ESSAYS ON PARENTS' SOCIOECONOMIC STATUS, CHILD HEALTH OUTCOMES AND SMOKING BEHAVIOURS

Ijeoma Peace Edoka

PhD Thesis

University of York
Department of Economics and Related Studies

March, 2013

Abstract

This thesis comprises a collection of three distinct essays on the relationship between parents' socioeconomic, child health outcomes and smoking behaviours.

Chapter 2 investigates the extent to which misclassification errors in self-reported smoking affects estimates of the impact of parental income on smoking in adolescents aged 11-15 years old. Smoking participation is modelled using self-reported smoking and cotinine-validated smoking as binary dependent variables in two separate probit models. A comparison of the marginal effects estimated from both models suggest that self-reported smoking is misreported leading to biased estimates of the impact of parental income on adolescent smoking. Estimates from the cotinine-validated smoking model are robust to different specifications of the model that account for exposure to second-hand smoke. Income-related inequality in smoking (the concentration index) is also underestimated due to variations in the extent of misclassification errors across income quantiles.

Chapter 3 uses three decomposition methods to decompose differences in the distribution of saliva cotinine between children/adolescents from high and low socioeconomic backgrounds. The decomposition methods applied are a mean-based (Oaxaca-Blinder) decomposition method and two decomposition methods that allow the decomposition of differences in quantiles (the quantile regression and recentered influence function regression decomposition methods). Group differences in the distribution of characteristics (composition effect) accounts for a larger proportion of the total difference in log cotinine compared to group differences in the impact of these characteristics on smoking (structural effect). The composition effect attributable to smoking within the home explains more of socioeconomic differences at lower quantiles, which are indicative of passive smoking compared to higher quantiles, which are indicative of active smoking. On the other hand, the composition effect of household income and parental smoking explains more of the socioeconomic differences in active smoking compared to passive smoking.

Chapter 4 uses the Vietnam Young Lives Survey to investigate the impact of small-scale weather shocks on child nutritional status as well as the mechanism through which weather shocks affect child nutritional status. The results shows that small-scale weather shocks negatively affect child nutritional status and total household per capita consumption and expenditure (PCCE) but not food PCCE. Disaggregating total food PCCE into consumption

of high-nutrient and energy-rich food shows that households protect food consumption by decreasing consumption of high-nutrient food and increasing consumption of affordable but low quality food. This suggests that the impact of small-scale weather shocks on child health is mediated through a reduction in the quality of dietary intake. Finally, chapter 4 shows evidence of a differential impact of weather shocks in children from different socioeconomic backgrounds. However, contrary to other studies, the impact of weather shocks is observed to be greater amongst children from wealthier households compared to children from poorer households.

Table of Contents

List of Tables

List of Figures

Acknowledgement

The success of this thesis is largely due to the support I received from my supervisors, my friends and family.

I thank my supervisors Professor Andrew Jones and Professor Nigel Rice and the third member of my Thesis Advisory Panel, Dr. Stephanie von Hinke Kessler Scholder for their guidance throughout the entire process of completing this work. I thank Professor Mireia Jofre-Bonet, whose support and encouragement made all the difference at the initial stages of my study.

My immense gratitude goes to my family, my loving mother, my brothers, my sister and sister-in-law, who always believed in me even when I had doubts. I thank them for their unconditional love and support and for inspiring me to achieve and never give up.

I dedicate this thesis to the loving memory of my father, Professor Benson Eziukwu Edoka, whose legacy still lives on in the pace he set for his children. I know he would have been very proud.

Finally, in deepest gratitude, I give thanks to the Most High, without Whose Grace and Mercy this work would never have come to completion.

Declaration

All chapters comprising this thesis are single-authored. Chapters 2 and 3 are available as Health Econometric and Data Group (HEDG), University of York, working papers[1].

Chapters 2 and 3 use data provided by the UK Economic and Social Data Service (ESDS), while chapter 4 uses data provided by Young Lives. I take sole responsibility for any errors and omissions therein.

The thesis was fully funded by the Economic and Social Research Council (ESRC).

[1] The working papers are available at:
http://www.york.ac.uk/res/herc/documents/wp/11_29.pdf
http://www.york.ac.uk/res/herc/documents/wp/12_29.pdf

Chapter 1

Introduction

This thesis comprises a collection of three distinct essays broadly covering themes on socioeconomic determinants of child/adolescent smoking behaviours and child health outcomes. The first two essays investigate the relationship between parents' socioeconomic status and child/adolescent smoking behaviours, while the third essay investigates the relationships between small-scale weather shocks, household consumption and child health. The following paragraphs briefly discuss the motivation for the empirical analyses implemented in each essay and outline the structure of the entire thesis.

In developed countries, such as in the United Kingdom, adolescent smoking is increasingly becoming an important health policy target, not least because strong associations have been demonstrated between early age of smoking initiation and smoking into adulthood as well as the adverse health consequences both in adolescence and during adulthood (Department of Health, 2010, The NHS Information Centre, 2010). Therefore, the focus of much research has been directed towards understanding the determinants of smoking initiation and participation amongst children and adolescents as well as in quantifying the impact of these determinants in order to inform anti-smoking laws[2].

The impact of a range of factors on smoking initiation and participation have been extensively investigated including the impact of tobacco price increases, peer influence, parental smoking behaviours and parental socioeconomic status. However, mixed findings have often been reported both with the presence and magnitude of an impact. For example while some empirical studies have shown that adolescent smoking is responsive to price increases[3], other studies have failed to reach similar conclusions[4]. In addition, differences in the magnitude of the price elasticity of adolescent smoking have been reported amongst empirical studies supporting a price responsiveness of adolescent smoking (some examples include Chaloupka and Grossman (1996), Chaloupka and Pacula (1998), Harris and Chan (1999), Emery et al. (2001), Ross and Chaloupka (2003)).

[2] For example, see Gruber and Zinman (2001).
[3] Some examples include Chaloupka and Grossman (1996), Chaloupka and Pacula (1998), Harris and Chan (1999), Emery et al. (2001), Ross and Chaloupka (2003).
[4] Some examples include Wasserman et al. (1991), Douglas and Hariharan (1994), DeCicca et al. (2002).

Similarly, studies investigating the relationships between parental socioeconomic status and adolescent smoking behaviours have also failed to report unanimous findings. While some studies report a negative association between parental socioeconomic status or parental income and the probability of smoking initiation and participation amongst adolescents (Tyas and Pederson, 1998, Soteriades and DiFranza, 2003, Gruber and Zinman, 2001), other studies do not observe an association (Blow et al., 2005). Although these divergent findings may reflect contextual differences between study populations and/or methodological differences between studies, the extent of misclassification errors within adolescent self-reported smoking data may partly explain these conflicting conclusions.

Smoking participation models typical use a binary 0/1 dependent variable where 1 represents smokers and 0 represents non-smokers. Under classical assumptions, measurement error in a continuous dependent variable may result in less statistical precision in the estimation of the coefficients using ordinary least squares, but does not lead to biased estimates of coefficients (Hausman, 2001). On the other hand, misclassification error in a binary 0/1 dependent variable will lead to biased estimates of the coefficients in a maximum likelihood estimation approach (Bound et al., 2001, Hausman et al., 1998). Therefore, when an adolescent misreports their smoking status, a positive response (smoker or 1) is miscoded as a negative response (non-smoker or 0) or vice versa and could result in biased and inconsistent estimates of the impact of the determinants of adolescent smoking.

The consequences of misclassification errors have been widely reported in other literatures including labour economics (Falaris, 2011, Hausman et al., 1998), epidemiology (Magder and Hughes, 1997, Höfler, 2005) and insurance (Artís et al., 2002). However, in health economics research, particularly in smoking participation models, the consequences of misclassification errors have received very little attention. One exception is a study by Kenkel et al. (2004) which showed that in a probit model, estimates of the effect of cigarette prices on adult smoking participation are biased when smoking participation is misclassified. Misclassification errors in adolescent smoking is likely to pose an even greater problem given high levels of inconsistencies observed in adolescent self-reported smoking behaviours in studies comparing self-reported smoking to objective biochemical indicators of smoking (Craig and Mindell, 2008, Wagenknecht et al., 1992, Kandel et al., 2006).

Chapter 2 contributes to this literature by highlighting the implications of misclassification errors in empirical studies of adolescent smoking participation, focussing largely on the

implications for the relationship between parents' socioeconomic status and adolescent smoking. Chapter 2 uses pooled cross-sections of the Health Survey for England (HSE) which contains both a self-reported smoking and an objective measure of smoking which was obtained from saliva cotinine assays. Cotinine is a biomarker of the extent of exposure to second-hand smoke and a quantitative indicator of active smoking. Saliva cotinine concentrations greater than or equal to 12ng/ml identifies active smoking with a high sensitivity of approximately 97% (Jarvis et al., 2008, Environmental Protection Agency, 1997). Therefore in chapter 2, cotinine-validated smokers are defined as those with saliva cotinine greater than or equal to 12ng/ml. This is used as the bench-mark or 'true' model of adolescent smoking participation and the magnitude of the impact of parental income and other characteristics on adolescent self-reported smoking is verified using this 'true' model. The results suggest that self-reported smoking is misreported resulting in biased estimates of the impact of parental income on adolescent smoking. In addition, income-related inequalities in smoking measured using the concentration index, are underestimated due to variations in the extent of misclassification errors across income quantiles.

Chapter 3 builds on the theme of chapter 2 by estimating the contributions of different determinants to socioeconomic differences in child and adolescent smoking. Chapter 3 uses the 1997/98 cross-section of the HSE and takes advantage of the entire distribution of saliva cotinine to decompose differences in the distribution of cotinine between two groups of children and adolescent. The two groups are defined based on the social class of the household head which was assigned using the Registrars General's Social Class (RGSC) classification system. Children living in households where the head of the household has a professional, managerial or technical occupation are classified as the high social class group, while those living in households where the head of the household holds a partly skilled, unskilled or any other occupation are classified as the low social class group.

The decomposition methods applied here allows socioeconomic differences in the distribution of saliva cotinine to be decomposed into a part explained by group differences in the distribution of characteristics (composition effect), and a part explained by group differences in the impact of these characteristics (structural effect). First, a mean-based decomposition method, the Oaxaca-Blinder decomposition method (Blinder, 1973, Oaxaca, 1973) is applied to decompose differences in mean (log) cotinine. Since lower quantiles of the (log) cotinine distribution is likely to comprise of non-smokers with moderate exposure to second-hand

smoke (passive smokers) and higher quantiles of active smokers (Jarvis et al. 2008), the empirical analysis is extended to decompose differences between quantiles of (log) cotinine (Firpo et al., 2009, Melly, 2005). The decomposition of quantiles allows further insights into variations in the extent to which different characteristics contribute to differences in passive and/or active smoking amongst children and adolescents from high and low socioeconomic backgrounds. The results show that the contributions made by different characteristics vary at different quantiles of the (log) cotinine distribution. For example, smoking within the home explains more of the socioeconomic difference at lower quantiles (passive smoking) and less of the difference at higher quantiles (active smoking). Conversely, parental smoking and income explains more of the difference at the higher quantiles (active smoking) compared to the contribution they make at lower quantiles (passive smoking).

Chapter 4 departs somewhat from the themes of chapters 2 and 3 (child/adolescent smoking behaviours) and investigates the mechanism through which small-scale weather shocks affect child nutritional status. Weather shocks in childhood have been implicated in several adverse child health outcomes. For example, following extreme drought in Zimbabwe, exposed children experienced slower growth rates (Hoddinott and Kinsey, 2001), the 1997 forest fire in Southeast Asia resulted in higher infant and child mortality in Indonesia (Jayachandran, 2009), while the 1998 Hurricane Mitch affecting large parts of Central America was associated with an increase in the prevalence of wasting and malnutrition amongst affected children in Honduras and Nicaragua (Barrios et al., 2000).

Previous studies have focused mainly on the impact of single large-scale weather shocks or natural disasters on child health, with fewer studies on the impact of smaller-scale weather shocks. However, repeated exposure to small-scale weather shocks is likely to have significant consequences for child health outcomes (Pörtner, 2010; Datar et al., 2011). In developing countries, weather shocks have been shown to result in significant reductions in both agricultural and non-agricultural wages several years after the occurrence (Mueller and Quisumbing, 2010, Mueller and Osgood, 2009, Jayachandran, 2006, Thomas et al., 2010). Therefore weather-induced household income shock is likely to represent an important mechanism through which small-scale weather shocks affect child health.

Chapter 4 uses the 2006 and 2009 panels of the Vietnam Young Lives Surveys (VYLS), which consist of a pro-poor sample of children aged 4 and 12 years in 2006. The impact of small-scale weather shocks on child nutritional status and household consumption is modelled using

community fixed effect models to control for time invariant unobserved community characteristics that may affect both the probability of exposure to weather shocks and child health/household consumption.

