the months February, June, and November of 2003, and the 2006-2007 cycle, which was conducted in the months of May and August of 2006 and in January 2007.

2.3.1 Measures of carton or pack buying

The dependent variable is derived from the survey question asking smokers whether they usually buy their cigarettes by the pack or carton. To be clear, a "pack" includes twenty cigarettes, while a "carton" includes ten packs of twenty or two-hundred cigarettes and usually involves a substantial quantity discount. Legitimate responses include "buy packs", "buy cartons", "buy both packs and cartons" or "do not buy own cigarettes". In our empirical work, we create three variables: "CARTON1" equals one if smokers buy cartons exclusively or buy both packs and cartons, and zero if they only buy packs; "CARTON2" equals one if smokers only buy cartons, and zero if they buy packs or buy both cartons and packs; "CARTON3" drops individuals who purchase both packs and cartons and

equals one if smokers only buy cartons, and zero if they only buy packs. We create these three measures since it is not clear how to categorize smokers who purchase both packs and cartons of cigarettes. Moreover, we drop individuals who report that they do not purchase their own cigarettes, which is a very small fraction of smokers, roughly 1.5 percent. As shown in Table 1, the percentage of smokers who usually buy cartons is consistently decreasing over our sample period. For example, CARTON1 decreases by around 1 percentage point each year off a base of roughly forty percent.

2.3.2 Cigarette taxes

The independent variable of interest is monthly cigarette excise taxes. We use data on taxes from Orzechowski and Walker (2008) to construct monthly state real cigarette taxes denominated in 2003 dollars. We merge this information to the TUS-CPS which contains data on state of residence and date of interview.

There were many large tax increases on cigarettes between 2003 and 2007 at the state-level. Given that cigarette taxes are virtually never reduced, these increases can be considered permanent, rather than temporary. Table 2 lists monthly cigarette taxes across each state. As we can see from Table 2, between February 2003 and January 2007, thirty out of fifty states and Washington, DC have increased their cigarette taxes. Over the entire period, and with respect to all increases, the average state excise tax increased from about 56 cents per pack to roughly 107 cents per pack, nearly doubling over this short period. Among them, the largest increase is 135 cents for Montana, while the smallest increase is 19

cents for Arkansas. Research in public finance suggests that the size and breadth of these tax increases were due to state budget shortfalls following the 2001-2002 recession in the United States (Maag and Merriman, 2003). To the extent that these increases, and those that followed them, were indeed driven by budgetary concerns, rather than state-specific anti-smoking sentiment, they likely represent more appropriate variation for estimating the causal effect of taxes.

2.3.3 Other independent variables and useful variables

Other control variables include demographic variables, socioeconomic variables, as well as indicators for state of residence and year and month of interview. With the exception of household size and number of household smokers, individual-level covariates are specified as a series of categorical variables to ensure a flexible functional form. In particular, excluding reference categories we include one indicator for gender, six for age, three for race, three for education, six for income, two for marital status. Finally, we control for the state-specific monthly unemployment rate as a proxy for economic conditions, since it is possible that fluctuations in the economy affect disposable income in a way that might influence the pack versus carton choice. In addition to these control variables, the data contain other variables which are useful for investigating heterogeneity in response to higher taxes (i.e., subsample analyses). These include smoker type, number of cigarettes smoked per day, and quitting intentions. Summary statistics for all variables are in Table 3.

2.3.4 Analysis sample

The 2003-2007 TUS-CPS files contain data on 483,557 individuals. Limiting our sample to everyday and someday smokers with valid state of residence information and who have provided information related to pack/carton buying results in a sample of 62,142 individuals, including 36,834 pack buyers (59.27 percent), 21,296 carton buyers (34.27 percent), and 4,012 smokers who buy both packs and cartons (6.46 percent).[9] Depending on different dependent variables and missing values of some regressors, sample size varies across specifications, but these figures represent our primary analysis samples. Note that income is most likely to have missing information. Since income is missing so often, we include a separate categorical variable for those whose income information is missing in order to preserve sample size. We employ complete case analysis for the remainder of the individual-level covariates.

2.4 Empirical strategy

In this section, we describe our strategy for identifying the impact of state cigarette taxes on the smoker's choice between buying their cigarettes via packs versus cartons[10]. Cross-sectional estimates of this relationship may be subject to serious omitted variables bias due to heterogeneity that may be correlated with state level cigarette excise taxes and cigarette purchasing behaviours simultaneously. For example, if we merely estimated cross-sectional models our estimates may be confounded by anti-smoking sentiment that is correlated with

[9] A total of 65,747 smokers were asked about their purchasing behaviour (pack vs. carton vs. both), so the response rate to this question is roughly 95 percent.
[10] Since the tax effect on quitting is very small and not significantly different from zero, our analysis sample is stable across time.

the tax rate and perhaps also the choice between single pack and carton purchasing. Put differently, we are concerned that smokers are different in higher tax states than in lower tax jurisdictions in ways that might lead cross-sectional estimation to yield biased results.

Because we have repeated cross-sectional data, we deal with this issue by including state-fixed effects in our regressions. To the extent that the heterogeneities in question are time-invariant, state-fixed effects will purge the correlation between the error term and the tax variable, eliminating potential omitted variables bias. In essence, the identification strategy is to compare the fraction of smokers who purchase cartons before and after tax increases, though we allow the treatment (i.e., the amount of the tax) to be a continuous variable. More specifically, we estimate two-way fixed effects specifications of the following form:

$$Carton_{ijmt} = \alpha + \beta_1 T_{jmt} + \beta_2 X_{ijmt} + \sigma_j + \mu_m + \tau_t + \varepsilon_{ijmt} \qquad (3)$$

where, i represents the individual, j stands for state, m is survey month and t is survey year. σ, μ, τ are state, month and year fixed effects, respectively. $Carton$ represents the carton or pack purchase behavior, T represents monthly state-specific cigarette excise taxes in constant dollars, and X includes all individual-level demographic variables as well as state-specific monthly unemployment rates to proxy for economic conditions in some specifications. Standard errors are clustered at the state-level to deal with within-state correlation in the errors.

Though they are useful in accounting for relevant heterogeneity, state, month and year fixed effects may absorb a great deal of variation in our cigarette excise tax variable. Put differently, there might be little identifying variation remaining after including fixed effects. To assess the correlation between tax and the set of fixed effects, we compute the variance inflation factor (VIF) which is $1/(1-R^2)$, where R^2 is computed from an auxiliary regression of the cigarette excise tax on the full set of fixed effects. Conventionally, it is assumed that if the variance inflation factor is greater than ten there is not enough independent variation left to perform the estimation (cf., Kennedy, 1994, p.183). In our case, it is roughly eight. This emphasizes the relatively high degree of within-state variation in cigarette taxes, alluded to earlier in this section. In the next section, we present our estimates.

2.5 Estimates

2.5.1 Main results

This section presents estimates from empirical models of smokers' pack versus carton purchasing behaviour in response to increases in cigarette excise taxes. As discussed in Sections 1 and 2, there are two broad theories regarding the pack versus carton choice. Public health research on "high price avoidance strategies" suggests that higher cigarette taxes will induce more smokers to buy cartons, since the per-pack price of cartons is much lower than that of a single pack. Intuition from behavioural economics suggests that smokers might switch to packs to the extent that higher taxes lead them to increase their willingness to

quit smoking. More directly, economic reasoning argues that the choice is affected by the relative price of cigarettes. For example, excise tax increases may lead to increased purchase of packs, because the per-pack relative price associated with a pack purchase will decrease if the ratio of cigarette tax pass-through rates between cartons and packs is greater than their original price ratio; please see the equations in Section 2.3 again for more detail. Overall, it is clear that the relationship is an empirical question.

To understand the likely impact of relative prices, which we discussed in Section 2.3, we report the pass-through rate ratio and original price ratio in Table 4. Table 4 presents estimates of the ratio of carton price (P_c) to pack price (P_p) for four different samples—2003, 2006, 2007 and 2003-2007, the latter of which represents all years of data. This information is repeated in Panels A, B and C; these panels correspond to our three measures of carton purchasing which are described in the footnotes of Table 4. Pass-through rates are generated from the same set of data and involve regressing price paid on cigarette tax. As can be seen in Panel A (i.e., CARTON1), there is little difference in the four estimated carton to pack price ratios (i.e., 0.735, 0.753, 0.731 and 0.739) and the ratio of the carton to pack pass-through rate (i.e., 0.754). This also holds, though slightly less strongly, for our other two definitions of carton purchasing behaviour as can be seen in Panels B and C. This pattern of results suggests that increasing cigarette excise taxes do very little, if anything, to change the relative price of cartons to

packs. If this is correct, it suggests that tax-induced relative price differentials in cartons and packs are not a likely mechanism by which our findings are generated.

Table 5 presents estimates of the effect of cigarette taxes on carton/pack purchasing behaviour for the three dependent variables discussed in Section 3.1[11]. Table 5 also presents estimates from two specifications; the upper panel is the basic specification represented by Equation (3) while the lower panel adds monthly state unemployment rate as a covariate. As can be seen in Table 5, a higher real cigarette tax is consistently associated with less carton buying across the three models.[12] Regarding model (3), a dollar increase implies a 1.65 percentage point decrease in carton buying. Relative to a base of 36.6%, this translates into a 4.5% reduction in the fraction of smokers who buy cartons. In model (1), because we classify smokers who buy both cartons and packs as carton buyers, the base of carton buying is larger. Therefore, the estimated 1.3 percentage point decrease translates to a 3.3% reduction. Finally, for Model (2) the roughly 1.5 percentage point decrease implies a 4.4% reduction in carton buying in response to a dollar increase in the cigarette tax. While these estimates are consistently negative, and imply substitution towards packs and away from cartons, they are not estimated precisely enough to achieve statistical significance

[11] Recall that CARTON1 equals one if smokers usually only buy cartons or both packs and cartons, zero if smokers usually only buy packs; CARTON2 equals one if smokers usually only buy cartons, zero if smokers usually only buy packs or both cartons and packs; CARTON3 equals on if smokers usually only buy cartons, zero if smokers usually only buy packs.

[12] Note also that there is little difference across the specifications, so from this point we focus on the specification that includes the unemployment rate and label it as our preferred specification.

at conventional levels.[13] In what follows, we perform selected subgroup analyses to investigate whether the imprecisely estimated substitution to pack purchasing is driven by the amount and/or frequency of smoking as well as if it is due to differences in income.[14] With regard to the latter we wish to understand whether the higher point-of-sale expenditure required to purchase a carton of ten packs is driving our estimates. If so, we expect to find a stronger negative relationship (i.e., more evidence of switching to packs) among the lowest income smokers for whom this type of constraint is most likely to bind.

2.5.2 Subgroup analyses

Because high frequency smokers may be more addicted to smoking relative to low frequency smokers, the former may have a stronger incentive to buy cheaper cigarettes in response to cigarette tax increases. In our context, heavier smokers might be more likely to realize the quantity discount associated with carton purchasing. Therefore, we define four subgroups: (1) everyday and someday smokers, where the former implies that a smoker reports smoking on every day and the latter implies that the smoker smokes on at least some days; (2) everyday smokers; (3) daily heavy and daily moderate smokers, where the former refers to smokers who report consuming more than thirty cigarettes per day and the later refers to smokers who report smoking between ten and thirty cigarettes

[13] The t-statistics for the estimated coefficients on the cigarette tax variable in models 1, 2, and 3 are -0.95, -1.07 and -1.14, respectively.
[14] From this point forward, all specifications include the monthly unemployment rate as a control for economic conditions.

per day; (4) daily heavy smokers[15]. Corresponding estimates are reported in Table 6. Three panels in Table 6 correspond to the three dependent variables. As can be seen, all effects are negative, which implies that the increase in cigarette tax leads smokers buy packs. However, none of the estimates is significantly different from zero, implying no systematic statistical relationship. If we compare the magnitude of marginal effects between column (1) and column (2), those in the second column are slightly larger than those in the first column. While neither coefficient is statistically different from zero at conventional levels, this pattern is not consistent with our intuition that everyday smokers have more incentive to cut cigarette expenditure by realizing the quantity discount, but it may suggest that everyday smokers have more incentive to cut cigarette expenditure by limiting cigarette consumption. If we compare the sizes between column (3) and column (4), this pattern is more obvious.

As noted, we also examine the relationship by household income level[16]. In particular, it is possible that very low income smokers might not be able to afford cartons following a tax increase, despite the lower per-pack price, though this implies a level of liquidity constraint that may not be realistic, given the relatively small amount of money in question. Estimates by income group are presented in Table 7. As can be seen, only smokers whose household income is between $25,000 and $35,000 have a significant tendency to switch to packs from carton

[15] We also have information on how soon smoker smokes his/her first cigarette of the day after waking up. Thus, we use it as a measure of addiction level and do subgroup study basing on it. Finally, we find similar pattern as that for the smoking frequency.

[16] We also adjust household income to account for household size. The results based on the adjusted income are similar to what we find for household income.

purchasing. No such systematic evidence is seen for those with household incomes less than $25,000 per year. Given that the inability to afford cartons represents an unlikely scenario (i.e., it would seem to imply an extreme liquidity constraint among smokers), this suggests that tax-induced income constraints are not responsible for the consistent inverse pattern of estimates we report in Table 5.

2.5.3 Digging deeper into competing explanations

While we find consistent evidence that smokers substitute towards packs and away from cartons when taxes increase, this evidence is not systematic in a statistical sense. Next, we examine the impact of higher cigarette taxes on the pack versus carton purchase choice by smokers stated motivation to quit. The purpose of this exercise is two-fold. First, it provides a better test of the behavioural hypothesis that smokers who are motivated to quit will substitute towards packs as a way of manifesting an increasing desire to quit. By contrast, those who have no intention to quit smoking, and whose intentions remain as such, should have a very different response to higher taxes if the behavioural story holds. In particular, they should not be more likely to substitute towards pack purchasing. Second, it provides an information on those who have no desire to quit smoking and may be differentially influenced by any anti-smoking policy including cigarette taxation.

To implement this exercise, we split our sample into two groups—one which is comprised of smokers who are ostensibly motivated to quit and one that is composed of smokers who have no intention of quitting. We use information

from the TUS-CPS to construct these groups. In particular, we label a smoker as "motivated to quit" if, in the past twelve months, they have attempted to quit smoking or at least seriously considered it. We label a smoker as "not motivated to quit" if they have neither attempted to quit nor seriously considered quitting in the past twelve months. In essence, this latter group represents smokers whose smoking status is relatively more permanent, so it is likely that they do not suffer from the sorts of self-control problems that smokers who say that they want to quit, but do not do so. Overall, the two groups are surprisingly similar demographically. For example, the average age of those who are labelled as "motivated to quit" is 41.6 years, while those not so motivated have an average age of 44.6 years. In terms of education, there are somewhat larger differences, but the two groups remain quite similar. For example, nearly 23.4 percent of the motivated group possesses a college degree or higher, while roughly 19 percent of the unmotivated group has the same level of formal education. Finally, the figures for household income are nearly identical as less than one percentage point separates each type of smoker (motivated and unmotivated) across all six household income categories (i.e., less than 15,000 per year, 15,000-25,000, 25,000-35,000, 35,000-50,000, 50,000-75,000 and greater than 75,000 per year).

Results from this exercise are included in Table 8 where column (1) includes estimates from a model that includes only those smokers who are defined as being motivated to quit, and column (2) includes estimates from a model that includes only those smokers whose responses suggest that they do not want to quit

smoking. By contrast, models in Table 5 include both types of these smokers. As before, Table 8 presents estimates for three different definitions of carton purchasing behaviour. The table presents estimates of the estimated marginal effect of a hypothetical one-dollar increase in the cigarette excise tax on whether the smoker purchases cartons or packs. As can be seen in all three models reported in column (1), the results imply that smokers who are motivated to quit systematically substitute away from cartons and into packs when cigarette excise taxes increase. For example, across the three models, a one-dollar increase implies roughly a five percentage point increase in the fraction who purchase their cigarettes in single packs. Off a base of roughly two-thirds, this represents about an eight percent increase in pack buying. By contrast, we find no evidence of substitution towards single pack purchasing for those who are not motivated to quit. As can be seen in column (2), we actually find a positive coefficient in all three models, which suggests that those not motivated to quit substitute away from single packs towards carton purchasing when cigarette taxes increase. That said, the estimates are not as systematically different from zero as those presented in column (1), but their pattern is striking nonetheless.[17] Overall, the findings presented in Table 8 are consistent with a behavioural economics story in that individuals who have expressed a desire to quit smoking substitute towards packs, while those who have expressed no such desire actually show some evidence of increasing their carton purchasing, consistent with the public health hypothesis,

[17] These estimates, along with all others, are nearly identical when a measure of "state anti-smoking sentiment" is added to the model (DeCicca et al., 2008).

when cigarette taxes increase. Note that these disparate results were, in essence, masked by pooling both types of smokers—those motivated to quit and those not motivated to quit—in the models associated with Table 5.

2.5.4 A simple simulation exercise

As discussed above, we allow the impact of taxes on the pack versus carton choice to vary across whether or not smokers are motivated to quit. While we find evidence that is consistent with a behavioural economics story (i.e., when taxes go up, motivated-to-quit smokers are more likely to purchase packs, while those who are not similarly motivated exhibit no such behaviour) rather than the one put forth by public health researchers, we recognize that motivation to quit smoking might well vary with taxes.[18] In particular, when taxes increase, it might be the case that smokers who previously did not want to quit, now wish to do so. If this is occurring in our data, then it will affect the composition of what we label "motivated-to-quit" and "not-motivated-to-quit" smokers. Indeed, we find evidence, that the fraction of "motivated-to-quit" smokers increases by about 2.2 percentage points, or nearly five percent when taxes increase by a hypothetical dollar.[19] While not a large effect, the change in composition represented by such marginal individuals might bias the findings we present in Table 8. In particular, any related bias will depend on how these marginal individuals compare to the

[18] Another, perhaps more basic, compositional issue is with respect to the impact of higher cigarette excise taxes on smoking participation since we examine only the behavior of smokers. We find no evidence in our sample that higher taxes lead to reduced smoking prevalence. This is consistent with many other adult tobacco use studies.

[19] In particular, the coefficient on the tax variable is -0.0225 with an associated standard error of 0.0120. This is also reported in Panel A of Table 8.

existing group of motivated-to-quit smokers. For example, if newly created motivated-to-quit smokers are more likely to buy their cigarettes in single packs than existing motivated-to-quit smokers, then the estimates in column (1) might reflect the change in motivation, rather than the impact of taxes on existing motivation. In a sense, they might be biased upward in magnitude. While this is possible, it seems unlikely that those newly induced to want to quit smoking are more likely to smoke packs than their fellow smokers who already wanted to quit. On the other hand, if tax-induced newly motivated-to-quit smokers are more likely to purchase cartons than existing motivated-to-quit smokers, then these results may be biased downwards in magnitude. Of course, it is not clear that this is such an important distinction since the measured effects in Table 8 ultimately reflect the impact of cigarette taxation on whether or not smokers choose packs or cartons. Moreover, it seems that given the relatively small response in quitting intentions (i.e., a five percent increase) the associated bias might be small. Nevertheless, to reinforce the idea, the issue is that the two groups—those motivated to quit and those not motivated—are not necessarily static when taxes increase. If they were fixed as taxes increase, as say groups defined by gender or age, then there would be no reason to address the implied compositional issues.[20]

To address this issue, we perform some simple simulations to address possible tax-induced compositional change in the two groups. In particular, we use estimates from our model of quitting intentions, where the dependent variable

[20] To simplify matters, we focus only on our dependent variable definition CARTON1 in what follows.

is consistent with how we define subsamples in Table 8. As noted above, we find a roughly 2.2 percentage point increase in the fraction of smokers who now are motivated to quit. This translates into an increase of 353 smokers who want to quit (i.e., those included in the column (1) models). To account for this, we randomly move 353 sample members from the set of individuals that is motivated to quit (i.e., represented in column (1) models) to the group who is not motivated to quit (i.e., represented in the column (2) models). The idea here is to replace those individuals who switch statuses—from not wanting to quit to wanting to quit—to hopefully recover the groups before the impact of the tax on quitting intentions. As discussed earlier, we need to assume something about the rate at which these individuals buy their cigarettes by the carton or pack. As can be seen in the five panels in Table 9, we allow for five hypothetical pack/carton purchasing ratios of the marginal smokers. Panel A (67%-33%) reflects the ratio of pack buyers to carton buyers among those who say they are motivated to quit (i.e., analogous to the sample of column (1) in Table 8) while Panel E (49%-51%) reflects this ratio for those who say they are not motivated to quit (i.e., analogous to the sample associated with column (2) in Table 8). These seem likely to be extreme fractions for the set of marginal smokers, so we also include three other, admittedly arbitrary, intermediate ratios (i.e., 62%-38%, 57%-42%, and 52%-48%). Moreover, for each estimate reported in each panel of Table 9 we estimate the model 100 times, drawing 100 different sets of 353 hypothetically marginal

smokers for each simulation. Table 9 presents median coefficient and associated standard error estimates.

When we perform the simulations, we find somewhat stronger results for those who are motivated to quit and these are presented in the first column of Table 9. As can be seen, the estimated marginal effects are larger in magnitude than the corresponding estimate in Table 8, but remain quite similar. By contrast, while we still find positive tax coefficients for the set of unmotivated smokers, as seen in the second column of Table 9, some are larger and some are smaller in magnitude, based on the assumed pack versus carton purchasing pattern of the marginal smokers (i.e., those who switch from not-motivated-to-quit to motivated-to-quit). That said, there are no dramatic departures from corresponding Table 8 estimates. Overall, the pattern of estimates shown in Table 9 remains consistent with a behavioural economics story—that is, higher cigarette taxes induce substitution to single packs from carton purchasing in those smokers who are motivated to quit. They are also consistent with claims from public health research that imply smokers will switch to carton purchasing to realize the implied quantity discount, though our findings suggest that this is only operative for those who are, in essence, committed smokers.

2.6 Conclusions

In this paper, we investigate the impact of higher cigarette excise taxes on carton versus single pack purchasing. Despite assertions from the public health literature, we found no systematic evidence that higher cigarette tax rates lead

smokers to a greater likelihood of carton purchasing. Instead, in our main models, we found weak evidence that smokers substituted from cartons to packs, consistent with the behavioural economics notion that higher taxes lead smokers to want to quit smoking and, in turn, they will be more likely to purchase single packs, since purchasing cartons would not be consistent with an intention to quit smoking. Also consistent with this idea, we find that smokers who are motivated to quit systematically substitute single pack purchasing for carton purchasing. However, consistent with suggestions from the public health literature, higher cigarette taxes induce those not motivated to quit to engage in carton purchasing.

Our results have interesting implications. Ostensibly, they would seem to imply that raising the per-pack price of cigarettes purchased by the carton, perhaps through a special carton tax, or outright banning the sale of cartons might be warranted. However, we find that only those smokers who have no intention of quitting smoking are impacted. In this light, such policies may not lead to improved outcomes, unless one believes that these adult smokers are, for example, myopic or misinformed. Finally, our findings also shed light on why it is that smokers actually purchase cigarettes in single packs when they could always realize a substantially lower per-pack price by purchasing cartons. In particular, our results suggest that taxes induce a greater desire to quit smoking. And, since purchasing cartons is likely seen as inconsistent with quitting, smokers are induced to switch to single pack purchasing. While this may not be too surprising, future work should investigate whether other tax avoidance behaviours (e.g.,

cross-border purchasing, switching to lower cost brands, etc.) depend on the motivation of the smoker to quit smoking.

References

Adda J. and F. Cornaglia (2006). "Prices, Cigarette Consumption and Smoking Intensity". *American Economic Review* 96(4): 1013-1028

Cummings, K. Michael, A. Hyland, E. Lewit, and D. Shopland (1997). "Use of Discount Cigarettes by Smokers in 20 Communities in the United States, 1988 - 1993." *Tobacco Control* 6 (Supplement 2): S25 - S30

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

DeCicca, Philip, D. Kenkel, F. Liu (2010). "Who Pays Cigarette Taxes? The Impact of Consumer Price Search." Working Paper #15942, NBER

Evans, W. N., M. C. Farrelly, (1998). "The compensating behavior of smokers: taxes, tar and nicotine." *RAND Journal of Economics* 29 (3): 578–595

Gruber, Jonathan and B. Koszegi (2001). "Is Addiction "Rational"? Theory and Evidence." *Quarterly Journal of Economics* 116(2): 1261-1303

Gruber, Jonathan and S. Mullainathan (2005). "Do Cigarette Taxes Make Smokers Happier." *Advances in Economic Analysis & Policy* 5(1): Article 4

Hyland, A., J. E. Bauer, Q. Li, et al. (2005). "Higher Cigarette Prices Influence Cigarette
Purchase Patterns." *Tobacco Control* 14: 86 – 92

Kennedy, P. (1994). *A Guide to Econometrics*, 3rd Ed. The MIT Press: Cambridge, MA.

Khwaja, Ahmed, D. Silverman, F. Sloan (2007). "Time Preference, Time Discounting, and Smoking Decisions" *Journal of Health Economics* 26: 927-949

Lovenheim, Michael F. (2008). "How Far to the Border? The Extent and Impact of
Cross-Border Casual Cigarette Smuggling." *National Tax Journal* 61 (1): 7 – 33

Maag, E., D. Merriman (2003). "Tax policy responses to revenue shortfalls." State Tax Notes (August 4)

Merriman, David (2010). "The Micro-geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chicago." *American Economic Journal: Economic Policy*, 2(2): 61–84

O'Donoghue, Ted and M. Rabin (2000). "The Economics of Immediate Gratification." *Journal of Behavioral Decision Making*, 13(2): 233-250

Orzechowski and Walker (2008). "*The Tax Burden on Tobacco: Historical Compilation*."
Orzechowski and Walker, Arlington Virginia

Thaler, Richard and H. Shefrin (1981). "An Economic Theory of Self-Control." *Journal of Political Economy* 89(2): 392–406

Wertenbroch, Klaus (1998). "Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice." *Marketing Science* 17 (4): 317-337

White Victoria M., E. A. Gilpin, M. M. White, J. P. Pierce (2005). "How Do Smokers Control Their Cigarette Expenditures?" *Nicotine & Tobacco Research* 7(4): 625-635

Tables

Table 1. Percentage of smokers who usually buy cigarettes by the carton

	2003	2006	2007
CARTON1	41.25	40.66	39.20
CARTON2	35.52	33.09	32.41
CARTON3	37.68	35.80	34.77

Notes: We have three different definitions of carton purchasing, due to the inclusion of a category labelled "both cartons and packs". CARTON1 equals one if smokers buy only cartons or buy both cartons and packs and it equals zero if smokers only buy packs. CARTON 2 equals one if smokers only buy cartons and equals zero if smokers buy only packs or buy both cartons and packs. CARTON3 equals one if smokers buy only cartons and equals zero if smokers buy only packs.

Table 2. Monthly state real per-pack cigarette taxes denominated in February 2003 dollars.

State	Tax 2/2003	Tax 6/2003	Tax 11/2003	Tax 5/2006	Tax 8/2006	Tax 1/2007
AL	0.17	0.17	0.16	0.39	0.38	0.38
AK	1.00	1.00	0.99	1.46	1.62	1.63
AZ	1.18	1.18	1.17	1.08	1.06	1.07
AR	0.34	0.59	0.59	0.54	0.53	0.53
CA	0.87	0.87	0.86	0.79	0.78	0.79
CO	0.20	0.20	0.20	0.77	0.76	0.76
CT	1.11	1.51	1.50	1.38	1.36	1.36
DE	0.24	0.24	0.55	0.50	0.50	0.50
DC	1.00	1.00	0.99	0.91	0.90	0.90
FL	0.34	0.34	0.34	0.31	0.31	0.31
GA	0.12	0.12	0.37	0.34	0.33	0.33
HI	1.20	1.20	1.29	1.28	1.26	1.44
ID	0.28	0.57	0.57	0.52	0.51	0.51
IL	0.98	0.98	0.97	0.89	0.88	0.88
IN	0.56	0.56	0.55	0.51	0.50	0.50
IA	0.36	0.36	0.36	0.33	0.32	0.33
KS	0.79	0.79	0.78	0.72	0.71	0.71
KY	0.03	0.03	0.03	0.27	0.27	0.27
LA	0.36	0.36	0.36	0.33	0.32	0.33
ME	1.00	1.00	0.99	1.82	1.80	1.81
MD	1.00	1.00	0.99	0.91	0.90	0.90
MA	1.51	1.51	1.50	1.38	1.36	1.36
MI	1.25	1.25	1.24	1.82	1.80	1.81
MN	0.48	0.48	0.48	1.35	1.34	1.34
MS	0.18	0.18	0.18	0.16	0.16	0.16
MO	0.17	0.17	0.17	0.15	0.15	0.15
MT	0.18	0.70	0.69	1.55	1.53	1.53
NE	0.64	0.64	0.64	0.58	0.58	0.58
NV	0.35	0.35	0.79	0.73	0.72	0.72
NH	0.52	0.52	0.52	0.73	0.72	0.72
NJ	1.50	1.50	2.03	2.19	2.32	2.32
NM	0.21	0.21	0.90	0.83	0.82	0.82
NY	1.50	1.50	1.49	1.37	1.35	1.35
NC	0.05	0.05	0.05	0.27	0.32	0.32
ND	0.44	0.44	0.44	0.40	0.40	0.40
OH	0.55	0.55	0.55	1.14	1.13	1.13
OK	0.23	0.23	0.23	0.94	0.93	0.93
OR	1.28	1.28	1.27	1.08	1.06	1.07

PA	1.00	1.00	0.99	1.23	1.22	1.22
RI	1.32	1.32	1.70	2.24	2.22	2.22
SC	0.07	0.07	0.07	0.06	0.06	0.06
SD	0.33	0.53	0.53	0.48	0.48	1.38
TN	0.20	0.20	0.20	0.18	0.18	0.18
TX	0.41	0.41	0.41	0.37	0.37	1.27
UT	0.70	0.70	0.69	0.63	0.63	0.63
VT	0.93	0.93	1.18	1.08	1.61	1.62
VA	0.03	0.03	0.02	0.27	0.27	0.27
WA	1.43	1.43	1.41	1.85	1.83	1.83
WV	0.17	0.55	0.55	0.50	0.50	0.50
WI	0.77	0.77	0.76	0.70	0.69	0.70
WY	0.12	0.12	0.60	0.55	0.54	0.54

Notes: Tax is for a single pack of twenty cigarettes and is rounded to the nearest penny.

Table 3. Summary statistics

Variable	Mean
Panel A:	
Never smoker	0.627
Everyday smoker	0.144
Some day smoker	0.033
Former smoker	0.196
Panel B:	
Real cigarette tax ($ per pack)	0.767
Female	0.521
Age 15-24	0.111
Age 25-34	0.196
Age 35-44	0.232
Age 45-54	0.233
Age 55-64	0.143
Age 65-79	0.077
Age 80+	0.009
Hispanic	0.063
Non- Hispanic white	0.799
Non- Hispanic black	0.083
Others	0.055
Less than high school	0.181
High school	0.402
Some college	0.207
College or higher	0.210
Household Income <15k	0.185
Household Income 15k-25k	0.132
Household Income 25k-35k	0.142
Household Income 35k-50k	0.151
Household Income 50k-75k	0.162
Household Income 75k+	0.136
Household Income missing	0.093
Married	0.446
Divorced, widowed, or separated	0.293
Never married	0.261
Employed	0.660
Unemployed	0.065
Retired	0.099
Not in the labor force	0.177
Household size	2.643
Number of smokers in the household	1.388
Year 2003	0.533

Year 2006	0.296
Year 2007	0.171
Some day smoker	0.171
Daily light smoker (<10)	0.114
Daily moderate smoker (10-30)	0.665
Daily heavy smoker (>30)	0.051
Smokers who have intentions to quit smoking	0.454
Smokers who do not have intentions to quit smoking	0.546

Notes: Statistics in Panel A are calculated from the entire sample of TUS-CPS, 2003-2007. Statistics in Panel B are calculated from the sample of smokers who have provided information related to pack/carton purchasing.

Table 4: Comparisons of real per-pack price ratio and passthrough rate ratio.

Panel A: CARTON1			
	Carton price	Pack price	Ratio
2003	$2.577	$3.505	0.735
2006	2.733	3.628	0.753
2007	2.958	4.049	0.731
2003-2007	2.686	3.637	0.739
	Carton buyers	Pack buyers	ratio
Pass-through rate	.716***	.949***	0.754
Panel B: CARTON2			
	Carton price	Pack price	Ratio
2003	$2.475	$3.486	0.710
2006	2.531	3.632	0.697
2007	2.716	4.064	0.668
2003-2007	2.530	3.632	0.697
	Carton buyers	Pack buyers	Ratio
Pass-through rate	.604***	.964***	0.627
Panel C: CARTON3			
	Carton price	Pack price	Ratio
2003	$2.475	$3.505	0.706
2006	2.531	3.628	0.698
2007	2.716	4.049	0.671
2003-2007	2.530	3.637	0.696
	Carton buyers	Pack buyers	Ratio
Pass-through rate	.604***	.949***	0.636

Notes: In Panel A: The related subsamples are defined by "CARTON1". CARTON1 equals one if smokers buy only cartons or buy both cartons and packs and equals zero if smokers buy only packs. In Panel B: The related subsamples are defined by "CARTON2". CARTON 2 equals one if smokers buy only cartons and equals zero if smokers buy only packs or buy both cartons and packs. In Panel C: the related subsamples are defined by "CARTON 3". CARTON 3 equals one if smokers buy only cartons and equals zero if smokers buy only packs. Pass-through rates are generated by regressing price paid on the cigarette tax variable. *** represents statistically significance at 1% level, ** represents statistically significance at 5% level, and * represents statistically significance at 10% level.

Table 5. Estimated effect of real cigarette tax on carton purchase

Dependent variable	(1) CARTON1	(2) CARTON2	(3) CARTON3
Real Cigarette tax	-0.0134	-0.0157	-0.0170
	(0.0133)	(0.0136)	(0.0138)
	[-1.01]	[-1.15]	[-1.23]
	{0.32}	{0.25}	{0.22}
Dependent mean	0.407	0.343	0.366
N	61951	61951	57946
Real Cigarette tax	-0.0133	-0.0151	-0.0165
	(0.0140)	(0.0141)	(0.0144)
	[-0.95]	[-1.07]	[-1.14]
	{0.35}	{0.29}	{0.26}
Dependent mean	0.407	0.343	0.366
N	61951	61951	57946

Notes: Dependent variable in column (1) is CARTON1. Dependent variable is column (2) is CARTON2. Dependent variable in column (3) is CARTON3. We present estimates from linear probability models; probit marginal effects are nearly identical in all cases. Coefficients on real cigarette tax are reported in the first row, standard errors are in parentheses in the second row, t-statistics are in brackets in the third row, p-values are in braces in the fourth row. The specification in the upper panel includes state-specific monthly real cigarette tax, sex, age, race, education, family income, marital status, employment status, household size, number of smokers within each household, state, month and year effects. In the lower panel we add the state-specific monthly unemployment rate as a control variable. In all tables that follow, we use the fuller specification represented in the lower panel.

Table 6: Estimated effect of real cigarette tax on carton purchase by smoking status.

	(1)	(2)	(3)	(4)
Panel A: CARTON1	Everyday & Someday	Everyday	Daily heavy & Daily moderate	Daily heavy
Real cigarette tax	-0.0133	-0.0144	-0.0125	-0.0417
	(0.0140)	(0.0153)	(0.0174)	(0.0440)
	[-0.95]	[-0.94]	[-0.72]	[-0.95]
	{0.35}	{0.35}	{0.48}	{0.35}
N	61951	51450	44094	3147
Panel B: CARTON2				
Real cigarette tax	-0.0151	-0.0159	-0.0140	-0.0294
	(0.0141)	(0.0159)	(0.0184)	(0.0465)
	[-1.07]	[-1.00]	[-0.76]	[-0.63]
	{0.29}	{0.32}	{0.45}	{0.53}
N	61951	51450	44094	3147
Panel C: CARTON3				
Real cigarette tax	-0.0165	-0.0181	-0.0151	-0.0489
	(0.0144)	(0.0162)	(0.0188)	(0.0473)
	[-1.14]	[-1.12]	[-0.80]	[-1.03]
	{0.26}	{0.27}	{0.43}	{0.31}
N	57946	47729	40705	2869

Notes: The dependent variables for Panel A, B, C are CARTON1, CARTON2, and CARTON3 respectively. Column (1) includes everyday and someday smokers. Column (2) includes only everyday smokers. Column (3) includes daily heavy and daily moderate smokers. Column (4) includes daily heavy smokers. We present estimates from linear probability models; probit marginal effects are nearly identical in call cases. Coefficients of real cigarette tax are reported in the first row of each panel, standard errors are in parentheses in the second row of each panel, t-statistics are in brackets in the third row of each panel, p-values are in braces in the fourth row of each panel. The model specification matches that in the lower panel of Table 5.

Table 7: Estimated effect of real cigarette tax on carton purchase by income groups.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: CARTON1	<15k	15k-25k	25k-35k	35k-50k	50k-75k	>75k
Real cigarette tax	-0.0179	0.0267	-0.0667**	-0.0127	-0.0445	-0.0041
	(0.0227)	(0.0282)	(0.0298)	(0.0298)	(0.0388)	(0.0278)
	[-0.79]	[0.95]	[-2.24]	[-0.43]	[-1.15]	[-0.15]
	{0.43}	{0.35}	{0.03}	{0.67}	{0.26}	{0.88}
N	11456	8181	8772	9353	10006	8433
Panel B: CARTON2						
Real cigarette tax	-0.0265	0.0217	-0.0608**	-0.0230	-0.0354	0.0169
	(0.0215)	(0.0339)	(0.0256)	(0.0323)	(0.0357)	(0.0265)
	[-1.23]	[0.64]	[-2.37]	[-0.71]	[-0.99]	[0.64]
	{0.22}	{0.52}	{0.02}	{0.48}	{0.33}	{0.53}
N	11456	8181	8772	9353	10006	8433
Panel C: CARTON3						
Real cigarette tax	-0.0247	0.0229	-0.0676**	-0.0217	-0.0462	0.0125
	(0.0226)	(0.0334)	(0.0260)	(0.0327)	(0.0385)	(0.0278)
	[-1.09]	[0.69]	[-2.60]	[-0.66]	[-1.20]	[0.45]
	{0.28}	{0.50}	{0.01}	{0.51}	{0.24}	{0.65}
N	10662	7630	8227	8782	9413	7989

Notes: The dependent variables for Panel A, B, C are CARTON1, CARTON2, and CARTON3 respectively. Column (1) includes smokers with household income <15,000 dollars. Column (2) includes smokers with household income 15,000-25,000. Column (3) includes smokers with household income 25,000-35,000. Column (4) includes smokers with household income 35,000-50,000. Column (5) includes smokers with household income 50,000-75,000. Column (6) includes smokers with household income >75,000. We present estimates from linear probability models; probit marginal effects are nearly identical in all cases. Coefficients of real cigarette tax are reported in the first row of each panel, standard errors are in parentheses in the second row of each panel, t-statistics are in brackets in the third row of each panel, p-values are in braces in the fourth row of each panel. *** represents statistically significance at 1% level, ** represents statistically significance at 5% level, and * represents statistically significance at 10% level.

Table 8: Estimated effect of real cigarette tax on quitting intentions and purchasing behavior.

Panel A: Effect on quitting intentions		
Real cigarette tax	0.0225*	
	(0.0120)	
	[1.87]	
	{0.07}	
N	63735	
	(1)	(2)
Panel B: Carton1	Motivated to quit	Not Motivated to Quit
Real cigarette tax	-0.0587***	0.0301**
	(0.0210)	(0.0141)
	[-2.79]	[2.14]
	{0.01}	{0.04}
N	27000	33707
Panel C: Carton2		
Real cigarette tax	-0.0461**	0.0176
	(0.0200)	(0.0168)
	[-2.31]	[1.05]
	{0.03}	{0.30}
N	27000	33707
Panel D: Carton3		
Real cigarette tax	-0.0533**	0.0240
	(0.0210)	(0.0159)
	[-2.54]	[1.51]
	{0.01}	{0.14}
N	25542	31260

Notes: The dependent variable for Panel A is binary quitting intention indicator which equals one if, during the past twelve months, smokers have attempted or seriously considered quitting, and zero if smokers have never attempted nor

seriously considered quitting in this time frame. The dependent variables for Panel B, C, D are CARTON1, CARTON2, and CARTON3, respectively, and remain as defined earlier. All regression estimates are derived from linear probability models; corresponding probit marginal effects are nearly identical in all cases. Column (1) of Panel B, C, D includes smokers who have intention to quit. Column (2) of Panel B, C, D includes smokers who have no intention to quit. Coefficients of real cigarette tax are reported in the first row of each panel, standard errors are in parentheses, t-statistics are in brackets, and p-values are in braces. *** represents statistically significance at 1% level, ** represents statistically significance at 5% level, and * represents statistically significance at 10% level.

Table 9. Simulation results of cigarette tax effects on carton purchasing

% of marginal smokers who buy packs vs. cartons:	(1)	(2)
Panel A:		
Real cigarette tax	-0.0646***	0.0227
(67% vs. 33%)	(0.0216)	(0.0144)
	[-2.99]	[1.58]
Panel B:		
Real cigarette tax	-0.0680***	0.0261*
(62% vs. 38%)	(0.0229)	(0.0142)
	[-2.98]	[1.84]
Panel C:		
Real cigarette tax	-0.0717***	0.0289**
(57% vs. 43%)	(0.0226)	(0.0145)
	[-3.17]	[1.99]
Panel D:		
Real cigarette tax	-0.0748***	0.0314**
(52% vs. 48%)	(0.0224)	(0.0138)
	[-3.34]	[2.28]
Panel E:		
Real cigarette tax	-0.0773***	0.0334**
(49% vs. 51%)	(0.0225)	(0.0136)
	[-3.43]	[2.46]

Notes: The dependent variable of each regression in this table is CARTON1. Column (1) contains estimates from models that include only those smokers who are motivated to quit smoking, while Column (2) contains estimates from models that include only those smokers who are not motivated to quit smoking, as defined in our Table 8 related discussion. Each panel represents a rate at which these marginal individuals buy their cigarettes by the carton or pack. The first row of each panel is the median of the simulated tax effects, of which we computed one-hundred times for each panel. Corresponding standard errors and t-statistics are in parentheses and brackets, respectively. *** represents statistically significance at 1% level, ** represents statistically significance at 5% level, and * represents statistically significance at 10% level.

Chapter 3

Investigating the Price Sensitivity of Older Smokers

3.1 Introduction

Since the second half of the 20th century, following the publication of seminal British and American reports on smoking and health, the focus of economic research on smoking has been shifted toward analysis regarding the potential public health damage associated with smoking. Evidence indicates that in recent decades, the percentage of smoking related deaths has increased, primarily due to lung cancer and heart disease. Increasing cigarette excise tax rates may be a very useful public policy because higher tax-induced prices might either reduce the

amount a smoker consumes or lead more smokers to quit smoking altogether. A common finding in the smoking literature is that price sensitivity varies inversely with age, primarily since the addictive nature of smoking makes adult smokers adjust less quickly to changes in price, which implies that older people are not as sensitive to price. Despite these findings, older adult smoking behavior remains empirically interesting, because many papers suggest that even later-in-life smoking cessation might yield substantial improvements in health.