Consistent with other studies, the results show that small-scale weather shocks including droughts, excessive rainfall or floods, erosions, landslides, frosts and storms are associated with poorer child nutritional status and lower total household per capita consumption and expenditure (PCCE). Total PCCE on food appears to be unaffected by weather shocks. However, disaggregating total food PCCE into PCCE on energy-rich (or low nutrient) food and PCCE on micronutrient-rich (or high nutrient) food shows a decrease in household consumption of high-nutrient food and an increase in the consumption of energy-rich food. This is indicative of a lower quality of dietary intake amongst exposed households. Since child nutritional status is a function of the quality of dietary intake, lower quality of dietary intake amongst exposed households provides a strong explanation for the impact of small-scale weather shocks on child nutritional status.

The results also show some evidence of a heterogeneous impact of small-scale weather shocks. However, contrary to other studies[5], the results suggest that the impact of small-scale weather shock is greatest amongst children living in wealthier households compared to children from poorer households. Given that the VYLS consists predominantly of disadvantaged households, wealthier households within the VYLS are likely to still face huge budget constraints and/or possess limited capacity to smooth consumption in response to household consumption shocks. On the other hand, for poorer households, exposure to weather shocks may be one of a host of other risky environmental conditions to which children are exposed. These conditions may have a greater impact on child nutritional status, thus explaining the failure to observe a significant impact of weather shocks in children living in poorer household.

Finally, the last chapter, chapter 5, concludes the thesis by summarising the key findings of the preceding chapters.

[5] An example is Hoddinott and Kinsey (2001).

Chapter 2

Parental Income and Adolescent Smoking Participation: Implications of Misclassification Errors in Empirical Studies of Adolescent Smoking Participation

2.1 Introduction

Tobacco smoking is recognized as the single largest cause of preventable diseases and deaths and is responsible for approximately 26 million years of life lost globally (Lopez, 2005). In England, an estimated 18% of deaths in adults aged 35 years and above were attributable to smoking in 2009 (The NHS Information Centre, 2010). Adolescent smoking is increasingly becoming an important target for health policies in developed countries because a high proportion of long-term adult smokers initiate smoking as children or adolescents. Over the past decade, the UK government has put great efforts into tackling smoking amongst children and adolescents by implementing various anti-smoking policies including the 2003/2004 ban of tobacco advertising and sponsorship in print, on billboards and on the internet, and an increase in the minimum tobacco purchase age from 16 to 18 years as well as a ban on smoking in public places in 2007 (Department of Health, 2010). One of the main aims of the February 2010 White paper, *A Smoke-free Future*, was to 'stop the inflow of young people recruited as smokers'. This White paper set out targets to reduce smoking initiation rates to 1% amongst 11-15-year-olds and to 8% among 16-17-year-olds by 2020. An understanding of the determinants of adolescent smoking initiation and participation is important to aid anti-smoking policies, and a wide range of studies have focused on investigating the role these determinants play in explaining adolescent smoking behaviours.

The impact of anti-smoking policies as well as the impact of other determinants of smoking is estimated using self-reported smoking data. However, high inconsistencies have often been found in self-reported adolescent smoking in studies comparing self-reported smoking to objective biochemical indicators of smoking, raising important questions about the validity of data (Craig and Mindell, 2008, Wagenknecht et al., 1992, Kandel et al., 2006). Misreporting of smoking behaviours has important consequences in empirical economic analyses when self-reported smoking is used as a binary 0/1 dependent variable. A binary variable is misclassified

when a one is miscoded as zero or vice versa. Thus, when an adolescent misreports their current smoking status, a positive response (one) is miscoded as a negative response (zero) or vice versa. In linear models, under classical assumptions, measurement error in a continuous dependent variable may result in less statistical precision in the estimation of the coefficients but does not lead to biased estimates of coefficients (Hausman, 2001). On the other hand, misclassification error in a binary 0/1 dependent variables will lead to biased and inconsistent estimates of the coefficients in a maximum likelihood estimation (MLE) approach (Bound et al., 2001, Hausman et al., 1998). Therefore estimates of the impact of the determinants of smoking initiation or participation as well as estimates of the impact of anti-smoking policies are likely to be biased in the presence of misclassification errors in self-reported smoking data.

Inconsistencies in reporting smoking participation may also undermine studies estimating inequalities in smoking between sub-groups, if there are systemic differences in misreporting patterns between the groups being compared. For example, Bauman and Ennett (1994) showed that under-reporting of smoking behaviour amongst African American adolescents partly accounted for the large racial difference typically seen in the prevalence of self-reported smoking between white and black adolescents. In terms of income-related inequalities in smoking, if misreporting of smoking behaviours vary with income, estimates of income-related inequality in smoking measurements are likely to be biased.

Smoking is concentrated in adults with lower education, lower income and lower socioeconomic status. In addition, long term smokers who initiate smoking at an earlier age are more concentrated in individuals from lower socioeconomic backgrounds. While the negative association between smoking and income is well established in adults, reports on the impact of parental socioeconomic status on adolescent smoking behaviours have often been mixed[6]. These studies use parental education, social class or income as proxies for parental socioeconomic status. While some studies conclude that higher parental socioeconomic status and income are negatively associated with the probability of smoking initiation and participation amongst adolescents (Tyas and Pederson, 1998, Soteriades and DiFranza, 2003, Gruber and Zinman, 2001), others have reported a higher probability to smoke in high school seniors with more educated parents (Gruber and Zinman, 2001). Other studies fail to observe any association between parental income and smoking prevalence amongst adolescents (Blow et al., 2005). Blow et al. (2005) used the British Household Panel Survey (BHPS) and the

[6] Tyas and Pederson (1998) provides an extensive review of the literature.

British Youth Survey to investigate the association between adolescent smoking and parental income. After controlling for parental education and adult smoking status, no significant association was observed between parental income and smoking participation in adolescents aged between 11 and 18 years (Blow et al., 2005).

These inconsistent conclusions may partly be attributable to contextual differences in study samples and/or methodological differences between studies as well as differences in the indicator of parental socioeconomic status. Since misclassification errors by as little as 2% in a binary dependent variable can result in significant levels of bias in estimated coefficients (Hausman et al. 1998), variations in the extent misclassification errors in adolescent smoking participation across datasets may also account for these divergent conclusions.

The consequences of misclassification errors have been widely reported in labour economics (Falaris, 2011, Hausman et al., 1998), epidemiology (Magder and Hughes, 1997, Höfler, 2005) and insurance (Artís et al., 2002). However, very little attention has been paid to the consequences of misclassification error in smoking participation models in health economics research. To the best of my knowledge, only one study has investigated how misclassification errors in self-reported smoking affect estimates of the impact of cigarette price on smoking participation in adults (Kenkel et al., 2004). Kenkel et al. (2004) showed that in a probit model, estimates of the effect of cigarette prices on smoking participation are biased when smoking participation is misclassified.

This chapter makes an important contribution to the empirical literature on adolescent smoking by investigating the extent to which misclassification errors bias estimates of the impact of parental socioeconomic status (as well as other characteristics) on the probability of smoking participation in adolescents aged 11-15 years. In addition, chapter 2 investigates if there are variations in adolescent misreporting behaviours across income quantiles and how this affects measurements of income-related inequality in smoking. Data is pooled from the 1997-2008 Health Survey for England (HSE), and annual household income is used as a proxy for parental socioeconomic status. The HSE contains both self-reported smoking and an objective measure of smoking obtained from saliva cotinine assays. Saliva cotinine concentrations are used to construct a binary variable representing the 'true' indicator of adolescent smoking participation. Cotinine-validated or 'true' smokers are defined as those with saliva cotinine concentration greater than or equal to 12ng/ml (Jarvis et al., 2008). A comparison of adolescent self-reported smoking participation to cotinine-validated smoking

participation suggests that self-reported smoking is misreported, resulting in biased estimates of the impact of parental income on adolescent smoking behaviours.

Epidemiological studies using objective measures of smoking to validate self-reported smoking behaviours have shown that age, ethnicity, frequency of exposure to other smokers and smoking status of parents and friends are predictors of misreporting smoking participation amongst adolescents (Griesler et al., 2008, Kandel et al., 2006). Younger adolescents are more likely to under-report smoking while those who perceive their friends to be smokers are less likely to under-report smoking (Kandel et al., 2006, Griesler et al., 2008). African Americans are more likely to under-report smoking compared to whites (Bauman and Ennett, 1994, Griesler et al., 2008, Kandel et al., 2006). Therefore, this chapter also examines the extent to which parental income as well as other observed characteristics predict under-reporting of smoking participation amongst adolescents. Although adolescents are also prone to over-report smoking participation, this chapter focuses on under-reporting for the following reasons. Since smoking is generally considered to be socially undesirable, adolescent non-smokers are more likely to report truthfully whereas adolescent smokers are less likely to disclose their true smoking status. This is particularly true if data is collected in settings where adolescents perceive their behaviour to be unacceptable (Adams et al., 2008, Griesler et al., 2008, Kandel et al., 2006). Given that this chapter uses a household survey and that household surveys have been shown to produce significantly lower estimates of adolescent smoking rates in comparison to school-based surveys (Craig and Mindell, 2008, Griesler et al., 2008), this suggest that under-reporting is likely to represent a more important problem in this study sample compared to over-reporting of smoking participation.

The rest of chapter 2 is organised as follows: Section 2.2 describes the model of smoking participation with misclassification and quantifies the extent to which coefficients (and marginal effects) in a probit model are biased as a result of misclassification errors. A description of the HSE and variables are detailed in section 2.3. The results are presented and discussed in section 2.4 and concluding comments are provided in the section 2.5.

2.2 Model of Smoking Participation with Misclassification

Economic models of smoking behaviour are based on the economic theory of demand where the tobacco demand equation is derived from a utility maximization process in which an

individual maximizes his/her utility subject to constraints that may comprise both economic and social factors. Behavioural models have evolved to incorporate important dimensions that reflect decision-making processes unique to children and adolescents. For example, in addition to economic factors such as cigarette prices and other anti-smoking policies, social factors may influence the decision-making process of adolescents. These social factors interact to create an environment which may either reduce the perceived costs of smoking or increase the perceived benefits[7]. The tobacco demand equation can be expressed as a latent variable model:

$$Y^* = X\beta + \varepsilon \quad (1)$$

$$Y = \begin{cases} 1 \; if \; Y^* \geq 0 \\ 0 \; otherwise \end{cases} \quad (2)$$

where Y^* is a latent variable and is a linear function of a vector of covariates X and an error term, ε, which is independent and identically distributed (i.i.d). An adolescent decides to smoke (Y=1) when the net benefits of smoking (for example, social acceptability amongst peers) outweighs the costs (i.e. if $Y^* \geq 0$).

Saliva cotinine concentration greater than or equal to 12ng/ml is indicative of active smoking in adolescents (Jarvis et al., 2008). Thus, a 'true' measure of adolescent smoking participation can be defined as a binary variable, Y^{CV}, which equals one when saliva cotinine concentrations are 12ng/ml or over and zero otherwise. The bias introduced to estimates of the coefficients as a result of misclassification errors in self-reported smoking can be quantified using this 'true' measure of smoking participation following the framework of Hausman et al. (1998) as follows:

When self-reported smoking (Y^{SR}) is misreported, two misclassification probabilities, p01 and p10 can be defined: the probability that a zero is misclassified as a one (equation 3) and the probability that a one is misclassified as a zero (equation 4), respectively:

[7] A detailed discussion is provided by O'donoghue and Rabin (2001)

$$p01 = \Pr(Y^{SR} = 1 \big| Y^{\mathrm{CV}} = 0) \qquad (3)$$

$$p10 = \Pr(Y^{SR} = 0 \big| Y^{\mathrm{CV}} = 1) \qquad (4)$$

The misclassification probabilities depend only on the true value, Y^{CV}, and are assumed to be independent of the covariates, X (Bound et al., 2001, Hausman et al., 1998).

The expected values of cotinine-validated smoking participation (Y^{CV}) and the self-reported smoking participation (Y^{SR}) can then be written as:

$$E(Y^{CV}|X) = \Pr(Y^{CV} = 1|X) = F(X\beta) \qquad (5)$$

$$E(Y^{SR}|X) = \Pr(Y^{SR} = 1|X) = p01 + (1 - p01 - p10)F(X\beta) \qquad (6)^{8}$$

where F is the cumulative distribution for a probit model.

When Y^{SR} is measured without error, p01 and p10 are equal to zero and $\Pr(Y^{SR}=1|X) = \Pr(Y^{CV}=1|X) = F(X\beta)$. However, in the presence of misclassification errors, equation (6) suggests that estimates of coefficient (β) are biased towards zero. Similarly, the marginal effects of each independent variable will be biased towards zero. The downwards bias in marginal effects can also be shown by estimating the partial derivative of equation 6 with respect to any independent variable:

$$\frac{\partial \Pr(Y^{SR} = 1|X)}{\partial x_j} = (1 - p01 - p10)\, f(X\beta)\beta_j \qquad (7)$$

Where x_j is the jth independent variable.