Using data from the U.S. Tobacco Use Supplements of the Current Population Surveys (TUS-CPS) during a time of large increases in state cigarette tax rates (2003-2007), we examine the relationship between cigarette excise taxes and smoking participation for older adults. Given the repeated cross-sectional nature of the data, we estimate the two-way fixed effect models that include controls for individual demographic and socioeconomic characteristics as well as other state-level policies implemented over the same period. In addition, we employ a novel measure of anti-smoking sentiment from DeCicca et al. (2008) to control for unobserved collective sentiment towards smoking which may affect both the cigarette tax and aggregate smoking behaviour in an area.

We find that higher cigarette taxes are strongly correlated with greater quitting intentions and reduced smoking participation for older adults. Consistent with DeCicca and McLeod (2008), our findings run contrary to most of the related literature and have important implications for public policy, especially since the

health gains associated with smoking cessation by older adults may be substantial.[21]

In the next section, we briefly review the existing literature on the impact of price on cigarette demand. Most of these studies rely on the cross-sectional variation in taxes and very few have used the most recent large tax increases which occurred primarily after 2000. Section 3 describes our data, including a description of the construction of key variables and our analysis sample. Section 4 describes our empirical strategy, which involves the estimation of two-way fixed effect models, and in effect relates changes in state cigarette excise taxes to changes in smoking participation and intentions to quit smoking. Section 5 presents our main estimates as well as some important extensions, while Section 6 concludes the paper.

3.2 Background

Tobacco use, particularly cigarette smoking, remains the leading cause of preventable illness and death in the developed world. In this context, cigarette taxes have attracted increasing interest by both government and public health officials during the past 30 years. The former are interested in using state–level excise taxes to increase government revenues, while the latter believe that increased taxes can either reduce amount smoked or aid in smoking cessation. The extent to which each of these goals can be met depends on the degree of price

[21] Comparing with DeCicca and McLeod (2008), we employ more recent large tax variations in US and study intentions to quit before the analysis of smoking participation of old smokers.

sensitivity of the smokers. In our case, we are interested in the price sensitivity of smoking by older, established smokers.

For a more complete literature review on the issue of cigarette taxes and smoking behaviors, we refer the handbook chapter by Chaloupka and Warner (2000) to readers for an excellent review. In the current review, we focus on the works that are closely related to our study. In the literature, economists have focused relatively more on youth smoking, and few papers provide evidence on the price-responsiveness of older adults. It is generally accepted that the price sensitivity of smoking is inversely related to age and older smokers are not sensitive to price.[22] One of the earliest studies [Lewit and Coate (1982)], estimated the effect of price on cigarette consumption for both extensive and intensive margins across the entire age distribution using cross-sectional data from the 1976 National Health Interview Survey[23]. In general, they found an overall price elasticity of cigarette consumption of -0.42 and an elasticity of smoking participation of -0.26. Moreover, the relationship between price elasticity and age was inversely related, which results in a total price elasticity of cigarette consumption for the young individuals more than double that of the relatively old people[24]. In addition, they found that price affected cigarette consumption of younger and older adults differently. For young adults, price mainly affected their decision to start smoking (participation elasticity of -0.74 and conditional demand

[22] Throughout our discussion we refer to the absolute value of the price elasticity of demand when we say "inversely" related.
[23] Extensive margin measures the effect on the smoking participation. Intensive margin measures the effect on the amount smoked conditional on smoking (conditional demand).
[24] Young people include those aged from 20 to 25. The old include those aged 26 and older.

elasticity of -0.20), but for those over 35 years of age, the price effect was about equally distributed (smoking participation and conditional demand elasticities of -0.15).

Several other early studies of youth and adult smoking have supported the conclusions reached by Lewit and his colleagues [Lewit et al. (1981), Lewit and Coate (1982), and Grossman et al. (1983)] in that the price elasticity of cigarette demand was inversely related to age and they find no systematic evidence that higher prices reduce smoking participation among old adults. Farrelly and his colleagues (1998) estimated the price elasticity of demand using thirteen waves of the National Health Interview Survey conducted between 1976 and 1992. They found that demand was more than six times as elastic for their sample of young adults, aged 18 to 24 years (a participation elasticity of -0.37), as for adults aged 40 or higher (a participation elasticity of -0.06). In essence, they found no evidence that higher taxes/prices reduced the smoking of individuals who were at least 40 years old. Using the data from the 1987 National Health Interview Survey, Evans and Farrelly (1998) found young smokers, aged 18 to 24, were much more responsive in changing smoking habits when tax increases than were older smokers. They estimated the effect of tax on the smoking participation of the older adults and found that it was not statistically different from zero. Their findings implied that cigarette excise taxes did not affect older adult smoking.

The relationship between price elasticity and age is not only revealed from the studies which investigated the price effect across the entire age distribution,

but also from studies that focus on the youth and older adults separately.

3.2.1 Previous research on youth and young adult smoking behavior

Using data from Cycle III of the US Health Examination Survey conducted in the period 1966-1970, Lewit et al. (1981) estimated the price elasticity of cigarette demand for the teenagers aged from 12 to 17. They found that the price elasticity of smoking participation was -1.20, and the conditional demand elasticity was -0.25. Their estimated total price elasticity of youth cigarette demand of -1.45 is more than three times that of Lewit and Coate's (1982) estimate for adults. More recently, Lewit et al. (1997) used the data from two cross-sectional, school-based surveys of ninth-grade students conducted in 21 North American communities in 1990 and 1992 to examine the effect of cigarette taxes on smoking participation and the intention to smoke among students[25]. They found that both smoking participation and the intent to smoke were inversely related to differences in cigarette prices, with estimated price elasticities of -0.87 and -0.95, respectively, which were similar to Lewit et al. (1981). Evans and Huang (1998) used state level aggregated data on smoking prevalence from Monitoring the Future surveys constructed from 1977 through 1992 to estimate the impact of cigarette excise tax on smoking participation of the youth. Unlike other studies of smoking participation that relied on the cross-sectional comparisons, their work utilized the state fixed effect models to construct

[25] Ninth-grade students aged from 13 to 16.

comparisons within each state. This way can efficiently deal with estimation bias caused by unobservable state-level differences that correlated with both tax and smoking behaviors. They found the price elasticity of youth smoking participation was -0.20. This estimated elasticity was small and they further analyzed the price elasticity across time. They concluded that smokers were becoming more price sensitive over time and the price participation elasticity for the youth was -0.5 when they restricted the sample between 1985 and 1992. Tauras and Chaloupka (1999) used the longitudinal data on high school seniors conducted from 1976 to 1993 as a part of the Monitoring the Future program to study the effect of cigarette price on smoking participation rate and conditional demand. With this data set, seven follow-ups were conducted on each senior class, so they could use individual fixed effects models to control for time-invariant selection bias. The estimation results indicated that cigarette price increase had significant effect on the reduction of both smoking participation and amount smoked with overall price elasticity of -0.79.

More recent studies provide somewhat mixed evidence of the tax-responsiveness of youth and young adults. Gruber and Zinman (2001) employed the first four waves of national Youth Risk Behavior Surveys (YRBS) to estimate the effect of cigarette taxes on youth cigarette consumption. They only found modest tax effects for the older youth. Using the data from National Educational Longitudinal Study (NELS), DeCicca et al. (2002) studied whether cigarette tax can substantially reduce youth smoking. They found that once the state-fixed

effects were controlled for, cigarette taxes did not have significant effects on the youth smoking initiation. Sloan and Trogdon (2004) estimated the price effect using the data from Behavior Risk Factor Surveillance System (BRFSS). They found that the adults, aged from 18 to 20, were significantly responsive to price even with state and year fixed effects in the model, while they found no effect for those aged 21-24. DeCicca et al. (2008) developed a direct measure of state-level anti-smoking sentiment to deal with the unobservable heterogeneity when estimating cross-sectional model. Their results showed a consistent pattern in that once differences in state anti-smoking sentiment were accounted for, youth smoking participation was statistically unrelated to cigarette prices. Carpenter and Cook (2008) employed two-way fixed effects models to estimate the effect of cigarette taxes on youth smoking behavior. They found that tax increases were significantly associated with reduction in participation and frequency.

3.2.2 Previous research on older adults smoking behavior

The conclusion that price sensitivity varies inversely with age were generally accepted until an influential 1991 Rand study by Wasserman et al. (1991). This paper used the data from smoking supplemental questionnaires administered between 1970 and 1985 of the National Health Interview Survey (NHIS) to examine the extent to which cigarette excise tax affected adult cigarette consumption and used the data from National Health and Nutrition Examination Survey II (NHNES II) to investigate the impact of cigarette tax on teenagers' demand. Using a generalized linear model, they found that the impact of cigarette

taxes for adults was unstable over time. In the earlier years of the data, adults were relatively not sensitive to price change with an overall price elasticity of 0.06 in 1970, but were becoming more price-sensitive latterly with an overall price elasticity of -0.23 in 1985. Based on the trend of estimation, they predicted that the overall price elasticity would be around -0.28 in 1988. In order to analyze smoking participation and conditional demand separately, this paper decomposed the overall price elasticity using a two-part model. Estimates indicated that the impact of price on smoking participation was almost two times that of the conditional demand. With respect to youth smoking, they did not find any significant effect of price. The authors attributed their relatively low estimates of price elasticity, particularly those for youth, to the inclusion of a state-specific index of restrictions on smoking. These restrictions had not been included in most previous studies of cigarette demand. Indeed, they obtained very similar estimates to Lewit and Coate (1982) when leaving the restriction index out of models estimated using the 1976 survey data.

More recent studies suggested somewhat mixed evidence of the price effect on older smokers' participation. Tauras (2006) used the repeated cross-sectional data from TUS-CPS (1992-1999) to study the effect of cigarette prices and smoking-free laws on adult cigarette consumption. In order to deal with the mass of non-smokers (zero cigarette consumption) in the data, he employed two-part model with probit method and generalized linear model with log-link and Gaussian distribution to model the participation and conditional demand,

respectively. For each part of the model, he also controlled for state fixed effects to deal with the unobservable difference in state anti-smoking sentiment that may bias the estimated effect. Finally, he found a small but precisely estimated effect of price on smoking participation (participation elasticity of -0.12), and even a smaller conditional demand elasticity of roughly -0.07. Moreover, he found no impact of the smoking-free laws on prevalence, and a significant but small effect on the amount smoked.

Most recently, DeCicca and McLeod (2008) focused exclusively on older adults aged 45 to 59. Using repeated cross-sectional data from BRFSS from 2000 to 2005, they investigated the responsiveness of older adult smoking using variation from recent historically large cigarette tax increases in the United States. These authors used two empirical strategies: two-way fixed effects (state and time) and they included a measure of state-level anti-smoking sentiment which we also use and will discuss at more length later. Contrary to other studies in this literature, they found relatively large estimates and consistent evidence that higher taxes reduce smoking for the older adults, especially those who were less educated and live in low-income households.

3.2.3 The impact of smoking cessation on health

Investigating the effect of tax on smoking behaviors among the older adults has important implications for public policy, especially since the health gains associated with smoking cessation by older adults may be substantial. In 1990, the U.S. Surgeon General completed the first comprehensive study of smoking

cessation (USDHHS, 1990). Among other things, it concluded that smoking cessation improved immediate and long-term health, including substantial reductions in mortality. Most relevant to our study, it provided evidence that even later-in-life smoking cessation increased longevity. In this regard, one of its key findings suggested that a 50-year old who quits smoking cuts the risk of dying by age 65 in half.

Several more recent studies also provided evidence of important health benefits of smoking cessation among the old. For instance, Ostbye et al. (2002) used two large nationally representative data sets, the Health and Retirement Study (HRS) and the Asset and Health Dynamics among the Oldest Old survey (AHEAD), to study the effect of smoking and other modifiable risk factors on ill health for middle-aged and older Americans. They found strong and consistent negative impacts of smoking on ill health and the magnitude of the effects were similar for the middle-aged and the elderly. In perhaps the most relevant study, Taylor et al. (2002) used the data from Cancer Prevention Study II to determine the life extension obtained from stopping smoking at various ages. They found that life expectancy among smokers who quit at age 35 exceeds that of continuing smokers by 6.9 to 8.5 years for men and 6.1 to 7.7 years for women. Smokers who quit at younger ages realized greater life extensions. However, even those who quit much later in life gained some benefits: among smokers who quit at age 65 years, men gained 1.4 to 2.0 years of life, and women gained 2.7 to 3.7 years. Their findings suggested that quitting smoking even at later ages still may provide

meaningful life extensions. More recently, Critchley et al. (2003) studied the effect of smoking cessation on the risk reduction in all-cause mortality for patients with coronary heart disease (CHD). Their results showed that quitting smoking was associated with a substantial reduction in all-cause mortality among patients with CHD, regardless of age.

Finally, another paper by Ostbye and Taylor (2004) used the data from Health and Retirement Study (HRS) and Asset and Health Dynamics Among the Oldest Old (AHEAD) to investigate the effect of smoking on life left and the quality of life left for the old, 50 to 89 years of age. Employing multivariate linear model, they related smoking and other explanatory variables to two outcome variables: years of life remaining and years of healthy life remaining and found that smoking was strongly and consistently related to both years of life and healthy life lost, which again implied that smoking cessation later in life may provide substantial health benefits.

3.3 Data

In order to investigate the impact of cigarette taxes on older adult smoking participation, we employ data from the Tobacco Use Supplements which are sponsored by the National Cancer Institute and administered as part of the Current Population Surveys (TUS-CPS). The TUS-CPS is a key source of national and state level data on smoking and other tobacco use in the United States because it uses a large, nationally representative sample that contains information on almost a quarter of a million individuals in any given survey period. The TUS-CPS is

repeated cross-sectional data and comprises seven cycles from 1992 to 2007. All cycles of TUS-CPS include demographic, socioeconomic, and detailed smoking information which are useful for tobacco related research purposes. Because many states initiated large cigarette tax increases between 2002 and the present, and given the availability of data, we focus on the cycles that were conducted from 2003 to 2007. In particular, the 2003 cycle was conducted in February, June, and November of that year and the 2006-2007 cycle was conducted in May and August of 2006 and January of 2007.

3.3.1 Measures of smoking participation and intentions to quit smoking

TUS-CPS asks all eligible respondents for their smoking status. The responses are either self-reported or answered by a family member and include four categories: “everyday smoker”, “some day smoker”, “former smoker”, and “never smoker”. Based on this information, we construct two smoking participation measures to reflect the respondent’s current smoking status. The first equals one if they report smoking every day and zero otherwise. The second equals one if respondents are either everyday smokers or someday smokers and zero otherwise. In other words, the second measure equals one if the individual smokes on at least some days, and hence is a more liberal definition of smoking participation. Before examining the effect on actual smoking participation, we study whether higher cigarette taxes motivate smokers to quit. We use information from the TUS-CPS to construct this variable. In particular, we label a smoker as

"motivated to quit" if, in the past twelve months, they have attempted to quit smoking or at least seriously considered it. We label a smoker as "not motivated to quit" if they have neither attempted to quit nor seriously considered quitting in the past twelve months. As shown in Table 1, the percentage of smokers who want to quit smoking is increasing from 2003 to 2007, while the participation rate among our sample decreases from 2003 to 2007.

3.3.2 Cigarette excise taxes

The independent variable of interest is monthly cigarette excise taxes. We use data on taxes from Orzechowski and Walker (2008) to construct monthly state real cigarette taxes denominated in 2003 dollars. We merge this information to the TUS-CPS which contains data on state of residence and date of interview.

There were several large tax increases on cigarettes between 2003 and 2007 at the state-level. Given that real cigarette taxes are virtually never reduced, these increases can be considered permanent, rather than temporary in nature. Table 2 lists monthly cigarette taxes across each state. As we can see from Table 2, between February 2003 and January 2007, thirty out of fifty states and Washington, DC increased their cigarette taxes. Over the entire period, and with respect to all increases, the average state excise tax increased from about 56 cents per pack to roughly 107 cents per pack, nearly doubling over this short period. Among them, the largest increase was 135 cents for Montana, while the smallest increase was 19 cents for Arkansas. Research in public finance suggests that the size and breadth of these tax increases were due to state budget shortfalls

following the 2001-2002 recession in the United States (Maag and Merriman, 2003). To the extent that these increases, and those that followed them, were indeed driven by budgetary concerns rather than state-specific anti-smoking sentiment, they likely represent more appropriate variation for estimating the causal effect of taxes on smoking behavior since they are more likely to represent pure changes in money price rather than the collective sentiment toward smoking in the given state.

3.3.3 Other independent variables

Other basic control variables include demographic variables, socioeconomic variables as well as indicators for state of residence and year and month of interview. With the exception of household size, individual-level covariates are specified as a series of categorical variables to ensure a flexible functional form. In particular, excluding reference categories we include one indicator for sex, three for race, three for education, six for income, two for marital status and one for each age. We also control for the state-specific monthly unemployment rate as a proxy for economic conditions, since previous work suggests that macroeconomic conditions may affect smoking behavior (Ruhm, 2005). Many studies suggest that cigarette tax rates are not random and they are closely related to the public sentiment towards smoking (cf., Warner, 1982; Hunter and Nelson, 1992; DeCicca et al., 2008). In particular, the concern is that unobserved public sentiment might bias the estimated effect of taxes away from zero if it is correlated with both the state level excise tax and smoking behaviors. In order to

deal with the different sentiment across state and time, DeCicca et al. (2008) used TUS-CPS data to construct a direct measure of monthly state anti-smoking sentiment basing on the survey questions about respondents' opinions towards smoking. We add this measure into analysis as a robustness check for our main results which rely on a two-way fixed effects approach which we describe in Section 4. Summary statistics for all variables are listed in Table 3.

3.3.4 Analysis sample

The 2003-2007 TUS-CPS files contain data on 483,557 individuals who have provided useable information related to smoking status. Limiting our sample to individuals aged 45 to 59 years old results in a sample of 129,565 individuals, including 73,007 never smokers (56.4 percent), 22,358 everyday smokers (17.3 percent), 3,883 someday smokers (3.0 percent), and 30,317 (23.4 percent) former smokers.[26] Depending on missing values of some regressors, sample size varies across model specifications, but these figures represent our primary analysis samples. Note that income is most likely to have missing information. Since income is missing relatively frequently, we include a separate categorical variable for those whose income information is missing in order to preserve sample size in our models. We employ complete case analysis for the remainder of the individual-level covariates.

3.4 Empirical strategy

[26] Following DeCicca and McLeod (2008), we exclude individuals who are older than 60 from our main analysis, because of the possibility of differential mortality after age sixty.

In this section, we describe our strategy for identifying the impact of state cigarette taxes on the smoker's intention to quit and their smoking participation. As noted elsewhere, cross-sectional estimates of this relationship may be subject to serious omitted variables bias due to heterogeneity that may be correlated with state level cigarette excise taxes and smoking behaviors. For example, if we merely estimated cross-sectional models our estimates may be confounded by state-level anti-smoking sentiment that is correlated with the tax rate and also the smoking behaviors. Put differently, we are concerned that individuals are unobservably different in higher tax states than in lower tax jurisdictions in ways that might lead cross-sectional estimation to yield biased results. For example, individuals in higher tax states have stronger sentiments against smoking than those in the lower tax states, so the lower smoking participation in the higher tax states may not be due to their high taxes per se, but because of their stronger anti-smoking sentiment.

Because we use repeated cross-sectional data, we deal with this issue by including state-fixed effects in our regressions. To the extent that the heterogeneities in question are time-invariant, state-fixed effects will purge the correlation between the error term and the tax variable, eliminating potential omitted variables bias. In essence, the identification strategy is to compare the fraction of smokers before and after tax increases within states, though we allow the treatment (i.e., the amount of the tax) to be a continuous variable. More

specifically, we estimate two-way fixed effects specifications of the following form:

$$S_{ijmt} = \alpha + \beta_1 T_{jmt} + \beta_2 X_{ijmt} + \delta Z_{jmt} + \sigma_j + \mu_m + \tau_t + \varepsilon_{ijmt} \qquad (3)$$

where, i represents the individual, j stands for state, m is survey month and t is survey year. σ, μ, τ are state, month and year fixed effects, respectively. S represents smoking participation or the quitting intention measures, T represents monthly state-specific cigarette excise taxes in constant dollars, and X includes all individual-level demographic and socio-economic measures. Z represents state-specific monthly unemployment rates to proxy for economic conditions. Standard errors are clustered at the state-level to deal with within-state correlation in the errors.

Though they are useful in accounting for relevant heterogeneity, state, month and year fixed effects may absorb a great deal of variation in our cigarette excise tax variable. That is, there might be little identifying variation remaining after including fixed effects due to collinearity between taxes and the fixed effects, specifically the state dummy variables. To assess the collinearity between tax and the set of fixed effects, we compute the variance inflation factor (VIF) which is $1/(1-R^2)$, where R^2 is computed from an auxiliary regression of the cigarette excise tax on the full set of fixed effects. Conventionally, it is assumed that if the variance inflation factor is greater than ten there is not enough independent variation left to perform the estimation (cf., Kennedy, 1994). In our case, it is roughly eight. While this is close to the acceptable level, it suggests a relatively

high degree of within-state and cross-state variation in cigarette taxes, especially in comparison with studies that use earlier data with much less within state variation. In the next section, we present our estimates.