For example if x_j is parental income and self-reported smoking (Y^{SR}) is misreported, then the marginal effect of parental income will be understated by a value 'ω', where ω = 1- p01- p10. Since the misclassification probabilities p01 and p10 are assumed to be constant across observed characteristics (Bound et al., 2001, Hausman et al., 1998), it is expected that the marginal effect of each covariate will be biased by a constant value, ω. In this chapter, the probability that a zero is misclassified as a one (p01) is approximately 0.02 and the probability that a one is misclassified as a zero is approximately 0.51. This implies that the marginal effects

[8]See Note A2.1 in the Appendix for proof.

(and coefficients) of observable characteristics in the self-reported smoking participation model will be underestimated by approximately 53% (p01+p10=0.53).

Self-reported and cotinine-validated smoking participation are modelled separately as a function of a set of observed characteristics in two separate probit models. Following the approach of Blow et al. (2005), only adolescent characteristics (age, gender and ethnicity) are controlled for in a basic model. The basic model is then extended by first including household characteristics and then, other observed parental characteristics which are likely to be correlated with parental income.

2.3 Data Description and Variables

This chapter uses the HSE and pools data from 1997 to 2008. The HSE is a series of annual cross-sectional surveys designed to collect data from adults and children living in a representative sample of households in England. Every year a nationally representative sample of households is drawn from the Postcode Address file and all adults over the age of 16 years and a random selection of two children aged between 0-15 years living within selected households are interviewed.[9] Each year, the survey includes a set of core health and lifestyle topics including smoking, drinking, and general health. Objective measures of health including saliva specimens for cotinine assay were collected in a nurse visit approximately one week after the first interview.

2.3.1 Self-reported smoking participation

Data on smoking participation was collected for all individuals aged 8 and above, using self-completed questionnaires. In this chapter, current smoking status of children aged 11-15 years was obtained from replies to two questions. First, 'Now read all the following sentences carefully and tick the box next to the one which best describes you: (1) I have never smoked; (2) I have only smoked once or twice; (3) I used to smoke sometimes, but I never smoke a cigarette now; (4) I sometimes smoke, but I don't smoke every week; (5) I smoke between one and six cigarettes a week and (6) I smoke more than six cigarettes a week. Second, 'Did you smoke any cigarettes last week?' Those who chose options 4, 5 or 6 in the first questions are

[9] A full description of the survey design can be found in Prescott-Clarke and Primatesta (1998).

classified as current smokers and those who chose options 1, 2, or 3 as non-smokers provided they did not answer 'yes' to the second question. Adolescents who chose options 1, 2 or 3 in the first question and answered 'yes' to the second questions were classified as current smokers[10]. A binary variable of self-reported smoking is defined as one for current smokers and zero for non-smokers.

2.3.2 Objective indicator of smoking participation

In all years except in 2000, saliva specimens were collected for cotinine assay in a nurse visit one week after the self-reported questionnaires were completed. Cotinine assays were performed using gas chromatography which can detect cotinine levels as low as 0.1ng/ml. Cotinine is a metabolite of nicotine and has a half-life of about 16-20 hours. It can be detected in saliva specimens of regular smokers and occasional smokers if smoking occurred a few days prior to the collection of the specimen. Cotinine is generally accepted as a quantitative indicator of tobacco intake with very high specificity (percentage of non-smokers classified as non-smokers) and sensitivity (percentage of smokers classified as smokers) (Benowitz, 1996, Jarvis et al., 2008). Recent studies have shown that a cut-point of 12ng/ml will detect active smoking in children aged 8-15 years with a sensitivity of 95.8% (Jarvis et al., 2008). In this chapter, cotinine-validated smokers are defined as those with saliva cotinine concentrations greater than or equal to 12ng/ml. However, because the cut-point may vary between 8-18ng/ml in non-smokers depending on the extent of exposure to second-hand smoke (Jarvis et al., 2008), a cut-point of 18ng/ml is used in a sensitivity analysis.

2.3.3 Socioeconomic measure and other variables

Equivalised annual household income is used as a proxy for parental socioeconomic status or parental income. From 1997 onwards, the HSE collected data on annual household income using cards displaying 31 income bands ranging from less than £520 to greater than £150,000 per annum. These cards were completed by a household reference person or their partner. Respondents were asked to estimate the total annual household income including their own income, the income of their partners and any other persons living within the household. Equivalised income was calculated using the McClements scoring system (McClements, 1977) to account for the number of persons in the household including children.

[10] Less than 1% of those who chose options 1,2 or 3 in the first question answered 'yes' to the second question.

A wide range of demographic and socioeconomic characteristics available in the HSE are used as control variables. These include demographic characteristics of adolescents: age at the time of the interview (11-15 years), gender (male/female) and ethnicity (white/non-white); household characteristics: home ownership and household location (rural/suburban/urban) Characteristics of parents controlled for include: fathers and mothers highest academic qualification, age, occupation, marital status and current smoking behaviour (number of cigarette packs (20/pack) smoked daily). These were obtained by linking parents' responses in the individual questionnaires to each adolescent. Parents' current smoking status and number of cigarette smoked were obtained from the questions: 'Do you smoke cigarettes at all nowadays?', if yes, 'about how many cigarettes a day do you usually smoke?'.

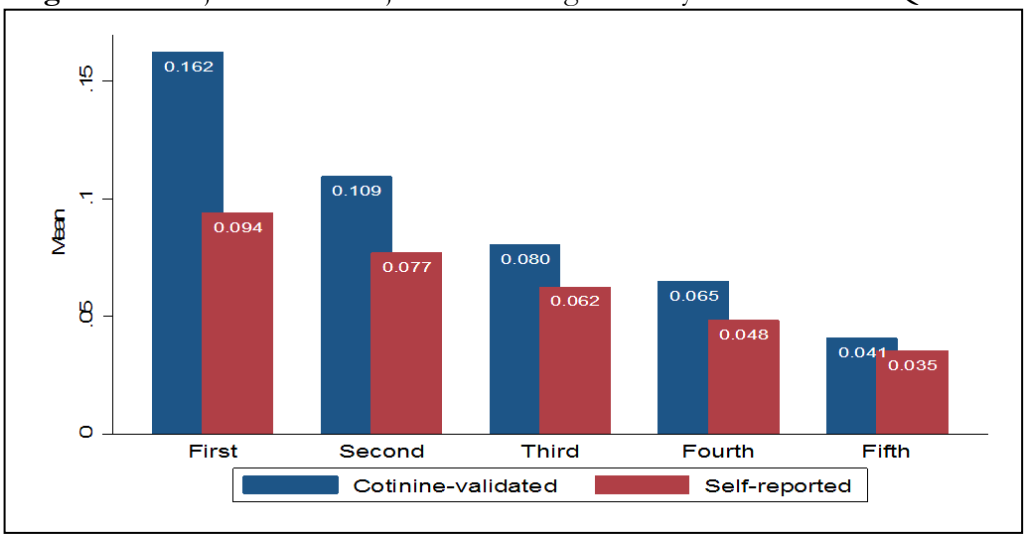
Finally, 'non-smoke-free' households were identified using the response to the question completed by the household reference person: 'Does anyone smoke inside this house/flat on most days?' If yes, 'How many people smoke inside this (house/flat) on most days?' Interviewers were explicitly asked to code as 'no' if smoking by any household member was reported, but occurred outside the home and to code as 'yes' if non-household members smoked within the home. Table A2.1 in the Appendix gives a description of all the variables.

The full sample consists of 7421 observations. This includes all adolescents who had valid cotinine values and complete information on parental income. Data from 1999, 2000 and 2004 were excluded because in 2000, saliva specimens were not collected and in 1999 and 2004, saliva specimens were collected from only adolescents in ethnic minority groups. In 1997 and 2002, the number of children surveyed was boosted by surveying more households. Although household questionnaires were completed by the head of the household, adults (including parents) from the boost sample were not surveyed. Therefore children from the boost sample had no information on parental characteristics. Dummy variables are included to account for missing information on parents' characteristics. Overall, parents' characteristics was missing if the adolescent is from the boost sample or if they live in a single parent household or in a two-parent household but one parent was absent during the interview.

To investigate the association between observed characteristics and the probability of under-reporting smoking, a subsample consisting of only cotinine-validated smokers (N=694) is used. For this subsample, a binary variable, Y^{UR}, is defined as one if a cotinine-validated smoker reports being a non-smoker and zero otherwise.

Table 2.1 shows a descriptive summary of the characteristics of adolescents in the full sample, as well as the descriptive summary of the characteristics of adolescents grouped by income quintiles. In the full sample (Table 2.1, column 1), the proportion of self-reported smokers (6.4%) is approximately 3 percentage points less than the proportion of cotinine-validated smokers (9.3%). Interestingly, the underestimation of the proportion of self-reported smokers in the full sample varies across income quintiles (Figure 2.1). The largest difference between the proportions of cotinine-validated and self-reported smokers is observed in the first income quintile (approximately 7%) while the smallest difference is observed in the highest income quintile. Overall, both self-reported and cotinine-validated smoking follows an income gradient with smoking participation highest amongst the poorest.

Figure 2.1 Subjective and Objective smoking status by Parent Income Quintiles

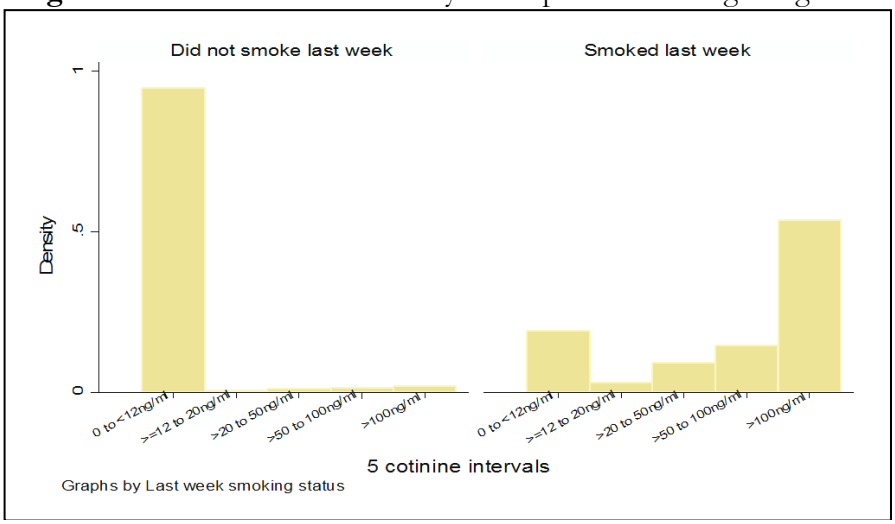


The distribution of cotinine by self-reported smoking behaviour (Figures 2.2a and 2.2b) shows further evidence of discrepancies in self-reported smoking. For example, very high levels of cotinine is detected in the saliva sample of a small proportion of those who reported to have 'not smoked last week' (Figure 2.2a) or in those who reported 'to have smoked sometimes before but never smoke now' (Figure 2.2b). On the other hand, in the proportion of those who reported to 'smoke sometimes, but not every week' (Figure 2.2b) or to have 'smoked last week' (Figure 2.2a), low levels of cotinine not indicative of active smoking is observed. This discrepancy could occur if a non-smoker reports smoking or if an occasional smoker smoked

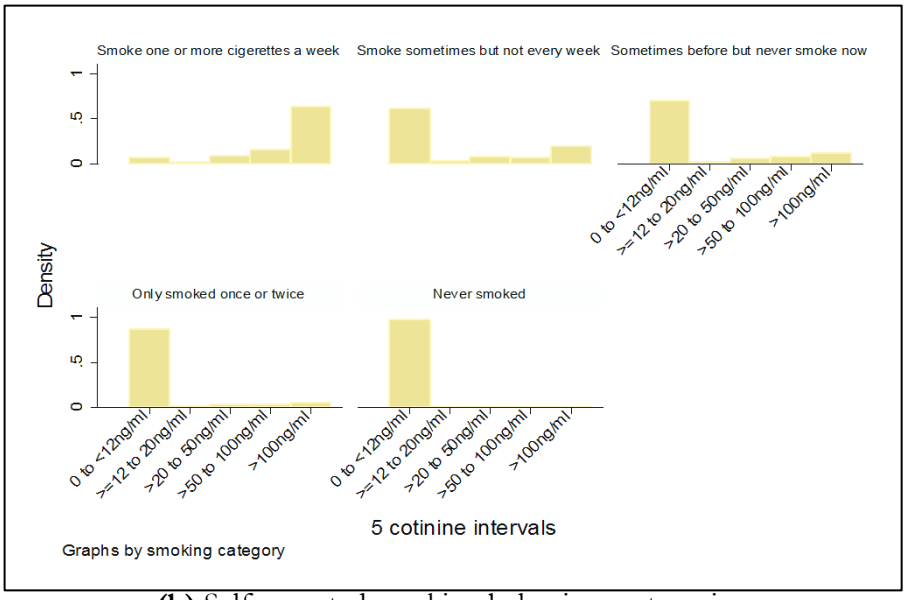
many days before providing a saliva sample. This may result in a failure of the assay to detect cotinine levels indicative of active smoking.