3.5 Estimates

3.5.1 Main results

In this section, we present estimates from empirical models of cigarette smoking as specified in equation (3). Table 4 contains our main estimates for two groups of older adults: individuals 45-59 years of age and individuals 45-64 years of age[27]. For each age group, we estimate three models. Model (1) estimates the impact of cigarette taxes on older smokers' quitting intentions. Recall that a smoker is defined as "motivated to quit" if they have either attempted to quit or seriously considered quitting in the past twelve months. Models (2) and (3) present estimates of everyday and at least someday smoking, respectively.

The first three columns present estimates for the individuals aged 45-59 years old. As can be seen from column 1, higher cigarette excise taxes are significantly associated with an increased intention of quitting. In particular, our estimate implies that a one-dollar increase in state cigarette tax rates will increase the proportion of smokers who want to quit smoking by nearly 4.8 percentage points. Relative to a base of 41.5 percent, this translates into a nearly 12.0 percent increase in the fraction of smokers with the intentions of quitting smoking. Although we prefer to report our results in this manner, we also present the price

[27] We restrict our sample to individuals less than 65, because the health benefits of smoking cessation decrease substantially with age.

participation elasticities since they are common in the literature. The corresponding elasticity for quitting intention is 0.42, which means that a one percent increase in the per-pack cigarette price will cause the fraction of smokers who want to quit smoking to increase by 0.42 percent.

The estimates for daily smoking participation and smoking on at least some days are consistently negative. Estimates in column 2 indicate that a one-dollar increase in tax will reduce the daily smoking prevalence by around 1.2 percentage points. Relative to an average daily smoking rate of 17.4 percent, this reduction implies a 7 percent decrease in the fraction of daily smokers. The elasticity corresponding to this estimate is -0.29 which is large in comparison with the general consensus in the literature (Evans and Farrelly, 1998; Farrelly et. al., 2001). Estimates for smoking on at least some days, as reported in column 3, show that a one-dollar increase in tax reduces this prevalence by roughly 1 percentage point. Relative to a base of 20 percent, this reduction implies a 5 percent reduction in the prevalence of some-day smoking. Contrary to most of the literature, our main findings provide some evidence that older adults are sensitive to tax-induced increases in cigarette price.

Estimates for a slightly larger group, those aged 45-64, which are presented from column 3 to column 6, have similar pattern as those for aged 45-59 but are smaller in magnitude suggesting less price-sensitivity among those 60-64 years old. With respect to the quitting intention, we find that a one-dollar increase in tax causes the fraction of smokers who want to quit smoking to increase by 9

percent, with a corresponding elasticity of 0.32. For daily smokers, our estimate implies a one-dollar increase reduces the proportion of daily smokers by around 6 percent, with a corresponding price participation elasticity of -0.25. Again, as with the 45-59 year old sample, the impact for smoking on at least some days is smaller in magnitude than that associated with daily smoking. The estimated effect implies that a one-dollar increase reduces the prevalence of such smokers by around 5 percent, with a corresponding price participation elasticity of -0.18.

As discussed in DeCicca et al. (2008), price participation elasticities can be explained as weighted average of initiation and cessation elasticities. With respect to youth, initiation dominates smoking participation, while for older adults, the group we study, it is likely that reduction in smoking participation represents quitting behavior. Overall from Table 4, we find sizeable impacts of cigarette tax increases on smokers' greater desire to quit; consistent with these stronger quitting intentions, we find consistent and sizeable reductions in daily and at least some-day smoking participation in response to cigarette tax increases, compared with the literature. In the next subsection, we examine whether there exists heterogeneity in these relatively large tax effects.

3.5.2 Heterogeneous tax effects

The cigarette tax is often criticized because it is a regressive tax where poorer people, who are much more likely to smoke, pay a relatively larger fraction of their disposable income. In Table 5, we present the estimates of quitting intention and smoking participation by different levels of personal income for the 45-59-

year old sample[28]. The first three columns report the estimates for respondents whose personal income is less than $25,000, while the last three columns contains individuals whose corresponding income is at least $25,000. As we can see, lower-income individuals are found to quit smoking at a much higher rate in response to higher taxes than their counterparts with higher personal income. With respect to the intention to quit, we find that a one-dollar increase in the cigarette tax will increase the fraction of smokers who intend to quit by 7.6 percentage points. Relative to a base of 41.6 percent, this translates into about an 18 percent increase in the desire to quit, with a corresponding elasticity of 0.64. In comparison with the effects for poorer smokers, the tax has a much smaller effect on higher-income smokers' intentions to quit. Here, a one-dollar increase is associated with a 4 percent increase in the intention to quit with a corresponding elasticity of 0.15, which is less than one-fourth that of the poorer smokers.

As noted, lower-income people are much more likely to smoke than their higher income peers. With respect to daily smoking participation, 22.4 percent of lower-income people are smokers, while 12.7 percent of the higher-income people, as we define them, are smokers. It is likely that the cigarette tax is effective in smoking reduction because poorer smokers are much more responsive to tax changes due to their lower levels of disposable income. The estimate in column 2 indicates that a one-dollar increase in the cigarette tax will reduce daily smoking participation by 2.6 percentage points. Relative to a base of 22.4 percent, this tax

[28] The corresponding estimates for those aged 45-64 are presented in Appendix 1. Appendix 1 has the same pattern as that from Table 4, but consistent with our main estimates, the estimated impacts are slightly smaller than those in Table 4.

increase reduces participation by nearly 12 percent, with a corresponding price participation elasticity of -0.44. However, the counterpart for higher-income respondents is much smaller in magnitude and is not statistically different from zero. The pattern for smoking participation on at least some days is similar as that for the daily smoking participation, but is even smaller, consistent with the notion that the observed patterns are being driven by an income effect. These estimates are also consistent with Gruber and Koszegi (2004) who reported that lower income individuals are more sensitive to price and provide some evidence that the tax burden for low-income individuals is bigger than that for high-income people. More directly, they also correspond to the findings of DeCicca and McLeod (2008).

Since income at a point in time may not be representative of the true purchasing power of individuals, especially older individuals, we also estimate our main models by educational attainment. Estimates for 45-59-year olds by education levels are reported in Table 6[29]. Consistent with our findings by income level, less-educated individuals, defined as those with a high school education or less, are much more price-sensitive than their more-educated counterparts. In particular, we find that a one-dollar increase in the cigarette tax will increase the fraction of smokers who want to quit by 16.8 percent (corresponding elasticity +0.59) while more-educated individuals increase the willingness to quit by only 6.5 percent (corresponding elasticity 0.24). With respect to smoking participation,

[29] The corresponding estimates for a slightly larger group which includes those aged 45-64 are presented in Appendix 2. Consistent with corresponding analyses by personal income, Appendix 2 has the similar pattern as that from Table 5, but the estimated impacts are smaller than those in Table 5.

our estimates imply that a one-dollar increase in tax will reduce the participation in daily smoking and smoking on at least some days by 14.7 percent (corresponding price participation elasticity -0.54) and 11.8 percent (corresponding price participation elasticity -0.43) respectively, while there is no evidence that the cigarette tax impacts the smoking participation of more-educated individuals. The strong effect on the smoking intentions and behaviour of less-educated individuals might suggest additional health gains from smoking cessation, to the extent that lower SES individuals are more susceptible to ill-health and have issues of health care access.

Finally, we allow the effect of cigarette taxes on smoking behavior to vary by sex, since this is a conventional breakdown in the literature[30]. As we can see from Table 7, while somewhat larger for males, both genders are basically similarly affected by taxes. In particular, we can only find significant impact of taxes on the intentions to quit among males.

3.6 Conclusion

Using the repeated cross-sectional data from TUS-CPS for three years over the period 2003-2007, this paper estimates several cigarette demand models for older adults. Existing work commonly finds that price sensitivity of smoking varies inversely with age, which generally suggests that old adults are insensitive to price changes. However, contrary to most of this literature, we find large responses among older adults in response to tax increases, especially for those

[30] The corresponding estimates for a slightly larger group which includes those aged 45-64 are presented in Appendix 3. Compared with Table 7, estimates in Appendix 3 are a bit smaller.

who are lower-income and less-educated. In particular, we find a one-dollar increase in the cigarette tax results in a 16.8 percent increase in the fraction of smokers who want to quit, a 14.7 percent decrease in the participation rate for daily smokers and a 11.8 percent reduction in the participation rate for smokers who smoke on at least some days for the less-educated individuals.

This paper provides new evidence to the literature on older smokers' quitting behaviors when taxes increase. Since lots of work suggest that quitting smoking even at old age can increase health, our study implies that tax increases may generate substantial health benefits for this group.

References

Carpenter, C., and P. J. Cook (2008), "Cigarette taxes and youth smoking: New evidence from national, state, and local Youth Risk Behavior Surveys", *Journal of health Economics*, 27: 287–299.

Chaloupka, F.J., and R.L. Pacula (1999), "Sex and race differences in youth smoking responsiveness to price and tobacco control policies", Tobacco Control, 8(4): 373-377.

Chaloupka, F., and K. Warner (2000), "The economics of smoking", In: Cuyler, A., Newhouse, J. (Eds.), Handbook of Health Economics. North-Holland, Elsevier Science, Netherlands.

Critchley, J.A., and S. Capewell (2003), "Mortality risk reduction associated with smoking cessation in patients with coronary heart disease: A systematic review", *Journal of American Medical Association*, 290: 86-97.

DeCicca, P., D. Kenkel, and A. Mathios (2002), "Putting Out the Fires: Will Higher Taxes Reduce the Onset of Youth Smoking?", *Journal of Political Economy* 110 (1): 144-169.

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

DeCicca, P., and L. McLeod (2008), "Cigarette taxes and older adult smoking: Evidence from recent large tax increases", *Journal of Health Economics*, 27(4): 918-929.

Evans, W.N., and M.C. Farrelly (1998), "The compensating behavior of smokers: Taxes, tar and nicotine", *RAND Journal of Economics*, 29(3):578-595.

Evans, W.N., and L.X. Huang (1998), "Cigarette taxes and teen smoking: New evidence from panels of repeated cross-sections", Working Paper (Department of Economics, University of Maryland).

Farrelly, M.C., and J.W. Bray (1998), "Office on Smoking and Health. Response to increases in cigarette prices by race/ethnicity, income, and age groups - United States, 1976-1993", *Morbidity and Mortality Weekly Report*, 47(29):605-609.

Farrelly, M.C., J.W. Bray, T. Pechacek, T. Woollery (2001), "Responses by adults to increases in cigarette prices by sociodemographic characteristics", *Southern Economic Journal*, 68 (1): 156–165.

Grossman, M., D. Coate, E.M. Lewit and R.A. Shakotko (1983), Economic and Other Factors in Youth Smoking (National Science Foundation, Washington).

Gruber, J., and B. Koszegi (2004), "Tax incidence when individuals are time-inconsistent: the case of cigarette taxes", *Journal of Public Economics*, 88: 1959–1988.

Gruber, J., and J. Zinman (2001), "Youth smoking in the United States: evidence and implications", In: Gruber, J. (Ed.), *Risky Behavior Among Youths: An Economic Analysis. University of Chicago Press, Chicago*, 69-120.

Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office of Smoking and Health. DHHS Publication No. (CDC) 90-8416.

Hunter, W.J., M.A. Nelson (1992), "The political economy of state tobacco taxation", *Public Finance/Finances Publiques*, 47(2): 214–228.

Kennedy, P. (1994), "A Guide to Econometrics, 3rd ed". *The MIT Press*, Cambridge, MA.

Lewit, E.M., and D. Coate (1982), "The potential for using excise taxes to reduce smoking", *Journal of Health Economics*, 1(2):121-145.

Lewit, E.M., D. Coate and M. Grossman (1981), "The effects of government regulation on teenage smoking", *Journal of Law and Economics*, 24(3):545-569.

Lewit, E.M., A. Hyland, N. Kerrebrock and K.M. Cummings (1997), "Price, public policy and smoking in young people", Tobacco Control 6(S2):17-24.

Maag, E., and D. Merriman (2003), "Tax policy responses to revenue shortfalls", State Tax Notes (August 4).

Orzechowski, W.P., and R. Walker (2008), "The tax burden on tobacco", Arlington, VA.

Ostbye, T., and D.H. Taylor (2004), "The effect of smoking on 'Years of Healthy Life' lost among middle-aged and older Americans", *Health Services Research*, 39 (3): 531–551.

Ostbye, T., D.H. Taylor, and S. Jung (2002), "A longitudinal study of the effects of tobacco smoking and other modifiable risk factors on ill health in middle-aged and old Americans: results from the health and retirement study and asset and health dynamics among the oldest old survey", *Preventive Medicine*, 34: 334–345.

Royal College of Physicians (1962), "Smoking and health", *Summary and Report of the Royal College of Physicians of London on Smoking in Relation to Cancer of the Lung and Other Diseases* (Pitman Publishing, New York).

Ruhm, C.J. (2005), "Healthy living in hard times", *Journal of Health Economics*, 24 (2): 341–363.

Sloan, F., and J. Trogdon (2004), "The impact of the master settlement agreement on cigarette consumption", *Journal of Policy Analysis and Management*, 23 (4): 843–855.

Tauras, J.A., (2006), "Smoke-free air laws, cigarette prices and adult cigarette demand", *Economic Inquiry* 44 (2), 333–342.

Tauras, J.A., and EJ. Chaloupka (1999), "Price, clean indoor air laws, and cigarette smoking: Evidence from longitudinal data for young adults", Working Paper Number 6937 NBER.

Taylor, D.H., V. Hasselblad, J. Henley, M.J. Thurn, F.A. Sloan (2002), "Benefits of smoking cessation for longevity", *American Journal of Public Health*, 92: 990–996.

US Department of Health, Education, and Welfare (1964), "Smoking and health", *Report of the Advisory Committee to the Surgeon General of the Public Health Service, PHS Publication* No. 1103 (U.S. Department of Health, Education, and Welfare, Public Health Service, Center for Disease Control).

U.S. Department of Health and Human Services (1990), "The health benefits of smoking cessation: a report of the Surgeon General", Atlanta: U.S. Department of Health and

Warner, K.E (1982), "Cigarette excise taxation and interstate smuggling: an assessment of recent activity", *National Tax Journal*, 35: 483–490.

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

Table 1. Descriptive statistics for intention to quit and participation by year.

	2003	2006	2007
Intentions to quit	40.00	43.02	42.89
Smoking on all days	17.60	17.20	16.48
Smoking on at least some days	20.48	20.40	19.41

Notes: Each cell presents the percent of smokers who want to quit or the participation rate for each year.

Table 2. Monthly state real per-pack cigarette taxes denominated in February 2003 dollars.

State	Tax 2/2003	Tax 6/2003	Tax 11/2003	Tax 5/2006	Tax 8/2006	Tax 1/2007
AL	0.17	0.17	0.16	0.39	0.38	0.38
AK	1.00	1.00	0.99	1.46	1.62	1.63
AZ	1.18	1.18	1.17	1.08	1.06	1.07
AR	0.34	0.59	0.59	0.54	0.53	0.53
CA	0.87	0.87	0.86	0.79	0.78	0.79
CO	0.20	0.20	0.20	0.77	0.76	0.76
CT	1.11	1.51	1.50	1.38	1.36	1.36
DE	0.24	0.24	0.55	0.50	0.50	0.50
DC	1.00	1.00	0.99	0.91	0.90	0.90
FL	0.34	0.34	0.34	0.31	0.31	0.31
GA	0.12	0.12	0.37	0.34	0.33	0.33
HI	1.20	1.20	1.29	1.28	1.26	1.44
ID	0.28	0.57	0.57	0.52	0.51	0.51
IL	0.98	0.98	0.97	0.89	0.88	0.88
IN	0.56	0.56	0.55	0.51	0.50	0.50
IA	0.36	0.36	0.36	0.33	0.32	0.33
KS	0.79	0.79	0.78	0.72	0.71	0.71
KY	0.03	0.03	0.03	0.27	0.27	0.27
LA	0.36	0.36	0.36	0.33	0.32	0.33
ME	1.00	1.00	0.99	1.82	1.80	1.81
MD	1.00	1.00	0.99	0.91	0.90	0.90
MA	1.51	1.51	1.50	1.38	1.36	1.36
MI	1.25	1.25	1.24	1.82	1.80	1.81
MN	0.48	0.48	0.48	1.35	1.34	1.34
MS	0.18	0.18	0.18	0.16	0.16	0.16
MO	0.17	0.17	0.17	0.15	0.15	0.15
MT	0.18	0.70	0.69	1.55	1.53	1.53
NE	0.64	0.64	0.64	0.58	0.58	0.58
NV	0.35	0.35	0.79	0.73	0.72	0.72
NH	0.52	0.52	0.52	0.73	0.72	0.72
NJ	1.50	1.50	2.03	2.19	2.32	2.32
NM	0.21	0.21	0.90	0.83	0.82	0.82
NY	1.50	1.50	1.49	1.37	1.35	1.35
NC	0.05	0.05	0.05	0.27	0.32	0.32
ND	0.44	0.44	0.44	0.40	0.40	0.40
OH	0.55	0.55	0.55	1.14	1.13	1.13
OK	0.23	0.23	0.23	0.94	0.93	0.93
OR	1.28	1.28	1.27	1.08	1.06	1.07

PA	1.00	1.00	0.99	1.23	1.22	1.22
RI	1.32	1.32	1.70	2.24	2.22	2.22
SC	0.07	0.07	0.07	0.06	0.06	0.06
SD	0.33	0.53	0.53	0.48	0.48	1.38
TN	0.20	0.20	0.20	0.18	0.18	0.18
TX	0.41	0.41	0.41	0.37	0.37	1.27
UT	0.70	0.70	0.69	0.63	0.63	0.63
VT	0.93	0.93	1.18	1.08	1.61	1.62
VA	0.03	0.03	0.02	0.27	0.27	0.27
WA	1.43	1.43	1.41	1.85	1.83	1.83
WV	0.17	0.55	0.55	0.50	0.50	0.50
WI	0.77	0.77	0.76	0.70	0.69	0.70
WY	0.12	0.12	0.60	0.55	0.54	0.54

Notes: Tax is for a single pack of twenty cigarettes and is rounded to the nearest penny.

Table 3. Summary statistics for years from 2003 to 2007

Variable	Mean
Panel A: *Dependent Variables*	
Smokers on all days	0.174
Smokers on some days	0.203
Smokers who have intentions to quit smoking	0.415
Panel B: *Independent Variables*	
Real cigarette tax ($ per pack)	0.815
Monthly state unemployment rate	0.051
Female	0.520
Age	51.6
Hispanic	0.071
Non- Hispanic white	0.790
Non- Hispanic black	0.082
Others	0.057
Less than high school	0.098
High school	0.311
Some college	0.178
College or higher	0.413
Household Income <15k	0.078
Household Income 15k-25k	0.067
Household Income 25k-35k	0.088
Household Income 35k-50k	0.127
Household Income 50k-75k	0.196
Household Income 75k+	0.319
Household Income missing	0.126
Married	0.695
Divorced, widowed, or separated	0.213
Never married	0.092
Employed	0.777
Unemployed	0.029
Retired	0.046
Not in the labor force	0.148
Household size	2.71

Table 4. Estimates of the impact of cigarette excise taxes on smoking participation.