Table 2.1 (column 7) also shows the descriptive statistics of self-reported smokers. Compared to the full sample, self-reported smokers are on average, older and live in poorer households. In addition, a higher proportion of self-reported smokers have parents who smoke and live in households were smoking is permitted within the home.

Figure 2.2 Cotinine distribution by self-reported smoking categories



(a) Self-reported smoking in the previous week



(b) Self-reported smoking behaviour categories

Table 2.1 Descriptive statistics of full sample and cotinine-validated (CV) sub-sample

	FULL SAMPLE							
		Income Quintiles						CV
	All	First	Second	Third	Fourth	Fifth	SR Smokers	Sub-sample
Adolescents characteristics								
Cotinine≥12ng/ml	0.0927	0.162	0.109	0.0805	0.0647	0.0408	0.729	-
Cotinine≥18ng/ml	0.0864	0.148	0.105	0.0794	0.0598	0.0351	0.711	-
SR smoking	0.0640	0.0939	0.0770	0.0624	0.0483	0.0354	-	0.503
Age	12.94	12.88	12.96	12.97	12.95	12.97	14.244	14.07
Age squared	169.6	167.8	170.0	170.1	169.8	170.3	203.799	199.1
Male	0.502	0.510	0.493	0.511	0.507	0.489	0.444	0.467
White	0.902	0.814	0.890	0.931	0.947	0.933	0.928	0.933
Household characteristics								
Income (£)	20386.4	5603.9	10566.4	15947.3	23650.0	47754.4	16013.128	14903.9
Log Income	9.634	8.568	9.256	9.671	10.06	10.71	9.398	9.324
Own a house	0.703	0.329	0.597	0.767	0.910	0.945	0.536	0.523
Urban	0.298	0.312	0.305	0.289	0.275	0.307	0.278	0.269
Suburb	0.490	0.538	0.501	0.505	0.488	0.418	0.502	0.539
Rural	0.212	0.150	0.194	0.207	0.238	0.276	0.220	0.192
Number of persons smoking within home								
No one	0.657	0.453	0.555	0.677	0.770	0.847	0.435	0.380
1 smoker	0.215	0.358	0.270	0.201	0.134	0.0999	0.286	0.351
2 smokers	0.107	0.145	0.151	0.107	0.0857	0.0425	0.173	0.177
≥3 smokers	0.0208	0.0440	0.0235	0.0138	0.0109	0.0104	0.106	0.0919
Fathers (F) and Mothers (M) characteristics								
F Degree	0.106	0.0157	0.0294	0.0687	0.136	0.290	0.067	0.0264
F Below degree	0.0790	0.0269	0.0605	0.0929	0.116	0.103	0.049	0.0651
F NVQ3/A levels	0.0649	0.0259	0.0547	0.0612	0.106	0.0800	0.070	0.0294
F NVQ2/Olevels	0.109	0.0650	0.110	0.158	0.140	0.0757	0.211	0.0875
F NVQ1/CSE	0.0314	0.0326	0.0310	0.0457	0.0295	0.0177	0.075	0.0305
F No qualification	0.0848	0.121	0.132	0.0845	0.0517	0.0297	0.159	0.115
M Degree	0.104	0.0206	0.0317	0.0642	0.131	0.283	0.034	0.0475
M Below degree	0.0743	0.0327	0.0500	0.0732	0.106	0.114	0.063	0.0357
M NVQ3/A levels	0.0860	0.0475	0.0722	0.0974	0.116	0.100	0.030	0.0674
M NVQ2/Olevels	0.230	0.181	0.264	0.291	0.256	0.159	0.101	0.233
M NVQ1/CSE	0.0579	0.0767	0.0811	0.0585	0.0411	0.0294	0.021	0.0777
M No qualification	0.113	0.224	0.172	0.0814	0.0462	0.0287	0.100	0.174
F Employed	0.417	0.133	0.337	0.491	0.565	0.583	0.277	0.275
F Unemployed	0.0390	0.0958	0.0558	0.0155	0.0107	0.0128	0.041	0.0584
F Sick	0.0189	0.0581	0.0252	0.00380	0.00439	0	0.035	0.0223
M Employed	0.465	0.179	0.432	0.534	0.609	0.591	0.398	0.392
M Unemployed	0.172	0.336	0.204	0.115	0.0728	0.120	0.201	0.197
M Sick	0.0282	0.0674	0.0363	0.0162	0.0148	0.00316	0.033	0.0468
P non smoker	0.732	0.631	0.667	0.730	0.779	0.863	0.619	0.578
P ≤1 cig pack/day	0.179	0.246	0.209	0.185	0.150	0.0948	0.234	0.270
P >1 cig pack/day	0.0822	0.106	0.118	0.0817	0.0631	0.0377	0.136	0.147
P married	0.513	0.302	0.444	0.543	0.624	0.667	0.378	0.376
P single	0.175	0.320	0.251	0.143	0.0883	0.0587	0.298	0.304

Table 2.1 (continued) Descriptive statistics

	FULL SAMPLE							
		Income Quintiles						CV
	All	First	Second	Third	Fourth	Fifth	SR Smokers	Sub-sample
F age ≤35 years	0.042	0.042	0.053	0.051	0.046	0.017	0.028	0.042
F age 36-45 years	0.272	0.162	0.243	0.305	0.345	0.314	0.222	0.214
F age≥46 years	0.160	0.082	0.121	0.152	0.189	0.264	0.102	0.098
M age ≤35 years	0.123	0.195	0.158	0.127	0.086	0.041	0.133	0.130
M age 36-45 years	0.421	0.319	0.424	0.426	0.479	0.464	0.394	0.400
M age≥46 years	0.122	0.069	0.089	0.115	0.132	0.210	0.104	0.106
Survey years								
1997-2001	0.389	0.511	0.431	0.388	0.341	0.263	0.455	0.443
2002-2005	0.420	0.376	0.398	0.421	0.461	0.451	0.419	0.421
2006-2008	0.190	0.113	0.171	0.191	0.198	0.286	0.126	0.136
Parents missing variable indicator								
Single mum	0.150	0.283	0.225	0.120	0.0640	0.0451	0.250	0.258
Single dad	0.0156	0.0189	0.0160	0.0175	0.0129	0.0123	0.035	0.0316
F not home	0.0657	0.0418	0.0564	0.0642	0.0724	0.0961	0.069	0.0575
M not home	0.0097	0.0089	0.0102	0.0101	0.00706	0.0100	0.005	0.00266
F&M missing	0.310	0.390	0.302	0.308	0.284	0.264	0.329	0.330
N	*7421*	*1490*	*1488*	*1474*	*1484*	*1485*	475	*694*

CV=cotinine-validated; SR=self reported; P=parent

2.4 Results and Discussion

2.4.1 Implications of misreporting smoking participation

Results of the cotinine-validated and self-reported smoking participation models are presented in Tables 2.2a and 2.2b. The results (showing marginal effects) are interpreted as estimates of the association between observable characteristics and the probability of smoking and do not reflect causality between observable characteristics and adolescent smoking behaviour. After controlling for adolescents' characteristics alone, a negative association is observed between parental income and the probability of smoking in both the cotinine-validated and self-reported smoking participation models (columns 1-4, Table 2.2a). An increase in parental income is associated with a decrease in the probability of being a smoker. This income effect is statistically significant in both models, but the magnitude of this effect is approximately 50% less in the self-reported smoking model compared to the cotinine-validated smoking model. A 1% increase in parental income is associated with a decrease in the probability of being a

cotinine-validated smoker by approximately 4 percentage points, but by approximately 2 percentage points in the self-reported smoking model.

Similar trends are observed with other covariates. For example, age is positively correlated with the probability of being a smoker in both models. However, while an increase in age is associated with an increase in the probability of being a cotinine-validated smoker by approximately 15 percentage points, the probability of being a self-reported smoker increases by approximately 7 percentage points. In the cotinine-validated smoking model, the probability of being a smoker is approximately 3 percentage points higher in white adolescents compared to adolescents from ethnic minority groups but by approximately 2 percentage points in the self-reported smoking model.

When the model is extended to include household characteristics, a reduction in the estimated magnitude of the association between cotinine-validated or self-reported smoking and parental income is observed (columns 5-8, Table 2.2a). After controlling for home ownership, area of residence and the number of persons who smoke within the home, the correlation between parental income and the probability of being a smoker remains negative and statistically significant in both models but decreases from 4 percentage points to approximately 2 percentage points in the cotinine-validated smoking model and from 2 percentage points to approximately 1 percentage point in the self-reported smoking model. In both models, the probability of being a cotinine-validated or a self-reported smoker is significantly higher in adolescents living in households where smoking is permitted within the home compared to smoke-free homes. However, this effect is higher in the cotinine-validated smoking model compared to the self-reported smoking model.

After controlling for the full set of observed parent's characteristics, the correlation between parental income and cotinine-validated smoking remains negative and significant (Table 2.2b). On the other hand, the income effect in the self-reported smoking model becomes insignificant (Table 2.2b). This result is similar to those reported by Blow et al. (2005) using the BHPS where the observed association between parental income and self-reported smoking disappear after controlling for parental education and smoking within the home.

Table 2.2a Probit model of cotinine-validated and self-reported (SR) smoking

	SR smoker		Cotinine ≥12ng/ml		SR smoker		Cotinine ≥12ng/ml	
Log Income	-0.0198**	(0.003)	-0.0424**	(0.004)	-0.00543*	(0.003)	-0.0174**	(0.004)
Age	0.0701+	(0.040)	0.147**	(0.048)	0.0679+	(0.038)	0.137**	(0.044)
Age squared	-0.00143	(0.002)	-0.00388*	(0.002)	-0.00143	(0.001)	-0.00359*	(0.002)
Male	-0.00807*	(0.004)	-0.0101+	(0.005)	-0.00803*	(0.003)	-0.0102*	(0.005)
White	0.0162**	(0.006)	0.0345**	(0.007)	0.0108+	(0.006)	0.0242**	(0.007)
2002-2005	-0.00173	(0.005)	0.00292	(0.007)	-0.00371	(0.004)	0.000686	(0.006)
2006-2008	-0.0167**	(0.005)	-0.0179*	(0.007)	-0.0172**	(0.005)	-0.0126+	(0.008)
Own a house					-0.0253**	(0.006)	-0.0338**	(0.007)
Suburb					-0.00290	(0.005)	0.00599	(0.007)
Rural					0.000745	(0.006)	0.000116	(0.008)
1 smoker					0.0204**	(0.006)	0.0599**	(0.009)
2 smokers					0.0337**	(0.008)	0.0605**	(0.012)
≥3 smokers					0.150**	(0.031)	0.211**	(0.038)
N	7421		7421		7421		7421	

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses.

Marginal effects estimated at sample mean. Standard errors adjusted for clustering at primary sample unit level. Omitted categories: urban and no one smokes within home for household characteristics; and 1997-2001 for survey year.

Table 2.2b Extended probit models of CV and SR smoking (full model)

Variables	SR smoker		Cotinine ≥12ng/ml	
Log Income	-0.00133	(0.003)	-0.0101**	(0.004)
Age	0.0680+	(0.035)	0.135**	(0.041)
Age squared	-0.00150	(0.001)	-0.00361*	(0.002)
Male	-0.00702*	(0.003)	-0.00911*	(0.005)
White	0.0112*	(0.005)	0.0218**	(0.007)
2002-2005	-0.00569	(0.004)	-0.000463	(0.006)
2006-2008	-0.0184**	(0.005)	-0.0117	(0.007)
Own a house	-0.0163**	(0.005)	-0.0220**	(0.007)
Suburb	-0.00311	(0.005)	0.00542	(0.006)
Rural	0.00308	(0.006)	0.00420	(0.008)
1 smoker	0.0179**	(0.006)	0.0480**	(0.010)
2 smokers	0.0385**	(0.010)	0.0611**	(0.014)
≥3 smokers	0.164**	(0.034)	0.222**	(0.042)
P ≤ 1 cig pack/day	-0.00285	(0.006)	0.00320	(0.009)
P >1 cig pack/day	-0.00524	(0.007)	-0.00210	(0.010)
F Below degree	0.0310*	(0.015)	0.0531*	(0.021)
F NVQ3/A levels	0.00465	(0.011)	0.00685	(0.016)
F NVQ2/O levels	0.0359*	(0.015)	0.0326*	(0.016)
F NVQ1/CSE	0.0118	(0.016)	0.0347	(0.023)
F No qualification	0.0211+	(0.013)	0.0464*	(0.019)
M Below degree	-0.00312	(0.006)	-0.00491	(0.006)
M NVQ3/A levels	-0.00258	(0.006)	0.00310	(0.007)
M NVQ2/O levels	-0.00507	(0.005)	0.00389	(0.006)
M NVQ1/CSE	0.000182	(0.007)	0.00901	(0.009)
M No qualification	-0.00229	(0.006)	0.00531	(0.006)
F Unemployed	-0.00862	(0.010)	0.00450	(0.012)
F Sick	0.0142	(0.012)	-0.0155	(0.016)
M Unemployed	0.00126	(0.005)	-0.00374	(0.007)
M Sick	-0.00103	(0.012)	0.0115	(0.014)
F age 36-45 years	0.000140	(0.012)	-0.0184	(0.013)
F age≥46 years	-0.00645	(0.013)	-0.0295*	(0.014)
M age 36-45 years	-0.00783	(0.007)	0.00200	(0.009)
M age≥46 years	-0.00915	(0.008)	0.00492	(0.011)
Parent single	0.0192+	(0.011)	0.0418**	(0.016)
	7421		7421	

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses.