	Age: 45-59			Age: 45-64		
VAR	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0481** (0.0196) {0.42}	-0.0122* (0.0076) {-0.29}	-0.0103 (0.0078) {-0.20}	0.0368** (0.0187) {0.32}	-0.0101 (0.0075) {-0.25}	-0.0088 (0.0077) {-0.18}
Monthly state unemployment rate	0.0079 (0.0086)	0.0005 (0.0038)	0.0011 (0.0035)	0.0046 (0.0064)	0.0008 (0.0032)	0.0017 (0.0032)
Female	0.0567*** (0.0074)	-0.0515*** (0.0024)	-0.0572*** (0.0031)	0.0580*** (0.0061)	-0.0504*** (0.0025)	-0.0562*** (0.0032)
White	-0.1135*** (0.0232)	0.1103*** (0.0053)	0.1052*** (0.0061)	-0.1054*** (0.0198)	0.1063*** (0.0047)	0.1027*** (0.0054)
Black	-0.0093 (0.0211)	0.0940*** (0.0157)	0.0945*** (0.0147)	-0.0035 (0.0171)	0.0898*** (0.0152)	0.0920*** (0.0138)
Other Race	-0.0373 (0.0242)	0.1295*** (0.0146)	0.1184*** (0.0158)	-0.0349* (0.0201)	0.1225*** (0.0134)	0.1140*** (0.0147)
High school	-0.0075 (0.0113)	-0.0369*** (0.0050)	-0.0410*** (0.0064)	0.0015 (0.0115)	-0.0352*** (0.0045)	-0.0383*** (0.0058)
Some college	0.0490*** (0.0116)	-0.0585*** (0.0052)	-0.0616*** (0.0067)	0.0496*** (0.0102)	-0.0547*** (0.0051)	-0.0569*** (0.0064)
College or Higher	0.0706*** (0.0131)	-0.1476*** (0.0053)	-0.1601*** (0.0069)	0.0766*** (0.0111)	-0.1379*** (0.0050)	-0.1481*** (0.0063)
Family income 15-25k	0.0110 (0.0154)	-0.0166*** (0.0038)	-0.0251*** (0.0043)	0.0065 (0.0146)	-0.0153*** (0.0032)	-0.0247*** (0.0037)
Family income 25-35k	0.0122 (0.0150)	-0.0218*** (0.0043)	-0.0289*** (0.0051)	0.0097 (0.0141)	-0.0225*** (0.0035)	-0.0307*** (0.0043)
Family income 35-50k	0.0055 (0.0138)	-0.0352*** (0.0045)	-0.0449*** (0.0051)	0.0057 (0.0131)	-0.0345*** (0.0034)	-0.0450*** (0.0040)
Family income 50-75k	0.0251* (0.0133)	-0.0530*** (0.0045)	-0.0637*** (0.0046)	0.0232* (0.0131)	-0.0523*** (0.0036)	-0.0640*** (0.0037)

Family income 75k+	0.0208 (0.0151)	-0.0934*** (0.0045)	-0.1053*** (0.0043)	0.0209 (0.0151)	-0.0898*** (0.0038)	-0.1028*** (0.0034)
Unmarried	-0.0082 (0.0078)	0.0860*** (0.0038)	0.1026*** (0.0041)	-0.0046 (0.0076)	0.0842*** (0.0032)	0.1006*** (0.0032)
Never married	-0.0328** (0.0129)	0.0176*** (0.0038)	0.0226*** (0.0045)	-0.0378*** (0.0117)	0.0190*** (0.0034)	0.0240*** (0.0039)
Unemployed	0.0484*** (0.0167)	0.0675*** (0.0072)	0.0936*** (0.0084)	0.0518*** (0.0149)	0.0610*** (0.0063)	0.0846*** (0.0071)
Retired	0.0112 (0.0205)	0.0029 (0.0061)	0.0048 (0.0064)	0.0215* (0.0116)	0.0076* (0.0039)	0.0091** (0.0040)
Not in labor force	0.0530*** (0.0083)	0.0329*** (0.0041)	0.0405*** (0.0044)	0.0544*** (0.0083)	0.0326*** (0.0036)	0.0401*** (0.0039)
Family size	-0.0033 (0.0029)	-0.0077*** (0.0011)	-0.0109*** (0.0011)	-0.0045 (0.0028)	-0.0057*** (0.0010)	-0.0085*** (0.0011)
Dep. Mean	0.415	0.174	0.203	0.412	0.167	0.196
N	20027	129565	129565	23824	158786	158786

Notes: Tax effects are estimated for different age groups. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Table 5. Estimates of the impact of cigarette excise taxes on smoking participation, by personal income

	Personal income: 0-25k			Personal income: 25k+		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0761**	-0.0259**	-0.0185	0.0168	0.0021	0.0022
	(0.0308)	(0.0116)	(0.0128)	(0.0270)	(0.0110)	(0.0115)
	{0.64}	{-0.44}	{-0.27}	{0.15}	{0.07}	{0.06}
Monthly state unemployment rate	0.0128	-0.0015	-0.0037	-0.0011	0.0055	0.0082*
	(0.0150)	(0.0053)	(0.0056)	(0.0136)	(0.0044)	(0.0046)
Female	0.0615***	-0.0693***	-0.0762***	0.0432***	-0.0372***	-0.0420***
	(0.0113)	(0.0043)	(0.0055)	(0.0097)	(0.0026)	(0.0028)
Non-Hispanic white	-0.1338***	0.1629***	0.1573***	-0.0767***	0.0455***	0.0324***
	(0.0228)	(0.0078)	(0.0087)	(0.0288)	(0.0045)	(0.0068)
Non- Hispanic black	-0.0281	0.1328***	0.1417***	0.0205	0.0220*	0.0134
	(0.0235)	(0.0202)	(0.0194)	(0.0377)	(0.0116)	(0.0114)
Other Races	-0.0466*	0.1708***	0.1704***	-0.0185	0.0602***	0.0357***
	(0.0246)	(0.0200)	(0.0210)	(0.0373)	(0.0106)	(0.0126)
High school	-0.0139	-0.0431***	-0.0462***	0.0411	-0.0459***	-0.0518***
	(0.0127)	(0.0065)	(0.0070)	(0.0295)	(0.0063)	(0.0070)
Some college	0.0473***	-0.0677***	-0.0683***	0.0888***	-0.0621***	-0.0685***
	(0.0153)	(0.0075)	(0.0088)	(0.0309)	(0.0056)	(0.0066)
College or higher	0.0472***	-0.1508***	-0.1633***	0.1264***	-0.1648***	-0.1776***
	(0.0145)	(0.0077)	(0.0089)	(0.0305)	(0.0090)	(0.0096)
Family income 15k-25k	0.0135	-0.0232***	-0.0330***			
	(0.0155)	(0.0045)	(0.0050)			
Family income 25k-35k	0.0024	-0.0267***	-0.0354***			
	(0.0186)	(0.0061)	(0.0068)			
Family income 35k-50k	0.0088	-0.0450***	-0.0573***	-0.0355	-0.0180***	-0.0200***
	(0.0162)	(0.0059)	(0.0068)	(0.0245)	(0.0066)	(0.0076)
Family income 50k-75k	0.0411**	-0.0801***	-0.0926***	-0.0148	-0.0102*	-0.0127**
	(0.0175)	(0.0057)	(0.0059)	(0.0220)	(0.0053)	(0.0062)

Family income 75k+	-0.0071 (0.0324)	-0.0873*** (0.0088)	-0.0977*** (0.0100)	-0.0140 (0.0274)	-0.0478*** (0.0061)	-0.0501*** (0.0068)
Divorced, widowed, or separated	-0.0099 (0.0105)	0.0981*** (0.0058)	0.1127*** (0.0063)	0.0184 (0.0177)	0.0651*** (0.0056)	0.0834*** (0.0056)
Never married	-0.0369** (0.0168)	0.0099 (0.0077)	0.0155* (0.0084)	-0.0191 (0.0204)	0.0201*** (0.0058)	0.0224*** (0.0064)
Unemployed	0.0557*** (0.0201)	0.0695*** (0.0103)	0.0968*** (0.0112)	0.0477 (0.0321)	0.0580*** (0.0100)	0.0857*** (0.0132)
Retired	0.0223 (0.0318)	-0.0078 (0.0079)	-0.0018 (0.0086)	-0.0271 (0.0293)	0.0036 (0.0075)	0.0021 (0.0089)
Not in labor force	0.0547*** (0.0111)	0.0383*** (0.0050)	0.0454*** (0.0055)	0.0663*** (0.0246)	0.0162*** (0.0059)	0.0222*** (0.0066)
Family size	-0.0036 (0.0038)	-0.0100*** (0.0016)	-0.0131*** (0.0015)	0.0083 (0.0098)	-0.0109*** (0.0021)	-0.0148*** (0.0025)
Dep. mean	0.416	0.224	0.259	0.425	0.127	0.153
N	11132	55616	55616	6961	57639	57639

Notes: Heterogeneous tax effects are estimated across samples with different personal income levels. All samples include individuals aged 45-59. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Table 6. Estimates of the impact of cigarette excise taxes on smoking participation, by education

	High school or less			Some college or higher		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0649**	-0.0356***	-0.0327***	0.0292	0.0018	0.0032
	(0.0310)	(0.0112)	(0.0116)	(0.0243)	(0.0080)	(0.0080)
	{0.59}	{-0.54}	{-0.43}	{0.24}	{0.06}	{0.09}
Monthly state unemployment rate	0.0187	0.0023	0.0034	-0.0041	-0.0007	-0.0003
	(0.0126)	(0.0059)	(0.0054)	(0.0115)	(0.0036)	(0.0035)
Female	0.0661***	-0.0692***	-0.0766***	0.0447***	-0.0393***	-0.0434***
	(0.0102)	(0.0053)	(0.0061)	(0.0102)	(0.0020)	(0.0024)
Non-Hispanic white	-0.1215***	0.1818***	0.1730***	-0.0984**	0.0507***	0.0406***
	(0.0202)	(0.0096)	(0.0102)	(0.0396)	(0.0041)	(0.0066)
Non- Hispanic black	-0.0168	0.1311***	0.1376***	-0.0015	0.0420***	0.0340***
	(0.0205)	(0.0222)	(0.0200)	(0.0420)	(0.0094)	(0.0104)
Other Races	-0.0852***	0.1785***	0.1618***	0.0225	0.0556***	0.0458***
	(0.0299)	(0.0225)	(0.0216)	(0.0412)	(0.0089)	(0.0143)
High school	-0.0047	-0.0657***	-0.0706***			
	(0.0111)	(0.0059)	(0.0070)			
College or higher				0.0190*	-0.0760***	-0.0864***
				(0.0111)	(0.0032)	(0.0034)
Family income 15k-25k	0.0142	-0.0250***	-0.0327***	0.0083	-0.0132**	-0.0234***
	(0.0208)	(0.0070)	(0.0078)	(0.0189)	(0.0059)	(0.0067)
Family income 25k-35k	-0.0084	-0.0333***	-0.0407***	0.0519**	-0.0191***	-0.0271***
	(0.0194)	(0.0074)	(0.0079)	(0.0241)	(0.0049)	(0.0056)
Family income 35k-50k	0.0145	-0.0497***	-0.0609***	0.0040	-0.0318***	-0.0418***
	(0.0156)	(0.0073)	(0.0081)	(0.0251)	(0.0051)	(0.0059)
Family income 50k-75k	0.0139	-0.0666***	-0.0759***	0.0475**	-0.0497***	-0.0623***
	(0.0211)	(0.0076)	(0.0075)	(0.0216)	(0.0043)	(0.0048)
Family income 75k+	0.0000	-0.0997***	-0.1093***	0.0485**	-0.0897***	-0.1040***
	(0.0229)	(0.0085)	(0.0084)	(0.0221)	(0.0056)	(0.0058)

Divorced, widowed, or separated	-0.0106 (0.0105)	0.1147*** (0.0056)	0.1274*** (0.0059)	-0.0032 (0.0110)	0.0623*** (0.0039)	0.0801*** (0.0045)
Never married	-0.0330* (0.0177)	0.0154** (0.0063)	0.0188*** (0.0070)	-0.0312 (0.0209)	0.0170*** (0.0046)	0.0225*** (0.0058)
Unemployed	0.0338 (0.0209)	0.0913*** (0.0122)	0.1120*** (0.0138)	0.0743*** (0.0251)	0.0459*** (0.0066)	0.0741*** (0.0087)
Retired	0.0240 (0.0276)	0.0031 (0.0106)	0.0022 (0.0106)	-0.0041 (0.0280)	0.0008 (0.0059)	0.0040 (0.0071)
Not in labor force	0.0525*** (0.0113)	0.0305*** (0.0064)	0.0372*** (0.0064)	0.0567*** (0.0130)	0.0389*** (0.0045)	0.0472*** (0.0051)
Family size	-0.0050 (0.0039)	-0.0068*** (0.0016)	-0.0100*** (0.0019)	-0.0017 (0.0049)	-0.0078*** (0.0013)	-0.0107*** (0.0014)
Dep. mean	0.388	0.245	0.279	0.447	0.123	0.149
N	11007	52965	52965	9020	76600	76600

Notes: Heterogeneous tax effects are estimated across samples with different education levels. All sample include individuals aged 45-59. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Table 7. Estimates of the impact of cigarette excise taxes on smoking participation, by gender

	Men			Women		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0592*	-0.0122	-0.0127	0.0360	-0.0123	-0.0082
	(0.0331)	(0.0084)	(0.0090)	(0.0259)	(0.0091)	(0.0090)
	{0.55}	{-0.26}	{-0.22}	{0.29}	{-0.33}	{-0.18}
Monthly state unemployment rate	0.0233*	0.0045	0.0081*	-0.0066	-0.0032	-0.0052
	(0.0135)	(0.0044)	(0.0044)	(0.0124)	(0.0046)	(0.0042)
Non-Hispanic white	-0.1072***	0.1046***	0.0904***	-0.1175***	0.1154***	0.1185***
	(0.0271)	(0.0058)	(0.0064)	(0.0310)	(0.0062)	(0.0075)
Non- Hispanic black	-0.0230	0.0947***	0.0928***	0.0039	0.0987***	0.1026***
	(0.0249)	(0.0155)	(0.0134)	(0.0323)	(0.0203)	(0.0203)
Other Races	-0.0363	0.1494***	0.1286***	-0.0374	0.1137***	0.1125***
	(0.0284)	(0.0155)	(0.0126)	(0.0357)	(0.0209)	(0.0249)
High school	-0.0101	-0.0418***	-0.0468***	-0.0031	-0.0324***	-0.0353***
	(0.0154)	(0.0051)	(0.0059)	(0.0154)	(0.0070)	(0.0088)
Some college	0.0537***	-0.0643***	-0.0675***	0.0463**	-0.0539***	-0.0567***
	(0.0157)	(0.0057)	(0.0068)	(0.0184)	(0.0062)	(0.0081)
College or higher	0.0815***	-0.1663***	-0.1815***	0.0630***	-0.1315***	-0.1413***
	(0.0175)	(0.0060)	(0.0070)	(0.0187)	(0.0069)	(0.0087)
Family income 15k-25k	-0.0030	-0.0097	-0.0158*	0.0253	-0.0218***	-0.0320***
	(0.0211)	(0.0080)	(0.0086)	(0.0213)	(0.0042)	(0.0045)
Family income 25k-35k	0.0302	-0.0187**	-0.0253***	0.0004	-0.0242***	-0.0315***
	(0.0217)	(0.0077)	(0.0080)	(0.0189)	(0.0051)	(0.0064)
Family income 35k-50k	0.0129	-0.0346***	-0.0413***	0.0046	-0.0349***	-0.0467***
	(0.0207)	(0.0070)	(0.0072)	(0.0157)	(0.0047)	(0.0055)
Family income 50k-75k	0.0317*	-0.0517***	-0.0609***	0.0239	-0.0528***	-0.0643***
	(0.0192)	(0.0073)	(0.0080)	(0.0177)	(0.0050)	(0.0048)
Family income 75k+	0.0296	-0.0999***	-0.1087***	0.0179	-0.0855***	-0.0992***
	(0.0218)	(0.0064)	(0.0063)	(0.0200)	(0.0053)	(0.0050)

Divorced, widowed, or separated	-0.0121 (0.0116)	0.1043*** (0.0059)	0.1214*** (0.0061)	-0.0032 (0.0108)	0.0742*** (0.0047)	0.0908*** (0.0052)
Never married	-0.0327** (0.0160)	0.0177*** (0.0052)	0.0244*** (0.0064)	-0.0294 (0.0204)	0.0212*** (0.0055)	0.0247*** (0.0056)
Unemployed	0.0648*** (0.0196)	0.0775*** (0.0112)	0.0998*** (0.0118)	0.0318 (0.0248)	0.0554*** (0.0093)	0.0859*** (0.0103)
Retired	0.0117 (0.0305)	-0.0014 (0.0083)	-0.0016 (0.0087)	0.0129 (0.0243)	0.0058 (0.0077)	0.0093 (0.0084)
Not in labor force	0.0548*** (0.0152)	0.0441*** (0.0066)	0.0559*** (0.0066)	0.0561*** (0.0118)	0.0262*** (0.0043)	0.0323*** (0.0048)
Family size	0.0005 (0.0038)	-0.0065*** (0.0012)	-0.0092*** (0.0013)	-0.0090** (0.0044)	-0.0086*** (0.0015)	-0.0121*** (0.0016)
Dep. mean	0.386	0.193	0.225	0.443	0.153	0.182
N	9910	62193	62193	10117	67372	67372

Notes: Heterogeneous tax effects are estimated for men and women separately. All sample include individuals aged 45-59. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 1. Estimates of the impact of cigarette excise taxes on smoking participation, by personal income

	Personal income: 0-25k			Personal income: 25k+		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0559**	-0.0164	-0.0114	0.0020	0.0008	0.0001
	(0.0249)	(0.0107)	(0.0117)	(0.0301)	(0.0097)	(0.0098)
	{0.48}	{-0.29}	{-0.17}	{0.02}	{0.03}	{0.00}
Monthly state unemployment rate	0.0021	-0.0027	-0.0048	0.0046	0.0054	0.0088**
	(0.0116)	(0.0046)	(0.0049)	(0.0145)	(0.0043)	(0.0042)
Female	0.0624***	-0.0672***	-0.0749***	0.0425***	-0.0368***	-0.0414***
	(0.0101)	(0.0044)	(0.0056)	(0.0094)	(0.0026)	(0.0029)
Non-Hispanic white	-0.1221***	0.1578***	0.1540***	-0.0944***	0.0443***	0.0326***
	(0.0192)	(0.0063)	(0.0072)	(0.0264)	(0.0056)	(0.0078)
Non- Hispanic black	-0.0202	0.1297***	0.1390***	0.0030	0.0192	0.0122
	(0.0196)	(0.0191)	(0.0179)	(0.0340)	(0.0119)	(0.0113)
Other Races	-0.0375*	0.1680***	0.1700***	-0.0508	0.0537***	0.0303**
	(0.0223)	(0.0182)	(0.0196)	(0.0349)	(0.0115)	(0.0129)
High school	-0.0018	-0.0414***	-0.0439***	0.0284	-0.0459***	-0.0501***
	(0.0123)	(0.0060)	(0.0068)	(0.0246)	(0.0050)	(0.0056)
Some college	0.0457***	-0.0621***	-0.0623***	0.0746***	-0.0604***	-0.0653***
	(0.0139)	(0.0074)	(0.0089)	(0.0251)	(0.0045)	(0.0053)
College or higher	0.0533***	-0.1402***	-0.1513***	0.1126***	-0.1576***	-0.1664***
	(0.0131)	(0.0070)	(0.0084)	(0.0258)	(0.0070)	(0.0075)
Family income 15k-25k	0.0086	-0.0216***	-0.0329***			
	(0.0153)	(0.0039)	(0.0045)			
Family income 25k-35k	-0.0016	-0.0263***	-0.0379***			
	(0.0175)	(0.0052)	(0.0058)			
Family income 35k-50k	0.0112	-0.0444***	-0.0582***	-0.0440*	-0.0128**	-0.0164**
	(0.0166)	(0.0048)	(0.0057)	(0.0245)	(0.0060)	(0.0072)
Family income 50k-75k	0.0348**	-0.0778***	-0.0922***	-0.0155	-0.0067	-0.0107*
	(0.0174)	(0.0050)	(0.0053)	(0.0199)	(0.0050)	(0.0056)

Family income 75k+	0.0021 (0.0309)	-0.0895*** (0.0076)	-0.1011*** (0.0084)	-0.0154 (0.0248)	-0.0400*** (0.0056)	-0.0446*** (0.0059)
Divorced, widowed, or separated	-0.0071 (0.0107)	0.0980*** (0.0045)	0.1118*** (0.0050)	0.0190 (0.0168)	0.0639*** (0.0045)	0.0825*** (0.0044)
Never married	-0.0465*** (0.0159)	0.0122* (0.0065)	0.0167** (0.0070)	-0.0206 (0.0191)	0.0222*** (0.0049)	0.0248*** (0.0054)
Unemployed	0.0573*** (0.0176)	0.0660*** (0.0090)	0.0901*** (0.0100)	0.0501* (0.0304)	0.0477*** (0.0089)	0.0735*** (0.0111)
Retired	0.0253 (0.0183)	0.0037 (0.0062)	0.0068 (0.0071)	0.0050 (0.0201)	0.0062 (0.0052)	0.0070 (0.0058)
Not in labor force	0.0571*** (0.0108)	0.0366*** (0.0049)	0.0434*** (0.0052)	0.0696*** (0.0216)	0.0186*** (0.0054)	0.0247*** (0.0061)
Family size	-0.0057 (0.0037)	-0.0071*** (0.0014)	-0.0099*** (0.0014)	0.0061 (0.0086)	-0.0095*** (0.0019)	-0.0128*** (0.0022)
Dep. mean	0.412	0.216	0.250	0.425	0.122	0.148
N	13401	68875	68875	8061	69256	69256

Notes: Heterogeneous tax effects are estimated across samples with different income levels. All samples include individuals aged 45-64. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 2. Estimates of the impact of cigarette excise taxes on smoking participation, by education

	High school or less			Some college or higher		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0396 (0.0302) {0.36}	-0.0354*** (0.0106) {-0.58}	-0.0327*** (0.0112) {-0.46}	0.0317 (0.0216) {0.26}	0.0057 (0.0076) {0.20}	0.0063 (0.0074) {0.18}
Monthly state unemployment rate	0.0079 (0.0115)	0.0009 (0.0049)	0.0021 (0.0049)	0.0027 (0.0104)	0.0004 (0.0033)	0.0013 (0.0034)
Female	0.0661*** (0.0095)	-0.0660*** (0.0049)	-0.0740*** (0.0058)	0.0472*** (0.0089)	-0.0388*** (0.0019)	-0.0428*** (0.0023)
Non-Hispanic white	-0.1052*** (0.0164)	0.1717*** (0.0081)	0.1660*** (0.0089)	-0.1062*** (0.0355)	0.0491*** (0.0035)	0.0399*** (0.0058)
Non- Hispanic black	-0.0080 (0.0175)	0.1227*** (0.0204)	0.1312*** (0.0190)	-0.0062 (0.0360)	0.0425*** (0.0090)	0.0358*** (0.0089)
Other Races	-0.0762*** (0.0266)	0.1649*** (0.0200)	0.1538*** (0.0195)	0.0133 (0.0376)	0.0528*** (0.0079)	0.0433*** (0.0130)
High school	0.0049 (0.0111)	-0.0620*** (0.0054)	-0.0655*** (0.0063)			
College or higher				0.0237** (0.0100)	-0.0714*** (0.0029)	-0.0804*** (0.0029)
Family income 15k-25k	0.0115 (0.0191)	-0.0228*** (0.0062)	-0.0324*** (0.0069)	-0.0022 (0.0208)	-0.0133** (0.0053)	-0.0238*** (0.0057)
Family income 25k-35k	-0.0135 (0.0184)	-0.0325*** (0.0066)	-0.0423*** (0.0071)	0.0543** (0.0236)	-0.0213*** (0.0037)	-0.0299*** (0.0045)
Family income 35k-50k	0.0088 (0.0150)	-0.0483*** (0.0058)	-0.0609*** (0.0066)	0.0122 (0.0253)	-0.0324*** (0.0041)	-0.0429*** (0.0046)
Family income 50k-75k	0.0125 (0.0187)	-0.0674*** (0.0063)	-0.0795*** (0.0063)	0.0484** (0.0218)	-0.0489*** (0.0036)	-0.0615*** (0.0039)
Family income 75k+	-0.0027 (0.0213)	-0.0942*** (0.0074)	-0.1062*** (0.0073)	0.0548** (0.0213)	-0.0871*** (0.0045)	-0.1016*** (0.0044)

Divorced, widowed, or separated	-0.0060 (0.0095)	0.1093*** (0.0045)	0.1222*** (0.0046)	-0.0011 (0.0116)	0.0624*** (0.0035)	0.0801*** (0.0038)
Never married	-0.0366** (0.0164)	0.0156*** (0.0051)	0.0176*** (0.0057)	-0.0363* (0.0203)	0.0188*** (0.0047)	0.0251*** (0.0056)
Unemployed	0.0293 (0.0202)	0.0837*** (0.0101)	0.1013*** (0.0113)	0.0894*** (0.0242)	0.0395*** (0.0063)	0.0659*** (0.0076)
Retired	0.0264** (0.0133)	0.0038 (0.0072)	0.0052 (0.0077)	0.0196 (0.0208)	0.0082** (0.0039)	0.0096** (0.0045)
Not in labor force	0.0562*** (0.0117)	0.0290*** (0.0055)	0.0361*** (0.0056)	0.0532*** (0.0119)	0.0394*** (0.0041)	0.0476*** (0.0044)
Family size	-0.0047 (0.0039)	-0.0042*** (0.0015)	-0.0069*** (0.0017)	-0.0045 (0.0049)	-0.0061*** (0.0012)	-0.0088*** (0.0013)
Dep. mean	0.387	0.231	0.264	0.445	0.119	0.146
N	13308	67360	67360	10516	91426	91426

Notes: Heterogeneous tax effects are estimated across samples with different education levels. All samples include individuals aged 45-64. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 3. Estimates of the impact of cigarette excise taxes on smoking participation, by gender

	Men			Women		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0512*	-0.0108	-0.0128	0.0206	-0.0095	-0.0052
	(0.0299)	(0.0079)	(0.0082)	(0.0249)	(0.0088)	(0.0092)
	{0.48}	{-0.24}	{-0.23}	{0.17}	{-0.26}	{-0.12}
Monthly state unemployment rate	0.0126	0.0035	0.0070	-0.0025	-0.0019	-0.0033
	(0.0127)	(0.0040)	(0.0043)	(0.0089)	(0.0040)	(0.0036)
Non-Hispanic white	-0.1130***	0.0997***	0.0876***	-0.0925***	0.1120***	0.1160***
	(0.0220)	(0.0054)	(0.0053)	(0.0293)	(0.0055)	(0.0066)
Non- Hispanic black	-0.0241	0.0877***	0.0878***	0.0214	0.0962***	0.1018***
	(0.0187)	(0.0150)	(0.0119)	(0.0303)	(0.0187)	(0.0185)
Other Races	-0.0500**	0.1388***	0.1214***	-0.0157	0.1097***	0.1102***
	(0.0222)	(0.0151)	(0.0118)	(0.0333)	(0.0207)	(0.0238)
High school	-0.0015	-0.0401***	-0.0437***	0.0060	-0.0311***	-0.0334***
	(0.0145)	(0.0045)	(0.0056)	(0.0158)	(0.0066)	(0.0078)
Some college	0.0556***	-0.0589***	-0.0611***	0.0460***	-0.0515***	-0.0536***
	(0.0151)	(0.0054)	(0.0066)	(0.0157)	(0.0060)	(0.0074)
College or higher	0.0853***	-0.1556***	-0.1686***	0.0711***	-0.1225***	-0.1300***
	(0.0148)	(0.0052)	(0.0062)	(0.0172)	(0.0065)	(0.0080)
Family income 15k-25k	-0.0035	-0.0097	-0.0187***	0.0172	-0.0197***	-0.0295***
	(0.0206)	(0.0064)	(0.0068)	(0.0202)	(0.0038)	(0.0044)
Family income 25k-35k	0.0323*	-0.0197***	-0.0291***	-0.0074	-0.0246***	-0.0322***
	(0.0194)	(0.0062)	(0.0067)	(0.0180)	(0.0042)	(0.0051)
Family income 35k-50k	0.0175	-0.0357***	-0.0454***	-0.0012	-0.0331***	-0.0442***
	(0.0181)	(0.0061)	(0.0065)	(0.0145)	(0.0036)	(0.0041)
Family income 50k-75k	0.0331*	-0.0517***	-0.0630***	0.0175	-0.0519***	-0.0637***
	(0.0187)	(0.0062)	(0.0066)	(0.0171)	(0.0041)	(0.0038)
Family income 75k+	0.0324	-0.0967***	-0.1082***	0.0140	-0.0819***	-0.0956***
	(0.0216)	(0.0054)	(0.0053)	(0.0188)	(0.0046)	(0.0042)

Divorced, widowed, or separated	-0.0062 (0.0111)	0.1020*** (0.0049)	0.1200*** (0.0049)	-0.0019 (0.0101)	0.0728*** (0.0039)	0.0885*** (0.0039)
Never married	-0.0363** (0.0153)	0.0194*** (0.0049)	0.0268*** (0.0061)	-0.0361* (0.0192)	0.0219*** (0.0052)	0.0247*** (0.0052)
Unemployed	0.0643*** (0.0186)	0.0694*** (0.0097)	0.0888*** (0.0101)	0.0405* (0.0231)	0.0509*** (0.0085)	0.0791*** (0.0096)
Retired	0.0209 (0.0190)	0.0036 (0.0051)	0.0037 (0.0054)	0.0248 (0.0163)	0.0101* (0.0052)	0.0129** (0.0053)
Not in labor force	0.0550*** (0.0128)	0.0446*** (0.0054)	0.0552*** (0.0056)	0.0583*** (0.0120)	0.0251*** (0.0039)	0.0315*** (0.0042)
Family size	-0.0002 (0.0040)	-0.0050*** (0.0010)	-0.0071*** (0.0011)	-0.0105** (0.0042)	-0.0059*** (0.0013)	-0.0095*** (0.0014)
Dep. mean	0.383	0.186	0.216	0.440	0.149	0.177
N	11701	76029	76029	12123	82757	82757

Notes: Heterogeneous tax effects are estimated for men and women separately. All samples include individuals aged 45-64. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 4. Estimates of the impact of cigarette excise taxes on smoking participation.

	Age: 45-59			Age: 45-64		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0443**	-0.0112	-0.0095	0.0337*	-0.0091	-0.0080
	(0.0205)	(0.0076)	(0.0077)	(0.0196)	(0.0073)	(0.0075)
	{0.38}	{-0.27}	{-0.19}	{0.29}	{-0.22}	{-0.16}
State anti-smoking sentiment	0.1707	-0.0456	-0.0355	0.1283	-0.0437	-0.0334
	(0.1100)	(0.0353)	(0.0334)	(0.0880)	(0.0327)	(0.0318)
Monthly state unemployment rate	0.0088	0.0004	0.0011	0.0053	0.0006	0.0016
	(0.0082)	(0.0037)	(0.0034)	(0.0061)	(0.0032)	(0.0032)
Female	0.0568***	-0.0515***	-0.0572***	0.0580***	-0.0504***	-0.0562***
	(0.0074)	(0.0024)	(0.0031)	(0.0061)	(0.0025)	(0.0032)
White	-0.1138***	0.1103***	0.1052***	-0.1056***	0.1063***	0.1027***
	(0.0232)	(0.0053)	(0.0061)	(0.0199)	(0.0047)	(0.0054)
Black	-0.0092	0.0940***	0.0945***	-0.0034	0.0898***	0.0920***
	(0.0210)	(0.0156)	(0.0147)	(0.0171)	(0.0152)	(0.0138)
Other Race	-0.0375	0.1296***	0.1184***	-0.0351*	0.1225***	0.1140***
	(0.0242)	(0.0146)	(0.0159)	(0.0201)	(0.0134)	(0.0147)
High school	-0.0075	-0.0369***	-0.0410***	0.0016	-0.0352***	-0.0383***
	(0.0113)	(0.0050)	(0.0064)	(0.0115)	(0.0045)	(0.0058)
Some college	0.0487***	-0.0585***	-0.0615***	0.0494***	-0.0546***	-0.0569***
	(0.0116)	(0.0052)	(0.0067)	(0.0102)	(0.0051)	(0.0063)
College or Higher	0.0702***	-0.1476***	-0.1601***	0.0763***	-0.1378***	-0.1481***
	(0.0131)	(0.0053)	(0.0069)	(0.0111)	(0.0050)	(0.0063)
Family income 15-25k	0.0109	-0.0166***	-0.0251***	0.0064	-0.0153***	-0.0247***
	(0.0154)	(0.0038)	(0.0043)	(0.0146)	(0.0032)	(0.0037)
Family income 25-35k	0.0122	-0.0218***	-0.0288***	0.0097	-0.0225***	-0.0307***
	(0.0150)	(0.0043)	(0.0051)	(0.0141)	(0.0035)	(0.0043)
Family income 35-50k	0.0054	-0.0352***	-0.0449***	0.0056	-0.0345***	-0.0450***
	(0.0138)	(0.0045)	(0.0051)	(0.0131)	(0.0034)	(0.0040)

Family income 50-75k	0.0250* (0.0133)	-0.0530*** (0.0045)	-0.0637*** (0.0046)	0.0231* (0.0131)	-0.0523*** (0.0036)	-0.0640*** (0.0037)
Family income 75k+	0.0208 (0.0151)	-0.0934*** (0.0045)	-0.1053*** (0.0043)	0.0209 (0.0151)	-0.0898*** (0.0038)	-0.1028*** (0.0034)
Unmarried	-0.0080 (0.0079)	0.0860*** (0.0038)	0.1026*** (0.0041)	-0.0045 (0.0076)	0.0842*** (0.0032)	0.1006*** (0.0032)
Never married	-0.0328** (0.0128)	0.0176*** (0.0038)	0.0226*** (0.0045)	-0.0378*** (0.0117)	0.0190*** (0.0034)	0.0240*** (0.0039)
Unemployed	0.0483*** (0.0166)	0.0675*** (0.0072)	0.0936*** (0.0084)	0.0516*** (0.0149)	0.0611*** (0.0063)	0.0847*** (0.0071)
Retired	0.0112 (0.0205)	0.0029 (0.0061)	0.0048 (0.0064)	0.0214* (0.0116)	0.0076* (0.0039)	0.0091** (0.0040)
Not in labor force	0.0530*** (0.0084)	0.0329*** (0.0041)	0.0405*** (0.0044)	0.0544*** (0.0083)	0.0326*** (0.0036)	0.0401*** (0.0039)
Family size	-0.0032 (0.0029)	-0.0077*** (0.0011)	-0.0109*** (0.0011)	-0.0045 (0.0028)	-0.0057*** (0.0010)	-0.0085*** (0.0011)
Dep. Mean	0.415	0.173	0.203	0.412	0.167	0.196
N	20027	129565	129565	23824	158786	158786

Notes: Tax effects are estimated for different age groups. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. In addition to main regressors as specified from Table 4 to Table7, we add monthly state anti-smoking sentiment into analysis. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 5. Estimates of the impact of cigarette excise taxes on smoking participation, by personal income

	Personal income: 0-25k			Personal income: 25k+		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0702**	-0.0246**	-0.0175	0.0149	0.0022	0.0022
	(0.0309)	(0.0111)	(0.0120)	(0.0271)	(0.0112)	(0.0118)
	{0.59}	{-0.42}	{-0.25}	{0.13}	{0.08}	{0.06}
State anti-smoking sentiment	0.2474*	-0.0516	-0.0417	0.0834	-0.0080	-0.0010
	(0.1433)	(0.0488)	(0.0551)	(0.1155)	(0.0469)	(0.0470)
Monthly state unemployment rate	0.0144	-0.0018	-0.0038	-0.0011	0.0055	0.0082*
	(0.0143)	(0.0053)	(0.0056)	(0.0136)	(0.0044)	(0.0046)
Female	0.0617***	-0.0693***	-0.0762***	0.0432***	-0.0372***	-0.0420***
	(0.0114)	(0.0043)	(0.0055)	(0.0097)	(0.0026)	(0.0028)
Non-Hispanic white	-0.1349***	0.1629***	0.1573***	-0.0765***	0.0455***	0.0324***
	(0.0230)	(0.0078)	(0.0088)	(0.0288)	(0.0045)	(0.0068)
Non- Hispanic black	-0.0285	0.1328***	0.1417***	0.0208	0.0220*	0.0134
	(0.0235)	(0.0202)	(0.0194)	(0.0377)	(0.0116)	(0.0114)
Other Races	-0.0479*	0.1711***	0.1705***	-0.0181	0.0602***	0.0356***
	(0.0245)	(0.0200)	(0.0210)	(0.0372)	(0.0105)	(0.0126)
High school	-0.0137	-0.0431***	-0.0462***	0.0408	-0.0459***	-0.0518***
	(0.0127)	(0.0065)	(0.0070)	(0.0293)	(0.0063)	(0.0071)
Some college	0.0471***	-0.0677***	-0.0683***	0.0882***	-0.0621***	-0.0684***
	(0.0153)	(0.0075)	(0.0088)	(0.0306)	(0.0056)	(0.0066)
College or higher	0.0470***	-0.1508***	-0.1632***	0.1259***	-0.1647***	-0.1776***
	(0.0145)	(0.0077)	(0.0089)	(0.0303)	(0.0090)	(0.0096)
Family income 15k-25k	0.0132	-0.0231***	-0.0330***			
	(0.0155)	(0.0045)	(0.0050)			
Family income 25k-35k	0.0025	-0.0267***	-0.0354***			
	(0.0186)	(0.0061)	(0.0068)			
Family income 35k-50k	0.0087	-0.0450***	-0.0573***	-0.0356	-0.0180***	-0.0200***
	(0.0162)	(0.0059)	(0.0068)	(0.0245)	(0.0066)	(0.0076)

Family income 50k-75k	0.0409**	-0.0801***	-0.0926***	-0.0148	-0.0102*	-0.0127**
	(0.0175)	(0.0057)	(0.0059)	(0.0220)	(0.0053)	(0.0062)
Family income 75k+	-0.0075	-0.0872***	-0.0977***	-0.0139	-0.0478***	-0.0501***
	(0.0324)	(0.0089)	(0.0100)	(0.0274)	(0.0061)	(0.0068)
Divorced, widowed, or separated	-0.0097	0.0980***	0.1127***	0.0184	0.0651***	0.0834***
	(0.0105)	(0.0058)	(0.0063)	(0.0176)	(0.0056)	(0.0056)
Never married	-0.0372**	0.0099	0.0155*	-0.0188	0.0201***	0.0224***
	(0.0168)	(0.0077)	(0.0084)	(0.0203)	(0.0058)	(0.0064)
Unemployed	0.0553***	0.0696***	0.0968***	0.0475	0.0580***	0.0857***
	(0.0200)	(0.0103)	(0.0113)	(0.0321)	(0.0100)	(0.0132)
Retired	0.0220	-0.0077	-0.0018	-0.0270	0.0036	0.0021
	(0.0319)	(0.0079)	(0.0086)	(0.0293)	(0.0075)	(0.0090)
Not in labor force	0.0546***	0.0382***	0.0453***	0.0663***	0.0162***	0.0222***
	(0.0111)	(0.0050)	(0.0055)	(0.0246)	(0.0059)	(0.0066)
Family size	-0.0035	-0.0100***	-0.0131***	0.0084	-0.0109***	-0.0148***
	(0.0038)	(0.0016)	(0.0015)	(0.0098)	(0.0021)	(0.0025)
Dep. mean	0.416	0.224	0.259	0.425	0.127	0.153
N	11132	55616	55616	6961	57639	57639

Notes: Heterogeneous tax effects are estimated across samples with different income levels. All samples include individuals aged 45-59. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. In addition to main regressors as specified from Table 4 to Table7, we add monthly state anti-smoking sentiment into analysis. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 6. Estimates of the impact of cigarette excise taxes on smoking participation, by education

	High school or less			Some college or higher		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0600*	-0.0324***	-0.0296***	0.0271	0.0014	0.0024
	(0.0322)	(0.0105)	(0.0111)	(0.0248)	(0.0084)	(0.0084)
	{0.55}	{-0.50}	{-0.39}	{0.22}	{0.05}	{0.06}
State anti-smoking sentiment	0.2404	-0.1465***	-0.1471***	0.0860	0.0193	0.0354
	(0.1824)	(0.0567)	(0.0562)	(0.1134)	(0.0337)	(0.0328)
Monthly state unemployment rate	0.0205*	0.0015	0.0028	-0.0039	-0.0006	-0.0003
	(0.0118)	(0.0059)	(0.0054)	(0.0116)	(0.0036)	(0.0035)
Female	0.0662***	-0.0692***	-0.0766***	0.0447***	-0.0393***	-0.0434***
	(0.0102)	(0.0053)	(0.0061)	(0.0102)	(0.0020)	(0.0024)
Non-Hispanic white	-0.1224***	0.1818***	0.1730***	-0.0982**	0.0507***	0.0406***
	(0.0205)	(0.0096)	(0.0102)	(0.0396)	(0.0041)	(0.0066)
Non- Hispanic black	-0.0172	0.1310***	0.1375***	-0.0011	0.0420***	0.0340***
	(0.0206)	(0.0221)	(0.0200)	(0.0421)	(0.0094)	(0.0104)
Other Races	-0.0858***	0.1788***	0.1620***	0.0226	0.0555***	0.0458***
	(0.0299)	(0.0225)	(0.0216)	(0.0412)	(0.0089)	(0.0143)
High school	-0.0047	-0.0657***	-0.0706***			
	(0.0111)	(0.0059)	(0.0070)			
College or higher				0.0190*	-0.0760***	-0.0865***
				(0.0111)	(0.0032)	(0.0034)
Family income 15k-25k	0.0142	-0.0249***	-0.0326***	0.0079	-0.0133**	-0.0234***
	(0.0208)	(0.0070)	(0.0078)	(0.0189)	(0.0059)	(0.0067)
Family income 25k-35k	-0.0083	-0.0333***	-0.0408***	0.0517**	-0.0191***	-0.0271***
	(0.0195)	(0.0074)	(0.0079)	(0.0241)	(0.0049)	(0.0055)
Family income 35k-50k	0.0144	-0.0497***	-0.0610***	0.0038	-0.0319***	-0.0418***
	(0.0156)	(0.0073)	(0.0081)	(0.0251)	(0.0051)	(0.0059)
Family income 50k-75k	0.0140	-0.0666***	-0.0760***	0.0473**	-0.0497***	-0.0623***
	(0.0211)	(0.0077)	(0.0075)	(0.0216)	(0.0043)	(0.0048)

Family income 75k+	0.0003 (0.0229)	-0.0997*** (0.0085)	-0.1093*** (0.0084)	0.0483** (0.0221)	-0.0897*** (0.0055)	-0.1040*** (0.0058)
Divorced, widowed, or separated	-0.0105 (0.0105)	0.1147*** (0.0056)	0.1274*** (0.0059)	-0.0031 (0.0110)	0.0623*** (0.0039)	0.0801*** (0.0045)
Never married	-0.0332* (0.0177)	0.0154** (0.0063)	0.0188*** (0.0070)	-0.0311 (0.0209)	0.0171*** (0.0046)	0.0225*** (0.0058)
Unemployed	0.0333 (0.0209)	0.0915*** (0.0122)	0.1122*** (0.0138)	0.0743*** (0.0251)	0.0459*** (0.0067)	0.0741*** (0.0087)
Retired	0.0242 (0.0276)	0.0029 (0.0105)	0.0020 (0.0106)	-0.0042 (0.0280)	0.0008 (0.0059)	0.0041 (0.0071)
Not in labor force	0.0525*** (0.0113)	0.0304*** (0.0064)	0.0372*** (0.0064)	0.0568*** (0.0131)	0.0389*** (0.0045)	0.0472*** (0.0051)
Family size	-0.0049 (0.0039)	-0.0068*** (0.0016)	-0.0100*** (0.0019)	-0.0016 (0.0049)	-0.0078*** (0.0013)	-0.0107*** (0.0014)
Dep. mean	0.388	0.245	0.279	0.447	0.123	0.149
N	11007	52965	52965	9020	76600	76600

Notes: Heterogeneous tax effects are estimated across samples with different education levels. All samples include individuals aged 45-59. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. In addition to main regressors as specified from Table 4 to Table7, we add monthly state anti-smoking sentiment into analysis. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Appendix 7. Estimates of the impact of cigarette excise taxes on smoking participation, by gender