Marginal effects estimated at sample mean. Standard errors adjusted for clustering at primary sample unit level. Omitted categories: urban and no one smokes within home for household characteristics; 1997-2001 for survey year; degree or equivalent for parents' qualification; employed for parents' employment status; non smoker for parents' smoking status; F/M age ≤35 years for parents' age; and married for parents' marital status.

2.4.2 Robustness checks

The empirical analysis outlined in the sub-section 2.4.1 relies on saliva cotinine concentration to identify 'true' smoking participation amongst adolescents. While cotinine assays have been shown to classify smokers and non-smokers with relatively high sensitivity and specificity, very high exposure to second-hand smoke may result in high levels of saliva cotinine in non-smokers (Jarvis et al., 2008). Therefore, some adolescents who do not smoke may present with saliva cotinine concentrations indicative of active smoking depending on the extent of exposure to second-hand smoke. In this sub-section, the robustness of the parental income effect in the cotinine-validated smoking model is further tested by accounting for possible mechanisms that may confound the correlation between parental income and cotinine-validated smoking.

First, a higher cut-point of 18ng/ml is used to define cotinine-validated smokers. In adolescents, high exposure to second-hand smoke can result in saliva cotinine concentrations as high as 18ng/ml in non-smokers (Jarvis et al., 2008). Therefore a cut-point of 12ng/ml may misclassify non-smokers as smokers if they are exposed to high levels of second-hand smoke. The results are robust to the cut-point for defining cotinine-validated smokers (Table 2.3, column 1). Using a cut-point of 18ng/ml, an increase in parental income is associated with a statistically significant lower probability of being a cotinine-validated smoker[11]. The magnitude of this effect is similar to the income effect observed with a lower cut-point of 12ng/ml. This is unsurprising given that the proportion of adolescents classified as cotinine-validated smokers using a cut-point of 12ng/ml and 18ng/ml do not differ markedly (9.26 and 8.63 percent respectively; Table 2.1, column 1). This implies that less than 1% of adolescents classified as smokers by a cut-point of 12ng/ml are classified as non-smokers by a cut-point of 18ng/ml.

Second, since children from lower income households are more likely to live with parents who smoke or in non-smoke-free homes compared to children from higher income households, the parental income effect may therefore, be due to higher exposure to second-hand smoke in lower income households. To control for the possibility of higher exposure in lower income households, parental income is interacted with the number of cigarettes smoked by parents and with the number of persons who smoke within the home. These interaction terms are used as controls in the cotinine-validated smoking model (using a cut-point of 12ng/ml).

[11] Here only results of the marginal effect on parental income are presented. However, the marginal effects of all covariates obtained using a cut-point of 18ng/ml are similar to those obtained using a cut-point of 12ng/ml.

Table 2.3 (columns 3 and 4) presents the marginal effect of the interaction terms estimated at each category of the number of cigarettes smoked by parents (column 3) and the number of smokers within the home (column 4). The marginal effect of the interaction terms can be interpreted as the average change in the conditional probability of being a cotinine-validated smoker when parental income increases by 1 percentage point in each category[12]. For example, the marginal effect of the interaction term, *P non smoker*Log Income*, captures the income effect in adolescents whose parents' are non-smokers (Table 2.3, column 3). In other words, for adolescents whose parents are non-smokers, an increase in parental income is associated with a lower probability of being a cotinine-validated smoker (Table 2.3, column 3). Similarly, the marginal effect of the interaction term, *No one*Log Income*, can be interpreted as the income effect in adolescents living in households where there are no smokers (Table 2.3, column 4), i.e. in households where there are no smokers, an increase in parental income is associated with a lower probability of been a cotinine-validated smoker. This parental income effect observed in smoke-free homes suggests that the effect of parental income in the cotinine-validated smoking model is not explained by less exposure to second-hand smoke (at least, exposure within the home) amongst adolescents living in higher income households.

Interestingly, the marginal effects of the interaction terms show some evidence of a differential effect of parental income in smoke-free and non-smoke-free households. The parental income effect appears to be higher in non-smoke-free households compared to the income effect in smoke-free households. For example, in adolescents whose parents smoke more than one pack of cigarette, the income effect is higher in comparison to the income effect in adolescents whose parents are non-smokers (2 vs. 1 percentage point, Table 2.3, column 3). Similarly the effect of parental income appears to be higher in adolescents living in households with three or more smokers compared to adolescents living in smoke-free homes (a difference of approximately 7 percentage points, Table 2.3, column 4). One possible explanation for these differences could be that higher income parents who smoke may perhaps adopt other compensatory behaviours that discourage smoking participation in their children[13].

[12] The marginal effects for the interaction term were estimated and interpreted following Karaca-Mandic et.al (2012).

[13] This idea is explored further in chapter 3.

Table 2.3 Robustness checks: Cotinine-validated smoking probit models

Variables	(1)		(2)		(3)		(4)	
Log Income	-0.0105**	(0.003)	-0.0100**	(0.004)	-0.0101**	(0.004)	-0.00966*	(0.004)
Spring‡			-0.00572	(0.007)				
Autumn‡			-0.00582	(0.007)				
Winter‡			-0.00671	(0.007)				
Parents(P) smoking								
P non smoker*Log Income					-0.0132*	(0.006)		
P ≤ 1 cig pack/day*Log Income					-0.0111	(0.010)		
P >1 cig pack/day*Log Income					-0.0245+	(0.015)		
P non smoker (base category)					-	-		
P ≤ 1 cig pack/day					0.00377	(0.009)		
P >1 cig pack/day					-0.00523	(0.010)		
Number of smokers within home								
No one*Log Income							-0.0104*	(0.005)
1 smoker*Log Income							-0.0214+	(0.012)
2 smokers*Log Income							-0.00644	(0.018)
≥3 smokers*Log Income							-0.0837*	(0.041)
No one (base category)							-	-
1 smoker							0.0466**	(0.010)
2 smokers							0.0649**	(0.015)
≥3 smokers							0.192**	(0.047)
	7421		7421		7421		7421	

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses. ‡Base category: Summer. All models include the full set of controls for child, household and parents' characteristics

Finally, during cold months, adolescents are more likely to spend more time indoors and exposure to second-hand smoke may be higher if parents change their smoking pattern to adapt to cold weathers by smoking inside the home. After controlling for potential seasonal variations in exposure to second-hand smoke using the month the adolescent was interviewed, the results of the cotinine-validated smoking model remain largely unchanged (Table 2.3, column 2).

2.4.3 Determinants of under-reporting smoking participation

Several factors may explain why adolescents misreport their smoking behaviour (Dolcini et al., 1996). Fear of disclosure and the perceived social undesirability of smoking may result in adolescents under-reporting their true smoking behaviour. Being interviewed by the same interviewer who interviewed a parent is significantly correlated with under-reporting of smoking participation in adolescents (Griesler et al., 2008, Kandel et al., 2006). The prevalence of smoking has been observed to be higher when anonymous questionnaires are administered in comparison to named questionnaires (Adams et al., 2008). Exit questionnaires completed by adolescents participating in an intervention study showed that approximately 4% of adolescents admit to have consistently misreported their smoking status during the intervention study (Stein et al., 2002). Of these, approximately 25% were worried the information they give will be relayed back to their parents (Stein et al., 2002). In other studies on substance-abuse, adolescents admit to denying the use of substances to physicians when their parents are present during consultation sessions (Friedman et al., 1990).

Parental smoking or living with other smokers may be associated with the misreporting of smoking behaviour. For example, in the 2008 Smoking, Drinking and Drug Use Survey, adolescents who smoked openly (i.e. those whose families were aware of their smoking behaviour) were less likely to perceive that their families would disapproved of their smoking in comparison to secret smokers (Fuller, 2009). In households where no other household member smoked, approximately 40% of adolescent smokers were open smokers while 60% were secret smokers. On the other hand, in households where three or more household members were smokers, as high as 71% were open smokers (Fuller, 2009). This suggests that adolescents living in households where other members smoke may be less likely to under-report their smoking. In adults, smoking has been shown to follow a socioeconomic gradient, with smoking concentrated more in adults from lower socioeconomic backgrounds. Taken together, it can be expected that adolescents from

higher socioeconomic backgrounds may have a greater motivation not to disclose their true smoking behaviour due to the perceived undesirability of their behaviour.

The setting in which data on smoking is collected is also important. The HSE consistently produces lower estimates of the prevalence of adolescent smoking in comparison to the Smoking Drinking and Drugs Use Survey, a school-based survey (Craig and Mindell, 2008). In the United States, a study comparing self-reported smoking status in a sample of adolescents within a school and household setting showed that approximately 9-20% of the same adolescents who had reported smoking in the last 12 months in the school-based survey reported never to have tried smoking in the household counterpart (Griesler et al., 2008).

Due to the sporadic and experimental nature of adolescent smoking, poor recall may result in an unintentional misreporting of smoking behaviour. For example, in the 2008 Smoking Drinking and Drugs Use Survey, 65% of those who reported usually smoking up to six cigarettes a week in the interviews had diary records of smoking more than seven cigarettes in the previous week. Similarly, 46% of those who reported smoking sometimes but less than once a week, had diary entries suggesting they had smoked at least one cigarette in the previous week (Fuller, 2009).

In this sub-section the association between observable characteristics and the probability of under-reporting smoking participation is further investigated using a subsample consisting of only cotinine-validated smokers (N=694). A binary variable Y^{UR}, was generated which equals one if a cotinine-validated smoker reports being a non-smoker and zero otherwise. The probability that Y^{UR} equals one conditional on a set of covariates, Pr (Y^{UR} =1|X) is estimated, using maximum likelihood estimation approach. Adolescents who, despite reporting being smokers, had cotinine levels less than 12ng/ml are not considered. This form of discrepancy (over-reporting) is less common and in this study sample the probability of over-reporting smoking is significantly lower (0.02) than the probability of under-reporting smoking (0.51).

Table 2.1 (last column) shows a summary of the characteristics of the sub-sample of cotinine-validated smokers. On average, in comparison to the full sample, the sub-sample consists of adolescents from a lower socioeconomic family background. For example, average household income is lower in the sub- sample compared to the full sample (£15,000 vs. £20,000 per annum). In addition, a lower proportion of cotinine-validated smokers (38% versus 85% in the full sample) live in smoke-free homes. Overall,

approximately 50% of the cotinine-validated smokers correctly reported their smoking behaviour (Table 2.1, last column) but no clear pattern is observed in the probability to under-report smoking participation across income quintiles (Figure 2.3). However, compared to those in lower income quintiles, adolescents in the highest income quintile have the highest probability of under-reporting smoking participation (Figure 2.3).

Figure 2.3 Reporting behaviour by Parent Income Quintiles (CV sub-sample only)

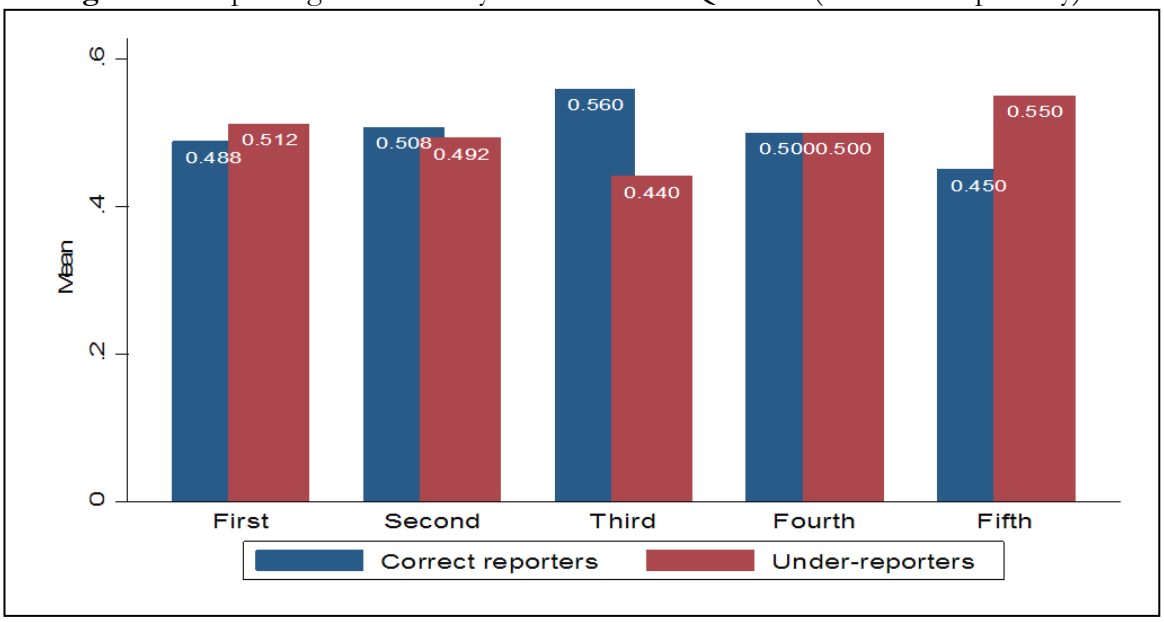


Table 2.4 shows the results (marginal effects) of the probit model of under-reporting smoking participation. In the basic model controlling for only adolescent characteristics, the association between parental income and under-reporting is positive, meaning that as parental income increases, the probability of under-reporting smoking participation increases. Although this effect is statistically insignificant, it is consistent with the notion that adolescents living in higher income households are more likely to under-report their smoking behaviour. When the model is extended to include household characteristics, the association between home ownership (another measure of socioeconomic status) and the probability of under-reporting smoking participation is positive and statistically significant. In addition, compared to adolescents living in smoke-free homes, adolescents living in households where three or more persons smoke within the home are statistically significantly less likely to under-report their smoking behaviour (column 5, Table 2.4).