	Men			Women		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
Cigarette tax	0.0544*	-0.0113	-0.0120	0.0334	-0.0112	-0.0072
	(0.0328)	(0.0087)	(0.0092)	(0.0263)	(0.0087)	(0.0087)
	{0.51}	{-0.24}	{-0.21}	{0.27}	{-0.30}	{-0.16}
State anti-smoking sentiment	0.2184	-0.0384	-0.0284	0.1142	-0.0529	-0.0434
	(0.1351)	(0.0449)	(0.0364)	(0.1578)	(0.0440)	(0.0409)
Monthly state unemployment rate	0.0242*	0.0044	0.0080*	-0.0059	-0.0035	-0.0053
	(0.0130)	(0.0044)	(0.0044)	(0.0122)	(0.0045)	(0.0041)
Non-Hispanic white	-0.1075***	0.1046***	0.0904***	-0.1178***	0.1154***	0.1185***
	(0.0271)	(0.0058)	(0.0064)	(0.0310)	(0.0062)	(0.0075)
Non- Hispanic black	-0.0228	0.0947***	0.0928***	0.0039	0.0986***	0.1026***
	(0.0250)	(0.0155)	(0.0134)	(0.0323)	(0.0203)	(0.0203)
Other Races	-0.0367	0.1494***	0.1286***	-0.0375	0.1138***	0.1126***
	(0.0284)	(0.0155)	(0.0126)	(0.0356)	(0.0210)	(0.0250)
High school	-0.0100	-0.0418***	-0.0468***	-0.0031	-0.0324***	-0.0353***
	(0.0154)	(0.0051)	(0.0059)	(0.0154)	(0.0070)	(0.0088)
Some college	0.0534***	-0.0643***	-0.0675***	0.0461**	-0.0539***	-0.0567***
	(0.0157)	(0.0057)	(0.0068)	(0.0184)	(0.0062)	(0.0081)
College or higher	0.0811***	-0.1663***	-0.1814***	0.0627***	-0.1314***	-0.1413***
	(0.0175)	(0.0060)	(0.0070)	(0.0187)	(0.0069)	(0.0087)
Family income 15k-25k	-0.0032	-0.0096	-0.0157*	0.0251	-0.0217***	-0.0320***
	(0.0211)	(0.0081)	(0.0086)	(0.0213)	(0.0042)	(0.0045)
Family income 25k-35k	0.0301	-0.0186**	-0.0253***	0.0005	-0.0242***	-0.0315***
	(0.0218)	(0.0077)	(0.0081)	(0.0189)	(0.0051)	(0.0064)
Family income 35k-50k	0.0127	-0.0346***	-0.0413***	0.0046	-0.0349***	-0.0467***
	(0.0208)	(0.0070)	(0.0072)	(0.0157)	(0.0047)	(0.0055)
Family income 50k-75k	0.0316*	-0.0517***	-0.0608***	0.0238	-0.0528***	-0.0643***
	(0.0192)	(0.0073)	(0.0080)	(0.0176)	(0.0050)	(0.0048)

Family income 75k+	0.0298 (0.0219)	-0.0999*** (0.0064)	-0.1087*** (0.0063)	0.0177 (0.0200)	-0.0854*** (0.0053)	-0.0992*** (0.0050)
Divorced, widowed, or separated	-0.0119 (0.0115)	0.1043*** (0.0059)	0.1214*** (0.0061)	-0.0031 (0.0108)	0.0742*** (0.0047)	0.0908*** (0.0052)
Never married	-0.0327** (0.0160)	0.0177*** (0.0052)	0.0244*** (0.0064)	-0.0294 (0.0204)	0.0212*** (0.0055)	0.0247*** (0.0056)
Unemployed	0.0643*** (0.0196)	0.0775*** (0.0112)	0.0998*** (0.0118)	0.0319 (0.0248)	0.0554*** (0.0093)	0.0859*** (0.0103)
Retired	0.0117 (0.0306)	-0.0014 (0.0083)	-0.0016 (0.0087)	0.0129 (0.0242)	0.0057 (0.0077)	0.0092 (0.0084)
Not in labor force	0.0548*** (0.0153)	0.0441*** (0.0066)	0.0559*** (0.0066)	0.0561*** (0.0118)	0.0262*** (0.0043)	0.0323*** (0.0048)
Family size	0.0006 (0.0038)	-0.0065*** (0.0012)	-0.0092*** (0.0013)	-0.0090** (0.0044)	-0.0086*** (0.0015)	-0.0121*** (0.0016)
Dep. mean	0.386	0.193	0.225	0.443	0.153	0.182
N	9910	62193	62193	10117	67372	67372

Notes: Heterogeneous tax effects are estimated for men and women separately. All samples include individuals aged 45-59. Observations with missing income are included by adding a separate indicator variable. Dependent variables in models (1), (2), and (3) are binary variables indicating intention to quit, smoker on all days, and smoker on some days, respectively. In addition to main regressors as specified from Table 4 to Table7, we add monthly state anti-smoking sentiment into analysis. All models are estimated by probit regression; marginal effects are reported in the first row, robust standard errors clustered on state of residence are in parentheses in the second row and price participation elasticities are in curly braces in the third row. The marginal effect associated with the cigarette tax gives the estimated impact of increasing it by $1.00 per pack. *** implies p<0.01, ** implies p<0.05 and * implies p<0.10.

Chapter 4

A Smooth Nonparametric Conditional Density Test for Categorical Responses

4.1 Introduction

Model specification tests can be found in every practitioner's toolkit. In the discrete outcome context, existing tests include Hausman & McFadden's (1984) test of correct specification of the multinomial logit model, Mora & Moro-Egido's (2008) test of correct specification of ordered discrete models, or Rodriguez-Poo, Sperlich & Vieu's (2004) test for correct specification of semiparametric models.

However, existing tests are applicable to a particular subset of categorical dependent variable models or are not consistent tests in that there exist alternative models for which the tests lack power or are not applicable to models containing both numeric and categorical covariates. An omnibus procedure capable of handling a rich array of categorical dependent variable models that is consistent and can admit both numeric and categorical data would have obvious appeal.

Fan, Li & Min (2006) have recently proposed a non-smoothing test for correct specification of parametric conditional distributions based on previous work by Zheng (2000) where non-smoothing here refers to non-smoothing of the response (they do smooth the categorical covariates). In the current paper we extend the approach of Fan et al. (2006) to obtain a smooth version of their non-smoothing test that displays finite-sample power improvements over its non-smoothing counterpart (i.e. we smooth the categorical/discrete response in addition to the categorical/discrete covariates).

Zheng (2000) proposed a kernel based nonparametric test for a parametric density function. Suppose that the data consist of $\{y_i, x_i\}_{i=1}^{n}$, an i.i.d. sample drawn from the distribution of (y, x) with the joint density function $p(y, x)$. Let $p(y|x)$ denote the conditional density function of y given x. We are interested in testing whether $p(y|x)$ belongs to a particular parametric family. Let $f(y|x, \theta)$ denote a parametric conditional density function with θ being an unknown $k \times 1$-dimensional parameter vector. The null hypothesis is given by

$$H_0: \Pr[p(y_i|x_i) = f(y_i|x_i, \theta_0)] = 1 \quad \text{for some } \theta_0 \in \Theta$$

Where Θ is a compact set in $\mathbb{R}^k$. The alternative hypothesis is the negation of the null:

$$H_1: \Pr[p(y_i|x_i) = f(y_i|x_i, \theta)] < 1 \quad \text{for some } \theta \in \Theta$$

Zheng (2000) proposed using the following quantity as a basis for a consistent test for H_0, which is simply a linearized version of the Kullback-Leibler information function $I(p, f)$ between two conditional distribution functions $p(y|x)$ and $f(y|x, \theta_0)$ and is given by

$$I_1(p, f) = E\left[\frac{p(y_i, x_i) - f(y_i|x_i, \theta_0)p_1(x_i)}{f(y_i|x_i, \theta_0)}\right] \qquad (1)$$

Zheng (2000) has shown that $I_1(p, f) \geq 0$ and the equality holds if and only if H_0 is true (i.e. $p(y|x)$ equals $f(y|x, \theta_0)$ almost everywhere). Therefore, $I_1(p, f)$ also serves as a proper measure to test for H_0.

Zheng (2000) only considered the case for which all variables (i.e. covariates and outcome) are continuous. Fan et al. (2006) extended Zheng's (2000) test to include both continuous and discrete explanatory variables (x is a mixed variable) while they also allow the dependent variable y to be either continuous or discrete.

Let $x = (x^c, x^d)$, where x^c is a $q \times 1$ continuous variable, and x^d is an $r \times 1$ discrete variable. Fan et al. (2006) use x_{is}^c (x_{is}^d) to denote the sth component of x_i^c (x_i^d). They further assume that x_{is}^d takes the values $0, 1, \ldots, c_s - 1$ (i.e. it can assume c_s different values).

For each discrete component of x_d, Fan et al. (2006) use the Aitchison & Aitken (1976) kernel given by $l(x_{is}^d, x_{js}^d, \lambda_s) = 1 - \lambda_s$ if $x_{is}^d = x_{js}^d$, and

$l(x_{is}^d, x_{js}^d, \lambda_s) = \lambda_s/(c_s - 1)$ if $x_{is}^d \neq x_{js}^d$, and for the $r \times 1$ variable x_d they use the product kernel given by

$L(x_i^d, x_j^d, \lambda) = \prod_{s=1}^r l(x_{is}^d, x_{js}^d, \lambda_s) = \prod_{s=1}^r \{\lambda_s/(c_s - 1)\}^{N_{is(x)}} (1 - \lambda_s)^{1-N_{is(x)}}$,

where $N_{is(x)} = I(x_{is}^d \neq x_{js}^d)$, in which $I(.)$ Is the usual indicator function and $\lambda = (\lambda_1, \dots \lambda_r)$ are the smoothing parameters for the discrete components which satisfy $0 \leq \lambda_s \leq (c_s - 1)/c_s$.

For the $q \times 1$ continuous component x^c, a standard (second order) product kernel function can be used and is given by $W(x_i^c, x_j^c, h) = \prod_{s=1}^q W((x_{is}^c - x_{js}^c)/h_s)/h_s$ where $W((x_{is}^c - x_{js}^c)/h_s)/h_s$ is a standard univariate kernel such as the Epanechnikov. Therefore, for the mix of continuous and categorical covariates $x = (x^c, x^d)$, the kernel function is defined by $K_{\gamma,ij} = K_\gamma(x_i, x_j) \stackrel{\text{def}}{=} W(x_i^c, x_j^c, h) \times L(x_i^d, x_j^d, \lambda)$, where $\gamma = (h, \lambda) \equiv (h_1, \dots, h_q, \lambda_1, \dots, \lambda_r)$.

Fan et al. (2006) propose estimating $p(y_i, x_i)$ and $p_1(x_i)$ using the following method. Assuming that y_i is a discrete variable, then Fan et al. (2006) estimate $p(y_i, x_i)$ and $p_1(x_i)$ by the following leave-one-out kernel estimators:

$$\hat{p}_{-i}(y_i, x_i) = \frac{1}{n}\sum_{j \neq i}^{n} I(y_i = y_j) K_\gamma(x_i, x_j) \quad \text{and}$$

$$\hat{p}_{1,-i}(x_i) = \frac{1}{n}\sum_{j \neq i}^{n} K_\gamma(x_i, x_j) \qquad (2)$$

A smoothed leave-one-out estimator of $f(y_i|x_i, \theta_0)$ is given by

$$\tilde{p}_{-i}(y_i, x_i) = \frac{1}{n}\sum_{j\neq i}^{n}\sum_{y} I(y_i = y)K_\gamma(x_i, x_j)f(y|x_j, \hat{\theta})$$

$$= \frac{1}{n}\sum_{j\neq i}^{n} K_\gamma(x_i, x_j)f(y_i|x_j, \hat{\theta}). \qquad (3)$$

Replacing $p(y_i, x_i)$, $f(y_i|x_i, \theta_0)p_1(x_i)$ and $f(y_i|x_i, \theta_0)$ by $\hat{p}_{-i}(y_i, x_i)$, $\tilde{p}_{-i}(y_i, x_i)$ and $f(y_i|x_i, \hat{\theta})$ in (1) leads to the following feasible test statistic.

$$J_{n,\gamma}^{0} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\left\{\frac{K_\gamma(x_i, x_j)}{f(y_i|x_i, \hat{\theta})}[I(y_i = y_j) - f(y_i|x_j, \hat{\theta})]\right\} \qquad (4)$$

4.2 A Smooth Nonparametric Conditional Density Test

When x contains discrete components, Fan et al. (2006) suggest also smoothing the discrete components of x. It is well known that one can choose not to smooth the discrete variables in nonparametric estimation or testing problems by resorting to sample-splitting. The rationale for smoothing the discrete components of x is that this avoids splitting the sample into discrete subsets (`cells') and, by borrowing information from nearby cells, one can realize efficiency gains in finite-sample settings. However, when y is a discrete variable, Fan et al. (2006) rely on sample splitting methods to deal with the dependent variable y. As was the case for the discrete components of x, it is reasonable to conjecture that by smoothing the discrete y variable in an appropriate manner one may obtain further finite-sample (power) gains for testing a conditional parametric densities' functional form. We will observe via Monte Carlo simulation

that the improvement in power can be non-trivial.

In this paper we propose smoothing y by replacing the indicator function $I(y_i = y_j)$ and $I(y_i = y)$ in Fan et al.'s (2006) test by the smooth kernel functions $l(y_i, y_j, \lambda_0)$ and $l(y_i, y, \lambda_0)$, where λ_0 is the smoothing parameter, and

$$l(y_i, y_j, \lambda_0) = (1 - \lambda_0) I(y_i = y_j) + \frac{\lambda_0}{c_0 - 1} I(y_i \neq y_j) \tag{5}$$

if y_i is an unordered variable, and

$$l(y_i, y_j, \lambda_0) = (1 - \lambda_0) \left[I(y_i = y_j) + \frac{\lambda_0^{|y_i - y_j|}}{2} I(y_i \neq y_j) \right] \tag{6}$$

if y_i is an ordered variable.

Then we estimate $p(y_i, x_i)$ and $p_1(x_i)$ by the following leave-one-out kernel estimator:

$$\hat{p}_{-i}(y_i, x_i) = \frac{1}{n} \sum_{j \neq i}^{n} l(y_i, y_j, \lambda_0) K_\gamma(x_i, x_j), \quad \hat{p}_{1,-i}(x_i) = \frac{1}{n} \sum_{j \neq i}^{n} K_\gamma(x_i, x_j), \tag{7}$$

$$\tilde{p}_{-i}(y_i, x_i) = \frac{1}{n} \sum_{j \neq i}^{n} \sum_{y} l(y_i, y, \lambda_0) K_\gamma(x_i, x_j) f(y | x_j, \hat{\theta}). \tag{8}$$

Using $\hat{p}_{-i}(y_i, x_i)$, $\hat{p}_{1,-i}(x_i)$ and $\tilde{p}_{-i}(y_i, x_i)$ just introduced, we define our test statistic as

$$J_{n,\gamma} = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \left\{ \frac{K_\gamma(x_i, x_j)}{f(y_i | x_i, \hat{\theta})} [l(y_i, y_j, \lambda_0) - \sum_{y} l(y_i, y, \lambda_0) f(y | x_j, \hat{\theta})] \right\}. \tag{9}$$

We suggest using cross-validation methods for choosing the smoothing parameter, and denote by $\hat{\gamma} = (\hat{h}_1, \dots, \hat{h}_q, \hat{\lambda}_0, \hat{\lambda}_1, \dots, \hat{\lambda}_r)$ the cross-validated smoothing parameter vector. The asymptotic behavior of $\hat{\lambda}$ is summarized in the following lemma.

Lemma 4.2.1 *Under the regularity conditions given in Hall, Racine and Li (2004),*

$n^{1/(q+4)}\hat{h}_s \xrightarrow{p} a_s^0$ *for* $1 \le s \le q$; $n^{2/(q+4)}\hat{\lambda}_s \xrightarrow{p} b_s^0$ *for* $0 \le s \le r$,

where $a_s^0 (> 0)$ *and* $b_s^0 (\ge 0)$ *are some finite constants.*

Theorem 4.2.1 *Suppose that Y is an unordered discrete variable. Under the regularity conditions given in Fan et al. (2006) and defining*

$$T_n = n\sqrt{\hat{h}_1 \dots \hat{h}_q} J_{n,\hat{\gamma}} / \sqrt{\hat{V}_a} \text{ (} \hat{V}_a \text{ is defined in the appendix),}$$

Then we have,

(i) *Under* H_0, $T_n \xrightarrow{d} N(0,1)$;

(ii) *Under* H_1, $\Pr[T_n > B_n] \to 1$ *for any non-stochastic sequence* $B_n = o(n(h_1 \dots h_q)^{1/2})$.

Theorem 4.2.2 *Suppose that Y is an ordered discrete variable. Under the same conditions as given in Fan et al. (2006) and defining*

$$T_n = n\sqrt{\hat{h}_1 \dots \hat{h}_q} J_{n,\hat{\gamma}} / \sqrt{(\hat{V}_b)} \text{ (} \hat{V}_b \text{ is defined in the appendix,)}$$

we have,

(i) *Under* H_0, $T_n \xrightarrow{d} N(0,1)$;

(ii) *Under* H_1, $\Pr[T_n > B_n] \to 1$, *for any non-stochastic sequence* $B_n = o(n(h_1 \dots h_q)^{1/2})$.

4.2.1 Parametric bootstrap Test

It is well known that consistent nonparametric tests often suffer from

substantial finite-sample size distortions. To overcome this problem, Fan et al. (2006) propose a bootstrap procedure to accurately approximate the finite-sample null distribution of T_n. It involves the following steps.

Step (i): Generate the ith bootstrap value of the dependent variable y from the parametric conditional distribution $f(\cdot\,|x_i, \hat{\theta})$. Denote this value by y_i^* $(i = 1, \dots, n)$. We have the complete bootstrap sample $\{x_i, y_i^*\}_{i=1}^n$.

Step (ii): Based on the parametric null model, estimate θ using the bootstrap sample. Let $\hat{\theta}^*$ denote the resulting estimator. Computing the bootstrap statistic T_n^* in the same way as T_n except that $\{y_i\}_{i=1}^n$ and $\hat{\theta}$ are replaced by $\{y_i^*\}_{i=1}^n$ and $\hat{\theta}^*$, respectively. Note that Fan et al. (2006) use the same cross-validation selected smoothing parameter $\hat{\gamma}$ in computing the bootstrap statistics. There is no re-cross-validation in computing $J_{n,\hat{\gamma}}^*$.

Step (iii): Repeat steps (i) and (ii) a large number of times, say B times, and use the empirical distribution of the B bootstrap statistics $\{T_n^*\}_{j=1}^B$ to approximate the null distribution of T_n.

Step (iv): The bootstrap test rejects H_0 at significance level α if T_n exceeds the empirical α-th percentile of $\{T_n^*\}_{j=1}^B$.

4.3 Monte Carlo Results

We consider a smoothed version of Fan et al.'s (2006) conditional density test in an ordered outcome setting. We compare the smooth and frequency (non-smooth) versions of the test incorporating a modification of the binary outcome DGP found in Fan et al. (2006). Note that Fan et al. (2006) consider $\beta_2 \in \{0,1\}$, while we consider $\beta_2 \in \{0,-1\}$, while the alternative DGP they consider involves $\beta_3 x^2$. Frequently their DGP lies to the right of the cutoff point, i.e., $\beta_0 + \beta_1 x_i + \beta_2 z_i + \beta_3 x^2 + u_i > 0$ for almost all n leading to $y_i = 1$ for almost all $i = 1, \dots, n$. By using $\beta_2 = \{0,-1\}$ to offset the positive intercept ($\beta_0 = 1$) and by using an alternative DGP involving $\sin(0.5\pi x)$ (which oscillates around zero taking positive and negative values) we obtain a DGP that is more `balanced' which is desirable in the ordered setting we consider below.

For the results reported below we consider an ordered probit specification in which the null DGP is given by $\zeta = \beta_0 + \beta_1 x_i + \beta_2 z_i + u_i$ with $x_i \sim N(0,1)$, $u_i \sim N(0,0.25)$, and $z_i \in \{0,1\}$ with $\Pr(z_i = 1) = 0.5$. We generate $y_i \in \{0,1,2,\dots,c-1\}$ with $c = c(3,6,7,9)$ where y_i is determined by the value of ζ_i lying in equal probability intervals defined by the value of c and cutoff values from the normal distribution. For instance, when $c = 3$, $y_i \in \{0,1,2\}$ when $\zeta_i \in [(-\infty,-0.43),[-0.43,0.43),[0.43,\infty)]$ representing 3 equally probable intervals. By changing the value of c we can assess the impact of smoothing y as the number of unique outcomes increases.

The alternative DGP is given by $\zeta = \beta_0 + \beta_1 x_i + \beta_2 z_i + \beta_3 \sin(\pi x_i/2) + u_i$. We let $\beta_3 \in \{0,1\}$. When $\beta_3 = 0$ we can investigate the test's size, while when

$\beta_3 = 1$ we investigate the test's power.

We conduct $B = 399$ bootstrap replications and $M = 1{,}000$ Monte Carlo replications and report results for sample sizes $n = (200{,}300{,}400)$. We report empirical rejection frequencies in tables 1 and 2 for levels $\alpha = (0.01{,}0.05{,}0.10)$.

Tables 1 and 2 reveal that the smoothed y test has reasonable size and also size comparable to the frequency y comparator. Furthermore, power for the smoothed and unsmoothed versions of the test increase with both n and c. However, it is evident that power for the smoothed version of the test proposed herein is substantially higher than that for the unsmoothed version thereby underscoring the utility of smoothing the outcome in conditional probability models in finite-sample settings.

4.4 Conclusion

Estimating a conditional density is one of the most important tasks in econometrics. Consequently, testing the correct specification of a parametric conditional density is of crucial importance for applied researchers.

In this paper, we generalize the nonparametric conditional density specification test proposed by Fan, Li, and Min (2006) to the extent that we also smooth the discrete dependent variable using data-driven bandwidth selection. The optimal data-driven bandwidth selection (i.e., least squares cross-validation method) is of crucial importance in nonparametric estimation and testing of econometric models. The least squares cross validation method has many advantages such as automatically removing irrelevant covariates. When there

exist discrete variables, it avoids sample splitting and often leads to substantial efficiency gains versus the conventional nonparametric frequency method. We establish the asymptotic null distribution of the test statistic with data-driven selected stochastic smoothing parameters. We also show that the test is consistent when the null hypothesis fails to hold. The method is applicable to popular parametric binary choice models such as the logit and probit specification and their multinomial and ordered counterparts along with parametric count models, among others. The test is valid when the conditional density function contains both categorical and real-valued covariates. Monte Carlo simulation results show that by smoothing the discrete dependent variable our proposed test enjoys substantial finite sample power gains compared with the test statistic proposed by Fan, Li and Min (2006).

References

Aitchison, J. and C. G. G. Aitken (1976), Multivariate binary discrimination by the kernel method, Biometrika 63, 413-420.

Fan, Y., Q. Li and I. Min (2006), A nonparametric bootstrap test of conditional distributions, Econometric Theory 22, 587-613.

Hall, P., J. Racine and Q. Li (2004), Cross-validation and the estimation of conditional probability densities, Journal of the American Statistical Association 99, 1015-1026.

Hausman, J. and D. McFadden (1984), Specification tests for the multinomial logit model, Econometrica 52(5), 1219-1240.

Mora, J. and A. I. Moro-Egido (2008), On specification testing of ordered discrete choice models, Journal of Econometrics 143, 191-205.

Rodriguez-Poo, J. M., S. Sperlich and P. Vieu (2004), An adaptive specification test for semiparametric models, SSRN eLibrary .

Zheng, J. X. (2000), A consistent test of conditional parametric distributions, Econometric Theory 16, 667-691.

Table 1. Size of the tests

		Smoothed y			Frequency y		
n	c	**0.010**	**0.050**	**0.100**	**0.010**	**0.050**	**0.100**
200	**3**	0.015	0.074	0.132	0.014	0.072	0.129
300	**3**	0.012	0.060	0.116	0.013	0.058	0.115
400	**3**	0.014	0.063	0.120	0.014	0.063	0.120
200	**5**	0.010	0.071	0.130	0.010	0.072	0.129
300	**5**	0.016	0.065	0.122	0.015	0.064	0.123
400	**5**	0.020	0.055	0.115	0.019	0.054	0.115
200	**7**	0.008	0.062	0.121	0.007	0.060	0.124
300	**7**	0.013	0.058	0.108	0.014	0.058	0.108
400	**7**	0.007	0.045	0.100	0.007	0.045	0.099
200	**9**	0.014	0.059	0.120	0.015	0.059	0.118
300	**9**	0.008	0.055	0.117	0.008	0.051	0.114
400	**9**	0.011	0.063	0.116	0.011	0.061	0.115

Notes: entries represent empirical rejection frequencies for the proposed bootstrap test of correct specification under the null. Columns corresponding to `Smoothed y' are those for the proposed test, while columns corresponding to `Frequency y' are those corresponding to the unsmoothed test of Fan et al. (2006).

Table 2. Power of the tests

		Smoothed y			Frequency y		
$\boldsymbol{n}$	$\boldsymbol{c}$	**0.010**	**0.050**	**0.100**	**0.010**	**0.050**	**0.100**
200	**3**	0.263	0.353	0.412	0.141	0.291	0.363
300	**3**	0.408	0.485	0.537	0.205	0.368	0.450
400	**3**	0.507	0.575	0.618	0.270	0.427	0.504
200	**5**	0.370	0.502	0.581	0.191	0.357	0.463
300	**5**	0.572	0.650	0.692	0.324	0.448	0.537
400	**5**	0.696	0.768	0.798	0.419	0.572	0.646
200	**7**	0.422	0.546	0.612	0.213	0.367	0.458
300	**7**	0.604	0.687	0.720	0.354	0.495	0.554
400	**7**	0.739	0.801	0.822	0.473	0.593	0.666
200	**9**	0.438	0.579	0.647	0.245	0.398	0.474
300	**9**	0.622	0.713	0.765	0.381	0.524	0.609
400	**9**	0.768	0.827	0.851	0.537	0.652	0.706

Notes: entries represent empirical rejection frequencies for the proposed bootstrap test of correct specification under the alternative. Columns corresponding to `Smoothed y' are those for the proposed test, while columns corresponding to `Frequency y' are those corresponding to the unsmoothed test of Fan et al. (2006).

Appendix A. Proof of Main Results

In the proof of Theorems 4.2.1 and 4.2.2, we will replace $\hat{h}_1, \dots, \hat{h}_q, \hat{\lambda}_0, \dots, \hat{\lambda}_r$ by their non-stochastic leading terms: $(h_1, \dots, h_q) = (a_1^0 n^{-1/(q+4)}, \dots, a_q^0 n^{-1/(q+4)})$ and $(\lambda_0, \dots, \lambda_r) = (b_0^0 n^{-2/(q+4)}, \dots, b_r^0 n^{-2/(q+4)})$. This will greatly simplify the arguments in the proof. By Lemma A.4 of Fan et al. (2006), we know that the conclusion holds provided $\hat{h}_s - h_s = o_p(h_s)$ $(s = 1, \dots, q)$ and $\hat{\lambda}_s - \lambda_s = o_p(\lambda_s)$ $(s = 0, \dots, r)$, which are true by Lemma 4.2.1.

Proof of Theorem 4.2.1. The test statistic is given by equation (9). We write $\hat{f}_{ij} = \hat{f}(y_i|x_j) \equiv f(y_i|x_j, \hat{\theta})$ and $\hat{f}_i = \hat{f}(y_i|x_i) \equiv f(y_i|x_i, \hat{\theta})$. Then substituting $l(y_i, y_j, \lambda_0)$ in (5) into equation (9), we obtain

$$
\begin{aligned}
J_{n,\gamma} &= \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} \frac{K_{\gamma,ij}}{\hat{f}_i} [(1-\lambda_0) I_{ij} + \frac{\lambda_0}{c_0 - 1}(1 - I_{ij}) \\
&\quad - \sum_y (1-\lambda_0) \hat{f}(y|x_j) I(y_i = y) - \sum_y \frac{\lambda_0}{c_0 - 1} \hat{f}(y|x_j)(1 - I(y_i - y))] \\
&= \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} \frac{K_{\gamma,ij}}{\hat{f}_i} [(1-\lambda_0)[I_{ij} - \hat{f}_{ij}] + \frac{\lambda_0}{c_0 - 1}(1 - I_{ij}) \\
&\quad - \frac{\lambda_0}{c_0 - 1} \sum_y \hat{f}(y|x_j) + \frac{\lambda_0}{c_0 - 1} \sum_y \hat{f}(y|x_j) I(y_i = y)] \\
&= \left[1 - \lambda_0 - \frac{\lambda_0}{c_0 - 1}\right] \left[\frac{1}{n(n-1)} \sum_i \sum_{j \neq i} \frac{K_{\gamma,ij}}{\hat{f}_i} (I_{ij} - \hat{f}_{ij})\right]
\end{aligned}
$$

$$= \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} \frac{K_{\gamma,ij}}{\hat{f}_i} \left[I_{ij} - \hat{f}_{ij} \right] + (s.o.)$$

$$= \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} \frac{K_{\gamma,ij}}{f_i} \left[I_{ij} - f_{ij} \right] + (s.o.) \qquad (10)$$

where $(s.o.)$ denotes smaller order terms, $f_{ij} = f(y_i|x_j, \theta_0)$ and $f_i = f(y_i|x_i, \theta_0)$. The fourth equality follows because $\lambda_0 = O\left(n^{-2/4+q}\right) = o(1)$, while the last equality follows from a Taylor expansion argument (see Fan et al. (2006) for detailed derivations).

We see that the leading term of our statistic given in (10) is the same as the leading term of Fan et al.'s (2006) test statistic. Hence, Theorem 4.2.1 follows from Theorem 2.1 of Fan et al. (2006) with $\hat{V}_a = \frac{2(h_1 \ldots h_q)}{n(n-1)} \sum_i \sum_{j \neq i} \{ K_{\hat{\gamma},ij} \left[I\left(y_i = y_j\right) - f\left(y_i \middle| x_j, \hat{\theta}\right) \right] / \hat{f}\left(y_i \middle| x_i, \hat{\theta}\right) \}^2$, which is a consistent estimator of $V_a = 2\left[\int W^2(v)dv\right] E\left[\left(1 - f(y_i|x_i, \theta_0)\right) f^{-1}(y_i|x_i, \theta_0) p_1(x_i)\right]$ (the asymptotic variance of $n\left(\hat{h}_1 \ldots \hat{h}_q\right)^{1/2} J_{n,\hat{\gamma}}$). This completes the proof of Theorem 4.2.1.

Proof of Theorem 4.2.2. When y is an ordered discrete variable, we use the Wang and van Ryzin kernel defined in (6). Substituting the ordered kernel function into equation (9), we have

$$J_{n,\gamma} = \frac{1 - \lambda_0}{n(n-1)} \sum_i \sum_{j \neq i} \frac{K_{\gamma,ij}}{\hat{f}_i} \{ I_{ij} + \frac{1}{2} I\left(y_i \neq y_j\right) \lambda_0^{|y_i - y_j|}$$

$$-\sum_{y}\hat{f}\,(y|x_j)[I(y_i=y)+\frac{1}{2}I(y_i\neq y)\lambda_0^{|y_i-y|}]\}$$

$$=J_{n1}-\lambda_0 J_{n1}$$

Where $J_{n1}=\frac{1}{n(n-1)}\sum_i\sum_{j\neq i}\frac{K_{\gamma,ij}}{\hat{f}_i}\{I_{ij}+\frac{1}{2}I(y_i\neq y_j)\lambda_0^{|y_i-y_j|}-\sum_y\hat{f}(y|x_j)\Big[I(y_i=y)+\frac{1}{2}I(y\neq y_i)\lambda_0^{|y_i-y|}\Big]\}$. The leading term of $J_{n,\gamma}$ is J_{n1}, because $\lambda_0=o(1)$. Therefore, we have

$$J_{n1}=\frac{1}{n(n-1)}\sum_i\sum_{j\neq i}\frac{K_{\gamma,ij}}{f_i}[I_{ij}-f_{ij}]$$

$$+\frac{1}{2n(n-1)}\sum_i\sum_{j\neq i}\frac{K_{\gamma,ij}}{f_i}\{I(y_i\neq y_j)\lambda_0^{|y_i-y_j|}$$

$$-\sum_y f(y|x_j)\,I(y\neq y_i)\lambda_0^{|y_i-y|}\}+(s.o.)$$

$$=J_{n1,1}+J_{n1,2}+(s.o.)$$

where the definition of $J_{n1,j}$ $(j=1,2)$ should be apparent, $(s.o.)$ denotes the smaller order terms of the expansion of $1/\hat{f}_i$ and the Taylor expansion of $\hat{f}(y|x_j)$. We see that $J_{n1,1}$ is identical to the leading term in (10). Hence, $n\sqrt{h_1\ldots h_q}J_{n1,1}\xrightarrow{d}N(0,V_a)$ by Theorem 4.2.1.

$J_{n1,2}$ can be written as second order U-statistic

$$J_{n1,2}=\frac{1}{2n(n-1)}\sum_i\sum_{j\neq i}J_{n,ij}=\frac{2}{n(n-1)}\sum_i\sum_{j>i}H_{n,ij} \qquad (11)$$

where $J_{n,ij}=\frac{K_{\gamma,ij}}{f_i}[I(y_i\neq y_j)\lambda_0^{|y_i-y_j|}-\sum_y\lambda_0^{|y_i-y|}f(y|x_j)I(y\neq y_i)]$ and

$H_{n,ij} = [J_{n,ij} + J_{n,ji}]/4$.

We will use the U-statistic's H-decomposition to analyze $J_{n1,2}$. We first evaluate $E[H_{n,ij}|z_i] = (1/4)E[J_{n,ij} + J_{n,ji}|z_i]$, where below we use the short-hand notation $\int [\dots]dx_j = \sum_{x_j^d}[\dots]dx_j^c$.

$$E[J_{n,ij}|z_i] = \frac{1}{f_i}\sum_{y_j}\int p(y_j, x_j)\, K_{\gamma,ij}\lambda_0^{|y_i - y_j|} I(y_i \neq y_j)dx_j$$

$$-\frac{1}{f_i}\sum_{y}\int p_1(x_j)\, f(y|x_j) K_{\gamma,ij}\lambda_0^{|y_i - y|} I(y \neq y_i)dx_j$$

$$= 0$$

because $p_1(x_j)f(y|x_j) = p(y, x_j)$.

Next, for $E[J_{n,ji}|z_i]$, we have

$$E[J_{n,ji}|z_i] = \sum_{y_j}\int p(y_j, x_j)\frac{K_{\gamma,ij}}{f_j}\lambda_0^{|y_i - y_j|} I(y_i \neq y_j)dx_j$$

$$-\sum_{y_j}\int p(y_j, x_j)\frac{K_{\gamma,ij}}{f_j}\sum_{y}\lambda_0^{|y_j - y|} f(y|x_i, \theta_0) I(y \neq y_j)dx_j$$

$$= A_{1n} - A_{2n}.$$

Define $m(y_i) = \sum_{y_j}\lambda_0^{|y_i - y_j|} I(y_i \neq y_j)$. Recalling that $f_j = f(y_j|x_j)$, we have

$$A_{1n} = m(y_i)\int p_1(x_j)K_{\gamma,ij}dx_j = m(y_i)E[K_{\gamma,ij}|x_i]$$

and

$$A_{2n} = \sum_{y}[f(y|x_i)m(y)]\int p_1(x_j)K_{\gamma,ij}dx_j$$

$$= E[m(y_i)|x_i]E\left[K_{\gamma,ij}\middle|x_i\right].$$

Therefore, we have shown that $E\left[J_{n,ji}\middle|z_i\right] = \epsilon_i E[K_{\gamma,ij}|x_i]$, where $\epsilon_i = m(y_i) - E[m(y_i)|x_i]$.

Note that $m(y_i) = \sum_y \lambda_0^{|y_i - y|} I(y_i \neq y) = \lambda_0 \sum_y \lambda_0^{|y-y_i|-1} I(|y - y_i| \geq 1) = \lambda_0 \xi(y_i)$, where $\xi(y_i) = \sum_y \lambda_0^{|y-y_i|-1} I(|y - y_i| \geq 1)$. Hence, $\epsilon_i = m(y_i) - E[m(y_i)|x_i] = \lambda_0 \upsilon_i$, where $\upsilon_i = \xi(y_i) - E[\xi(y_i)|x_i]$. Hence, we have shown that

$$E\left[J_{n,ji}\middle|z_i\right] = \lambda_0 \upsilon_i E\left[K_{\gamma,ij}\middle|x_i\right].$$

$E[\upsilon_i|x_i] = 0$ implies that $E\left[J_{n,ji}\right] = 0$. Also, $E\left[J_{n,ij}\middle|z_i\right] = 0$ implies that $E\left[J_{n,ij}\right] = 0$. Hence, $E\left[H_{n,ij}\right] = 0$. Thus, by H-decomposition we have

$$J_{n1,2} = \frac{2}{n(n-1)} \sum_i \sum_{j>i} H_n\left(z_i, z_j\right)$$

$$= \frac{\lambda_0}{2n} \sum_{i=1}^{n} \upsilon_i \, E\left[K_{\gamma,ij}\middle|x_i\right] + (s.o.) \qquad (12)$$

where the partial term has an order $O_p(n^{\frac{-8-q}{8+2q}})$, $(s.o.) = O_p(n^{\frac{-12-q}{8+2q}})$.

By the Liapunov CLT, we have that

$$\frac{\sqrt{n}}{\lambda_0} J_{n1,2} \xrightarrow{d} N(0, V_2) \qquad (13)$$

Where $V_2 = (1/4)E[\upsilon_i^2 p_1^2(x_i)] + o(1)$ because $Var\left(\upsilon_i E\left[K_{\gamma,ij}\middle|x_i\right]\right) = E\left[\upsilon_i^2 E^2\left(K_{\gamma,ij}\middle|x_i\right)\right] = E[\upsilon_i^2 p_1^2(x_i)] + o(1)$.

To show that $Cov\left(J_{n1,1}, J_{n1,2}^0\right) = 0$, where $J_{n1,2}^0$ is the leading term of $J_{n1,2}$, we use the fact that $J_{n1,1}$ can be written as a degenerate U-statistic, i.e., $J_{n1,1} =$

$\frac{2}{n(n-1)}\sum_i \sum_{j>i} H_{n,ij}^a$, where $H_{n,ij}^a = \frac{K_{\gamma,ij}}{2}\left[\frac{I_{ij}-f_{ij}}{f_i} + \frac{I_{ji}-f_{ji}}{f_j}\right]$.

It is easy to verify that $E\left[H_{n,ij}^a \middle| z_i\right] = 0$, $z_i = (x_i, y_i)$ .Hence, $Cov\left[H_{n,ij}^a, v_i p_1(x_i)\right] = E\{v_i p_1(x_i) E\left[H_{n,ij}^a \middle| z_i\right] = 0$. Therefore, $Cov\left(J_{n1,1}, J_{n1,2}^0\right) = 0$. $J_{n1,1}$ and $J_{n1,2}^0$ are asymptotically independent.

Thus we have that $n\sqrt{h_1 \dots h_q} J_{n,\gamma} \xrightarrow{d} N(0, V_b)$, where $V_b = V_a + c_0^2 V_2$ with $c_0 = \left(a_1^0 \dots a_q^0\right)^{1/2} b_0^0 = \lim_{n\to\infty} n\left(h_1 \dots h_q\right)^{1/2} \lambda_0/\sqrt{n}$ (a_s^0 and b_0^0 are defined in Lemma 2.1). A consistent estimator of V_b is $\hat{V}_b = \hat{V}_a + \hat{c}_0^2 \hat{V}_2$, where $\hat{c}_0 = \left(n\hat{h}_1 \dots \hat{h}_q\right)^{1/2} \hat{\lambda}_0$, $\hat{V}_2 = \frac{1}{4n}\sum_{i=1}^n \hat{v}_i^2\, \hat{p}_1^2(x_i)$, $\hat{v}_i = [\hat{\xi}(y_i) - \sum_y \hat{\xi}(y_i) f(y|x_i, \hat{\theta})]$ and $\hat{\xi}(y_i) = \sum_y \hat{\lambda}_0^{|y_i - y|-1} I(|y - y_i| \geq 1)$.

Chapter 5

Conclusion

This thesis contains two parts. The first and the second essays focus on the economics of smoking and the last essay is about nonparametric econometrics. Cigarette smoking has been consistently considered as an important risk factor for public health. Research on the impact of anti-smoking policies, like the cigarette excise tax, on smokers' behaviors have received substantial interest over the last two decades.

In the first essay, we study the impact of cigarette taxes on smokers' pack versus carton purchasing behavior. Because carton purchase is associated with a substantial quantity discount, smokers can potentially avoid higher tax-induced prices by purchasing cartons, as suggested in the public health literature. First, we

estimate a two-way fixed effect model for all smokers, and find that, instead of carton purchasing, smokers tend to switch to single pack purchasing in response to tax increases, although the impact is not statistically significantly different from zero. Then, we perform selected subgroup analyses to investigate why smokers do not switch to cartons when they experience tax increases. The initial subgroup analyses are done by smoking frequency and income levels of smokers, because high frequency smokers may have stronger incentive to realize the quantity discount associated with carton purchasing, while the relatively greater expenditure required for carton purchasing may make poorer smokers less likely to purchase cartons. The corresponding estimates, however, do not provide consistent evidence supporting the above hypotheses. We then conduct further subgroup analyses by smokers' intentions to quit smoking consistent with notions from the behavioral economics literature. Estimates by quitting intentions give us clear and consistent evidence. For smokers who say that they want to quit smoking, higher taxes significantly cause smokers to substitute to pack purchase, consistent with behavior economics literature, while for smokers who do not want to quit smoking, higher taxes result in greater likelihood of carton purchasing, consistent with public health literature, which suggests that they might do so to take advantage of the quantity discount. In addition to these findings, our results suggest that future studies should investigate whether other tax avoidance behaviors also depends on quitting intentions.

The second essay focuses on the impact of the cigarette excise tax on the

smoking participation of older adults, particularly those aged 45 to 59. This is an interesting topic, since quitting smoking even at an older age can still generate substantial health benefits according to available evidence. We first explore the effect of the cigarette excise tax on the quitting intentions of older smokers, and then investigate its effect on actual smoking participation. Overall, we find that tax increases have sizeable effects on intentions to quit smoking. And, contrary to the existing literature, we find tax increases have significant effects in reducing the smoking participation of these older, longer-term smokers. We further analyze the heterogeneity of the effect of tax. Consistent with expectations, we find that smokers with lower income are much more likely to quit smoking in response to tax increases than higher-income smokers. This pattern is even stronger when we compare more-educated vs. less-educated individuals, since education is likely a better proxy for permanent income. The strong effect on the smoking behaviour of less-educated individuals might suggest additional health gains from smoking cessation, to the extent that lower SES individuals are more susceptible to ill-health and have issues of health care access. Finally, we allow the effect of cigarette taxes on smoking behavior to vary by gender, since this is a conventional breakdown in the literature. While tax effects are somewhat larger for males, both genders are similarly affected.

The last essay generalizes the nonparametric model specification test proposed by Fan et al. (2006). In particular, we smooth the discrete dependent variable. By smoothing the discrete variable, the method avoids sample splitting

and may provide efficiency gains. We derive the distribution of the test statistics under the null hypothesis and show that the test is consistent even when the null hypothesis fails to hold. Finally, a Monte Carlo study is conducted to evaluate the finite sample performance of the smoothed test statistics. From the simulation, we find that by smoothing discrete dependent variable, the test enjoys substantial power gains compared with the conventional frequency method.

The first two essays add to research on the economics of smoking. However, in order to fully understand the role of cigarette excise tax in the anti-smoking policies, more research is needed. The last essay extends the use of smooth method and provides new evidence of the benefit of cross validation in selecting bandwidths. Further generalization of this method into panel data model would be of special interest to applied researchers.