Table 2.4 Probit model of under-reporting smoking behaviour

Variables	(1)		(2)		(3)	
Log Income	0.0385	(0.029)	-0.0200	(0.035)	-0.0368	(0.037)
Age	-0.601	(0.493)	-0.654	(0.507)	-0.751	(0.517)
Age squared	0.0165	(0.018)	0.0184	(0.019)	0.0215	(0.019)
Male	0.0475	(0.040)	0.0521	(0.040)	0.0418	(0.042)
White	0.0558	(0.077)	0.0561	(0.079)	0.0314	(0.082)
2002-2005	-0.0379	(0.046)	-0.0254	(0.047)	-0.00986	(0.050)
2006-2008	0.0659	(0.064)	0.115	(0.075)	0.166*	(0.081)
Own a house			0.116*	(0.052)	0.0778	(0.055)
Suburb			0.0667	(0.058)	0.0860	(0.060)
Rural			0.0380	(0.073)	0.0369	(0.075)
1 smoker			-0.00462	(0.051)	-0.0420	(0.060)
2 smokers			-0.0837	(0.064)	-0.136+	(0.074)
≥3 smokers			-0.293**	(0.071)	-0.356**	(0.072)
P ≤ 1 cig pack/day					0.0692	(0.073)
P >1 cig pack/day					0.0339	(0.094)
F Below degree					0.0224	(0.167)
F NVQ3/A levels					0.0845	(0.190)
F NVQ2/O levels					0.0971	(0.160)
F NVQ1/CSE					0.161	(0.188)
F No qualification					0.207	(0.161)
M Below degree					-0.0618	(0.134)
M NVQ3/A levels					0.0173	(0.119)
M NVQ2/O levels					0.103	(0.091)
M NVQ1/CSE					-0.149	(0.111)
M No qualification					0.0582	(0.103)
F Unemployed					0.0350	(0.119)
F Sick					-0.235	(0.159)
M Unemployed					-0.0544	(0.068)
M Sick					0.132	(0.125)
F age 36-45 years					-0.198	(0.137)
F age≥46 years					-0.0931	(0.149)
M age 36-45 years					0.230**	(0.072)
M age≥46 years					0.246*	(0.099)
Parent single					0.00216	(0.072)
N	694		694		694	

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses

Marginal effects estimated at sample mean. Standard errors adjusted for clustering at primary sample unit level. Omitted categories: urban and no one smokes within home for household characteristics; 1997-2001 for survey year; degree or equivalent for parents' qualification; employed for parents' employment status; non smoker for parents' smoking status; F/M age ≤35 years for parents' age; and married for parents' marital status.

With respect to adolescents' characteristics, no statistically significant effects of age, ethnicity and gender is observed. Nevertheless, the direction of the association between age and reporting behaviour is worth noting. After controlling for the full set of covariates, older adolescents are less likely to under-report smoking participation. This is consistent with findings from recent epidemiological studies which have reported negative

associations between age and the probability to under-report smoking in adolescents (Griesler et al., 2008, Kandel et al., 2006).

2.4.4 Misreporting smoking participation and income-related inequality in smoking

In addition to producing biased estimates of the effects of observable characteristics on the probability of smoking participation, misclassification errors are likely to affect measures of income-related inequality in smoking if there are systemic differences in misreporting smoking amongst adolescents from different socioeconomic backgrounds. Figure 2.1 shows that the underestimation of average rates of smoking participation varies across parental income quintiles. The difference between cotinine-validated and self-reported smoking participation is largest at the lowest parental income quintile and smallest at the highest parental income quintile.

The concentration index has been widely used in measuring income-related health inequality and depends on the relationship between the variable of interest and the rank position in the income distribution (Wagstaff et al., 1991, Kakwani et al., 1997). When the variable of interest is unbounded and measured on a ratio scale, the value of the concentration index ranges from -1 (in which case the variable of interest is disproportionately concentrated amongst the poor) and +1 (in which case the variable of interest is disproportionately concentrated amongst the rich). The concentration index, *C*, can be defined as:

$$C = \frac{2}{\bar{y}}\mathrm{cov}(Y_i, R_i) \qquad (10)$$

where Y_i is the variable of interest (in this study, the smoking status for the *i*th individual), R_i is the fractional rank of the *i*th individual in the income distribution and $\bar{y}$ is the mean of Y across all individuals.

However, when the variable of interest is binary, such as in smoking participation, the possible range of values the concentration index can take will depend on $\bar{y}$ such that as $\bar{y}$ increases, the range the concentration index takes becomes narrower (Wagstaff, 2005, Erreygers, 2009). This poses a problem when comparing the degree of inequality between populations with different means. Therefore, to ensure comparability of the concentration indices estimated using cotinine-validated smoking and self-reported smoking, the corrected concentration index is applied (Erreygers, 2009). The corrected concentration index, *CCI* is defined as:

$$CCI = \frac{4\bar{y}}{z - w} C \qquad (11)$$

where *C* is the concentration index estimated from equation (10), z and w are the upper (one) and lower (zero) bounds of *Y*, respectively.

Table 2.5 Corrected concentration index for smoking participation

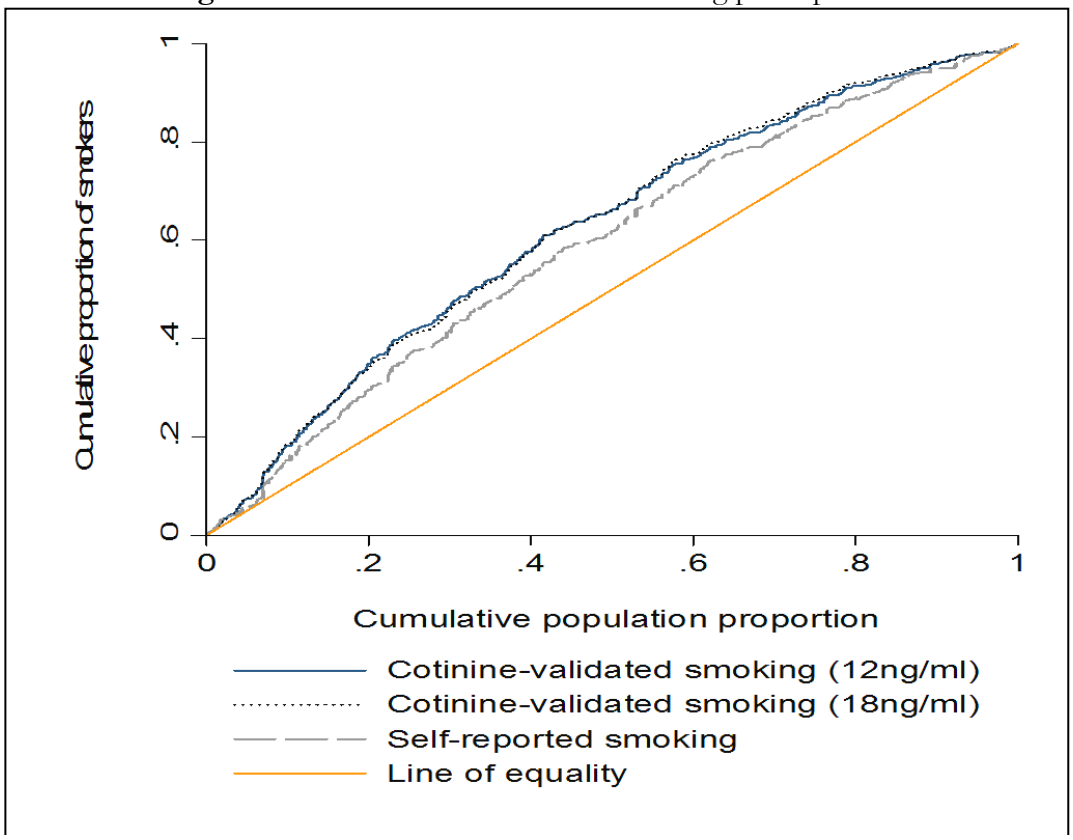
Year	SR smoking	Cotinine>12ng/ml	Cotinine>18ng/ml	N
1997-2001	-0.0316**	-0.0837**	-0.0785**	3149
2002-2005	-0.0505**	-0.0890**	-0.0865**	2783
2006-2008	-0.0487**	-0.0867**	-0.0791**	1489
All years	-0.0468**	-0.0904**	-0.0855**	7421

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Estimates of the corrected concentration indices based on cotinine-validated and self-reported smoking participation are presented in Table 2.5. The corrected concentration indices for both cotinine-validated and self-reported smoking are negative, meaning that smoking is concentrated disproportionately in the poor. However, parental income-related inequality in self-reported smoking is less than inequality in cotinine-validated smoking by approximately 50%. This is unsurprising given that on average, self-reported smoking participation is underestimated to a larger extent in adolescents from lower income households compared to adolescents from higher income households (Figure 2.1).

Figure 2.4 shows a graphical representation (concentration curves) of income-related inequality in self-reported and cotinine-validated smoking. The concentration curve plots the cumulative proportion of smoking by the cumulative proportion of adolescents ranked from the poorest to the richest. The concentration curves for both cotinine-validated and self-reported smoking lie above the diagonal line of equality, meaning that smoking is disproportionately concentrated in the poor. However, the concentration curve of cotinine-validated smoking lies everywhere above the concentration curve of self-reported smoking, suggesting that income-related inequality in cotinine-validated smoking is greater. This may have significant implications when comparing inequality in smoking between populations or within a population across different time periods. If the extent or pattern of misreporting of smoking participation varies significantly across different populations or time periods, incorrect conclusions may be reached on the extent of or trends in income-related inequalities in adolescent smoking.

Figure 2.4 Concentration curves of smoking participation



2.5 Conclusion

This chapter demonstrates the implications of misclassification errors in empirical studies on adolescent smoking participation. Using saliva cotinine concentration to define a 'true' model of adolescent smoking participation, the results show that misclassification errors in self-reported smoking results in biased estimates of the impact of observed characteristics on the probability of smoking participation. The marginal effects of parental income (and other characteristics) are biased in the presence of misclassification errors. The observed parental income effect is robust to different specifications of the cotinine-validated smoking model that account for potential variations in exposure to second-hand smoke. In addition, this chapter highlights the implications of misclassification errors in measuring income-related inequality in self-reported smoking, when misclassification errors vary across income quantiles. From a policy maker's perspective, this represents an important problem when comparing inequality in smoking between populations or across time periods.

Although cotinine assays are likely to produce more reliable data on adolescent smoking participation, collecting such data is not without its problems. It is expensive, more difficult

to obtain and may result in large proportions of missing data if participants are unable to provide sufficient saliva samples. Self-reported smoking questionnaires, on the other hand, are easier and cheaper to administer and represent an important method of collecting data on smoking behaviour in population surveys. Therefore there is the need to develop suitable methods or to adapt existing methods to account for misclassification errors in empirical studies on smoking participation in adolescents. This will ensure that the impact of observable characteristics including anti-tobacco smoking policies is consistently estimated using self-reported smoking data.

2.6 Appendix

Note A2.1 Misclassification errors in self-reported smoking

$$\Pr(Y^{SR}=1|X)=\Pr(Y^{SR}=1,Y^{CV}=1|X)+\Pr(Y^{SR}=1,Y^{CV}=0|X)$$

$$=\Pr(Y^{SR}=1|\,Y^{CV}=1,X)*\Pr\,(Y^{CV}=1|X)+\Pr(Y^{SR}=1|\,Y^{CV}=0,X)*\Pr\,(Y^{CV}=0|X)$$

$$=\Pr(Y^{SR}=1|\,Y^{CV}=1)*\Pr\,(Y^{CV}=1|X)+\Pr(Y^{SR}=1|\,Y^{CV}=0)*\Pr\,(Y^{CV}=0|X)$$

$$=(1-\ p10)\Pr(Y^{CV}=1|X)+p01\,(1-\Pr(Y^{CV}=1|X))$$

$$[where\ p01=\Pr(Y^{SR}=1|Y^{\mathrm{CV}}=0)\quad and\quad p10=\ \Pr(Y^{SR}=0|Y^{\mathrm{CV}}=1)]$$

$$=p01+(1-p01-p10)\Pr(Y^{CV}=1|X)$$

Table A2.1 Description of Variables

Variable name	Variable Label
Cotinine≥12ng/ml	1 if cotinine ≥12ng/ml
Cotinine≥18ng/ml	1 if cotinine ≥18ng/ml
SR Smoking	1 if a self-reported smoker
Under-reporting	1 if cotinine≥12ng/ml and SR non-smoker
Age	Age of adolescent at last birthday
Age squared	Age squared
Male	1 if male
White	1 if white
Household characteristics	
Income	Annual total household income
Log Income	Log household income
Own a house	1 if owns outright/mortgage/shared ownership part rent/mortgage
Urban (Base group)	1 if lives in inner city/other dense urban or city centre
Suburb	1 if lives in a suburb residential (city/large town outskirts)
Rural	1 if lives in rural residential/village centre/isolated dwelling
Number of persons smoking within home	
No one (Base group)	1 if no one smokes inside the house/flat on most days
1 smoker	1 if one person smokes inside the house/flat on most days
2 smokers	1 if two persons smokes inside the house/flat on most days
≥3 smokers	1 if three or more persons smokes inside the house/flat on most days
Fathers (F) highest qualification	
F Degree (Base group)	1 if NVQ4/NVQ5/Degree or equivalent
F Below degree	1 if higher education below degree
F NVQ3/A levels	1 if NVQ3/GCE A level equivalent
F NVQ2/O levels	1 if NVQ2/GCE O level or equivalent
F NVQ1/CSE	1 if NVQ1/CSE/other grade equivalent/foreign qualification
F No qualification	1 if no qualification
Mothers (M) highest qualification	
M Degree (Base group)	1 if NVQ4/NVQ5/Degree or equivalent
M Below degree	1 if higher education but below degree
M NVQ3/A levels	1 if NVQ3/GCE A level equivalent
M NVQ2/O levels	1 if NVQ2/GCE O level or equivalent
M NVQ1/CSE	1 if NVQ1/CSE/other grade equivalent/foreign qualification
M No qualification	1 if no qualification
Father's employment status	
F Employed (Base group)	1 if in paid employment or self employed
F Unemployed	1 if unemployed/looks after home/retired/full-time education
F Sick	1 if unable to work due to long-term sickness/disability

Table A2.1 (Continued) Description of Variables

Variable name	Variable Label
Mother's employment status	
M Employed (Base group)	1 if in paid employment or self employed
M Unemployed	1 if unemployed/looks after home/retired/full-time education
M Sick	1 if unable to work due to long-term sickness/disability
Parents' age	
F/M age ≤35 years (Base group)	1 if father/mother's age is less than or equal to 35 years
F/M age 36-45 years	1 if father/mother is aged between 36 and 45 years
F/M age≥46 years	1 if father/mother's age is greater or equal to 46 years
Number of cigarette packs (20 cigarette/pack) smoked by Parents daily	
P non smoker (Base group)	1 if parent is a non smoker
P ≤ 1 cig pack/day	1 if parent smokes up to one pack of cigarette per day
P >1 cig pack/day	1 if parent smokes more than one pack of cigarette per day
Parents marital status	
Parent married (Base group)	1 if parents are married/cohabiting
Parent single	1 if parent is single/divorced/widowed/separated
Survey years	
1997-2001 (Base group)	1 if year of survey is from 1997-2001
2002-2005	1 if year of survey is from 2002-2005
2006-2008	1 if year of survey is from 2006-2008
Parents missing variable indicator	
Single mum	1 if father information are missing because mother is single
Single dad	1 if mother information are missing because father is single
F not home	1 if father information missing because father not home
M not home	1 if mother information missing because mother not home
Both parent missing	1 if information on both parents are missing (boost sample)

SR=self-report

Chapter 3

Decomposing Differences in Cotinine Distribution between Children and Adolescents from High and Low Socioeconomic Backgrounds

3.1 Introduction

The adverse health consequences of passive and active smoking in children and adolescents are well established. Exposure to second-hand smoke or passive smoking has been associated with several adverse health outcomes in children including respiratory illnesses (bronchitis, pneumonia, asthma, coughing and wheezing), recurrent middle ear infection, brain tumours, leukaemia and meningitis (Tobacco Advisory Group of the Royal College of Physicians, 2010, US Department of Health and Human Services, 2006). In children, passive smoking has also been linked to impairments in mental development, affecting both reading and reasoning skills (Yolton et al., 2004) and repeated absence from school due to respiratory illnesses (Gilliland et al., 2003, Charlton, 1996). Impairments in both the physical and mental development of the child could in turn have important consequences for future health outcomes and labour market participation in adulthood (Eriksen, 2004 , Graham and Power, 2004). Similarly, active smoking in adolescents has been linked to several adverse health outcomes both in adolescence and later in adulthood (Center for Disease Control and Prevention, 2004).

The relationship between socioeconomic status, health risk behaviours and health is well established. In adults, higher socioeconomic status (education, income or occupation) is associated with better health outcomes and explaining the association between health and socioeconomic status has been the focus of much research (some examples include Balia and Jones (2008), Contoyannis and Jones (2004), Vallejo-Torres and Morris (2010)). These studies have shown the existence of a strong and robust correlation between socioeconomic status and health risk behaviours, suggesting that the socioeconomic gradient in health can be explained by socioeconomic-related inequalities in health risk behaviours and lifestyle choices such as smoking, excessive alcohol consumption and lack of physical exercise (Balia and Jones, 2008, Contoyannis and Jones, 2004, Vallejo-Torres and Morris, 2010). Furthermore, limited knowledge of the adverse health consequences of health risk behaviours and lifestyle choices may provide further explanations for the link between socioeconomic status and health. For example Kenkel (1991) showed that higher years of schooling is associated with better knowledge of the relationship between lifestyle

choices and health outcomes, resulting in higher allocative efficiency in the production of health (Kenkel, 1991). Other recent studies have shown links between education, health knowledge and lifestyle choices (Cutler and Lleras-Muney, 2010, Peretti-Watel et al., 2007). Cutler and Lleras-Muney (2010) showed that education increases cognitive ability which in turn improves health behaviours. Peretti-Watel et al. (2007) demonstrated that persons with less education are more likely to underestimate the health risk of smoking.

The relationship between child health and parental socioeconomic status has also been widely reported (recent examples include Cameron and Williams (2009), Condliffe and Link (2008) and Currie et al. (2007)). Grossman (1972, 2000), using a health capital model, describes how parental socioeconomic status can affect child health. Child health can be 'produced' using a set of health inputs, the choice of which is determined by the child's parent, subject to a budget constraint and parental preferences. Parental socioeconomic status can therefore affect child health directly or indirectly through its effect on the choice of the health inputs that go into the child health production function. For example the effect of parental socioeconomic status on child health may arise directly because parents with lower income are unable to afford better quality healthcare or high nutritional food for the child, or indirectly, due to parental preferences for risky health behaviours, which in turn impact adversely on child health. On the other hand, parents with lower socioeconomic status may simply have different health beliefs that make them treat health inputs differently from parents with higher socioeconomic status (Currie, 2009).

Smoking is a major factor contributing to socioeconomic variations in adult health and children from lower socioeconomic backgrounds are likely to be exposed to environmental factors that increase both the probability of them initiating smoking or increase the probability of exposure to second-hand smoke. Therefore parental smoking behaviour may, at least in part, explain the socioeconomic gradient in child health. However, recent evidence appears to suggest that the correlation between parental socioeconomic status and child health is not mediated through parental smoking (Frijters et al., 2011, Reinhold and Jürges, 2011).

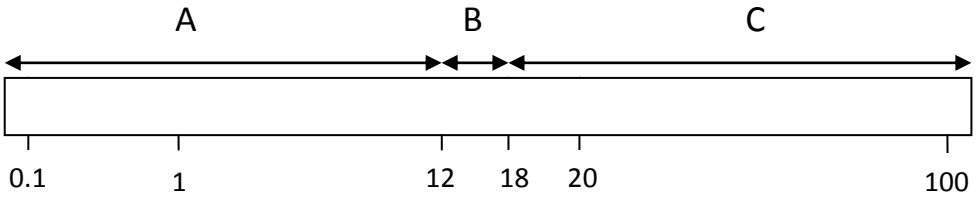
While the association between adult smoking behaviour and socioeconomic status has been widely studied, very few studies investigate the relationship between parental socioeconomic status and passive/active smoking in children and adolescents. In a recent study, Frijters et al. (2011) showed that household income is negatively associated with passive smoking (measured using saliva cotinine). Existing studies on the relationship between active smoking in adolescents and parental socioeconomic status have often

reported conflicting findings (Tuinstra et al., 1998, Glendinning et al., 1994, Soteriades and DiFranza, 2003, Edoka, 2011, Blow et al., 2005). While some studies report a robust negative correlation between parental socioeconomic status and active smoking amongst adolescents (Soteriades and DiFranza, 2003, Edoka, 2011), other studies fail to find an association (Glendinning et al., 1994, Tuinstra et al., 1998) or the association disappears after controlling for parental smoking (Blow et al., 2005). Differences in the indicators of parental socioeconomic status (Gruber and Zinman, 2001, Currie et al., 2008, Tyas and Pederson, 1998), contextual differences of samples as well as the extent of misclassification errors in adolescent self-reported smoking participation (Edoka, 2011), may explain these divergent findings.

The unequal distribution of the determinants of passive and active smoking amongst individuals from different socioeconomic background may explain the social gradient of smoking in children and adolescents. For example, the attenuation or disappearance of the negative association between parental socioeconomic status and adolescent smoking, after controlling for parental smoking, suggest that parental smoking is an important mediator of the socioeconomic gradient in adolescent smoking (some examples include Blow et al. (2005) and Soteriades and DiFranza (2003)). Children from lower socioeconomic backgrounds are more likely to have parents or other family members or friends who smoke, are more likely to live in non-smoke free homes and are more likely to live in deprived neighbourhoods (Action on Smoking and Health (ASH), 2011). In addition to increasing the risks of exposure to second-hand smoke, these factors have been shown to increase the probability of active smoking amongst adolescents (Powell et al., 2005, Loureiro et al., 2010, Powell and Chaloupka, 2005, Sims et al., 2010).

This chapter aims to contribute to a further understanding of the role different determinants play in explaining socioeconomic variations in passive and active smoking amongst children and adolescents. Two groups of children and adolescents are defined using parental socioeconomic status (high and low socioeconomic status) and three decomposition methods are applied to decompose differences in saliva cotinine between the two groups. Saliva cotinine, a major metabolite of nicotine and a quantitative indicator of passive and active exposure to tobacco, is used as a proxy for active and passive smoking (Jarvis et al., 2008, Environmental Protection Agency, 1997).

Figure3.1 Saliva cotinine cut-points (ng/ml)



A: Non-smoker with moderate exposure
B: Non-smoker with higher exposure/light smoker
C: Smoker

Decomposition methods, which were originally developed and applied in the labour economics literature (for example, in explaining gender, regional and inter-country differences in wages, as well as in explaining changes in wage inequalities across time)[14], have now found wider application in other fields including health economics. These decomposition methods allow socioeconomic differences in the distribution of log cotinine to be decomposed into a part explained by differences in the distribution of characteristics (composition effect), and a part explained by differences in the impact of these characteristics (structural effect). Therefore, in addition to quantifying the extent to which the distribution of characteristics explain socioeconomic differences in smoking amongst children and adolescents, the extent to which variations in the impact of these characteristics contribute to socioeconomic differences in smoking, are also identified.

In the first instance, a mean-based decomposition approach (Blinder, 1973, Oaxaca, 1973) is used to decompose differences in mean log cotinine. Then, the empirical analysis is extended to decompose differences between quantiles of log cotinine (Firpo et al., 2009, Melly, 2005). The decomposition of the entire distribution of cotinine allows the simultaneous identification of contributions made by each determinant to socioeconomic differences in both active and passive smoking. Figure 3.1 shows a diagramatic representation of saliva cotinine cut-points for identifying active and passive smoking in the cotinine distribution (based on the findings of Jarvis et al. (2008)). The lower end of the log cotinine distribution is likely to comprise of non-smoking children/adolescents with moderate exposure to second-hand smoke while the top end of the distribution comprises of active smokers.

[14] Fortin et al. (2011) provide an extensive review of the decomposition methods and the applications in labour economics.

The results show that different determinants make greater or lesser contributions at different quantiles of the log cotinine distribution that are indicative of passive and active smoking. For example, the results suggest that smoking within the home explains more of the socioeconomic difference at the lower end of the log cotinine distribution and less of the difference at the upper end of the distribution. Conversely, parental smoking explains more of the difference at the upper end of the log cotinine distribution compared to its contribution at the lower end of the distribution.

The rest of this chapter is organized as follows: section 3.2 gives a description of the data and variables. An empirical framework motivating the choice of variables is also outlined in section 3.2. Section 3.3 broadly defines the parameters of interest in the decomposition analysis and describes the estimation procedures. The results are presented and discussed in section 3.4 and section 3.5 concludes the chapter.

3.2 Data, Variables and Empirical Framework

3.2.1 Data and variables

This chapter uses the 1997/98[15] cross-section of the Health Survey for England (HSE). The HSE is a series of annual cross-sectional surveys which includes a nationally representative sample of households in England. Households are drawn from the Postcode Address file and all adults over the age of 16 years and a random selection of two children aged between 0-15 years living within selected households are interviewed.[16] In addition to individually self-completed questionnaires, each consenting household received a nurse visit during which objective measures of health were taken and saliva specimens collected for cotinine assay. Cotinine assay was performed using gas chromatography which detects cotinine levels as low as 0.1ng/ml. Cotinine is a metabolite of nicotine and with a half-life of approximately 16-20 hours, it can be detected in saliva specimens of regular or occasional smokers or in individuals exposed to second-hand smoke.

The two groups of children and adolescents are defined based on the social class of the household head which was assigned using the Registrars General's Social Class (RGSC) classification system. The RGSC classification system is based on six categories of

[15] Although more recent years of the HSE collected saliva specimens for cotinine assays, specimens were only collected for a small proportion of children. In addition to the nationally representative sample of children, a boost sample of children were surveyed in 1997 and saliva specimens collected, thus providing a larger sample of children with valid cotinine measurements.

[16] A full description of the survey design can be found in Prescott-Clarke (1998).

occupation: professional (I), managerial/technical (II), non-manual skilled (IIIa), manual skilled (IIIb), partly skilled (IV), unskilled (V) and other (VI). Children and adolescents living in households where the head of the household had a professional or managerial/technical occupation were classified as the 'high social class' (HSC) group, while children and adolescents living in households where the head of the household had a partly skilled, unskilled, or any other occupation, were classified as the 'low social class' (LSC) group.

A wide range of demographic and socioeconomic characteristics are available in the HSE. These characteristics include child or adolescent characteristics: age group (8-10/11-12/13/14-15years), gender (male/female) and ethnicity (white/non-white); household characteristics: household location (rural/suburban/urban), non-smoke free homes (defined as whether smoking by members or non-members of the household was permitted within the home) and household income. Finally, parental characteristics include: parental smoking behaviour, highest academic qualification, marital status and age group[17]. These were obtained by linking parents' responses in the individual questionnaires to each child.

Figure 3.2 Log cotinine distributions by social class

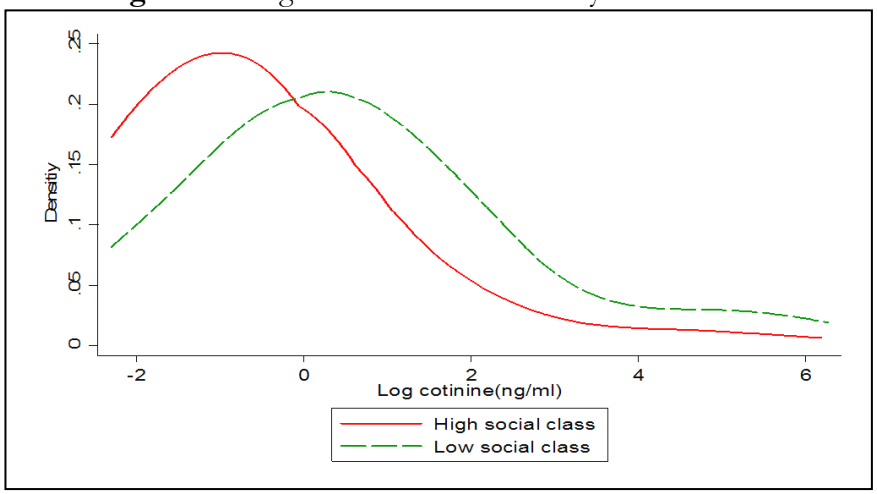


Figure3.2 shows distribution of log cotinine by social class. The log cotinine distribution in the LSC group lies to the right of the log cotinine distribution of the HSC group indicating higher levels of cotinine in children and adolescents in the LSC group compared to those in the HSC group at all quantiles. Table 3.1 shows a comparison of average characteristics of both groups.

[17] A full description of all variables including parents' characteristics is provided in Table A3.1 of the Appendix.

Table 3.1 Mean characteristic by parents' socioeconomic class

Variables	High social class	Low social class	Difference
Log household income	9.872	8.925	0.947**
Household income	23976.72	9247.87	14728.85**
8-10 years old	0.387	0.424	-0.037
11-12 years old	0.257	0.254	0.003
13 years old	0.113	0.104	0.009
14-15 years old	0.243	0.218	0.025
White	0.928	0.888	0.039**
Male	0.497	0.517	-0.019
Urban	0.102	0.177	-0.075**
Suburb	0.590	0.647	-0.057*
Rural	0.308	0.175	0.132**
Non-smoke free homes	0.238	0.587	-0.348**
Mother(M) smokes	0.105	0.251	-0.145**
Father (F) smokes	0.089	0.138	-0.048**
M Degree	0.143	0.020	0.123**
M Below degree	0.077	0.023	0.054**
M NVQ3	0.079	0.025	0.054**
M NVQ2/NVQ1	0.193	0.246	-0.053**
M No qualification	0.067	0.194	-0.127**
F Degree	0.204	0.011	0.193**
F Below degree	0.102	0.027	0.075**
F NVQ3	0.064	0.022	0.043**
F NVQ2/NVQ1	0.083	0.098	-0.015
F No qualification	0.029	0.136	-0.106**
M age ≤35 years	0.091	0.235	-0.144**
M age 36-45 years	0.364	0.230	0.135**
M age≥46 years	0.103	0.045	0.058**
F age ≤35 years	0.044	0.076	-0.033**
F age 36-45 years	0.268	0.163	0.106**
F age≥46 years	0.171	0.056	0.115**
Parent single	0.087	0.239	-0.152**
M missing (single father/mother not home)	0.017	0.016	0.002
F missing (single mother/father not home)	0.094	0.230	-0.136**
Boost sample	0.423	0.476	-0.053*
Observations	1397	958	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

On average, children and adolescents in the HSC group, have more educated and older parents compared to those in the LSC group (Table 3.1). In addition, household income is significantly higher in the HSC group compared to the LSC group (Table 3.1). A significantly higher proportion of mothers (25% vs. 11%) and fathers (14% vs. 9%) smoke within the LSC group compared to the HSC group (Table 3.1). Similarly, a higher proportion of families in the LSC group permit smoking within homes compared to those in the HSC group (59% vs. 24%; Table 3.1).

In 1997, the number of children surveyed was boosted by surveying more households[18]. Although household questionnaires were completed by the head of the household, adults (including parents) from the boost sample were not surveyed. Parent information is therefore missing for all those in the boost sample as well as for those living in a single parent household or in a two-parent household but one parent was absent during the interview. Missing parents are accounted for by including dummy variables for missing information on parental smoking status, highest academic qualification, marital status and age group. The final sample consists of 2355 children and adolescents with 1397 observations from the HSC group and 958 observations from the LSC group.

3.2.2 Empirical framework

The decomposition methods rely on modelling log cotinine as a function of a set of covariates. The conceptual framework for the model adopted in this chapter is based on the Bommier and Stecklov (2002) approach which states that individuals should have equal opportunities to achieve their health potential and inequalities in health arise only due to inequalities in the distribution of unobserved natural factors (or "luck"). These natural factors generally reflect circumstances that are largely beyond the individual's control. The empirical analysis described in this paper focuses on variations in the log cotinine distribution of a young cohort (8-15 year olds) that can be explained by variations in circumstances such as family/parental socioeconomic background that are beyond the individuals' control. These circumstances form the social environment which either reduces the perceived cost of smoking (for active smokers) or increases exposure to second-hand smoke (for non-smokers). Log cotinine (COT) is modelled as a function of a set of covariates that reflect these circumstances.

$$COT_g = \theta_g(X_g, \varepsilon_g)$$

where g denotes HSC or LSC group membership; X is a vector of characteristics including demographic characteristics of the child/adolescent (age, gender and ethnicity) and other characteristics that define the social environment of the child including household characteristics (smoking within homes, home location and household income) and parental characteristics (parents' age, academic qualification and current smoking status); ε is a vector of unobservable characteristics.

[18] Approximately 45% of the final sample comprise of children from the boost sample.

These characteristics have been shown to influence active smoking participation as well as passive smoking amongst children and adolescents. For example parental smoking has been shown to increase the probability of active and passive smoking in children and adolescents (Frijters et al., 2011, Loureiro et al., 2010) while parental income has been shown to be negatively correlated with both passive and active smoking (Edoka, 2011, Frijters et al., 2011, Soteriades and DiFranza, 2003).

3.3 The Decomposition Methods

To decompose differences in log cotinine between children and adolescents from the HSC and LSC groups, three methods are applied: a mean-based decomposition method, the Oaxaca-Blinder (OB) decomposition method (Blinder, 1973, Oaxaca, 1973), and two methods that allow the decomposition of differences in distributional statistics other than the mean, the quantile regression (QR) decomposition method (Machado and Mata, 2005, Melly, 2005) and the recentered influence function regression (RIFR) decomposition method (Firpo et al., 2009, Fortin et al., 2011). These methods allow socioeconomic differences in log cotinine to be decomposed into a part attributable to group differences in the distribution of characteristics (composition effect) and a part attributable to group differences in coefficients (structural effect). In the following sub-sections the restriction assumptions required for identification of the composition and structural effects are outlined formally[19].

3.3.1 Identification

The decomposition methods rely on estimating unconditional counterfactual distributions of the outcome variable. For two mutually exclusive groups, HSC (H) and LSC (L) groups, with observed log cotinine distributions COT_H and COT_L respectively, the unconditional counterfactual distribution is constructed to simulate what the log cotinine distribution of individuals in the HSC group would be if they belonged to the LSC group, or, conversely, what the log cotinine distribution of individuals in LSC group would have been if they belonged to HSC group[20]. To construct these counterfactual distributions, the decomposition methods explore the relationship between log cotinine and a set of observed and unobserved characteristics.

[19] Fortin et al. (2011) provides an extensive discussion of these assumptions. Only assumptions that apply to the decomposition methods applied in this chapter are highlighted here.
[20] In this chapter, the former unconditional counterfactual distribution is constructed.

$$COT_g = \theta_g(X_g, \varepsilon_g), \quad g \in \{H, L\}$$

$$\Delta_{COT} = COT_H - COT_L = [\theta_H(X_H, \varepsilon_H)] - [\theta_L(X_L, \varepsilon_L)] \qquad (1)$$

where X_H and X_L are vectors of observable characteristics, θ_H and θ_L are the functional forms of the log cotinine equation and ε_H and ε_L are vectors of unobservable characteristics for the HSC and LSC groups respectively.

The unconditional counterfactual distribution of log cotinine is generated by integrating the conditional distribution of log cotinine given a set of covariates in one group over the marginal distribution of covariates in the other group. If the unconditional distribution of log cotinine of each group is given by:

$$F_{COT_g}(cot) = \int F_{COT_g|X}(cot|X = x) \, . \, dF_{X_g}(x), \quad g \in \{H, L\} \qquad (2)$$

(where $F_{COT_g|X}(cot|X = x)$ is the conditional distribution of log cotinine and $F_{X_g}(x)$ is the marginal distribution of *X*), the unconditional counterfactual distribution can be generated by either replacing the conditional distribution of log cotinine in one group with the corresponding conditional distribution of the other group or by substituting marginal distribution of covariates. In this chapter the LSC is used as the reference group and a counterfactual distribution, $F^C_{COT_H}$, is constructed by replacing $F_{COT_H|X}(cot|X = x)$ with $F_{COT_L|X}(cot|X = x)$ in equation (2) when $g = H$:

$$F^C_{COT_H}(cot) = \int F_{COT_L|X}(cot|X = x) \, . \, dF_{X_H}(x) \qquad (3)$$

The unconditional counterfactual distribution $F^C_{COT_H}(cot)$ represents the distribution of log cotinine that would have prevailed in the HSC group if the distribution of characteristics were similar to the LSC group.

From equation (1), it follows that the total difference in log cotinine between the two groups can be written as:

$$\Delta_{COT} = \Delta_\theta + \Delta_X + \Delta_\varepsilon$$

where Δ_θ captures group differences in the θ functions **(A)**, Δ_X captures group differences in the distribution of observable characteristics **(B)**, and Δ_ε captures group differences in the distribution of unobservable characteristics **(C)**.

In constructing the unconditional counterfactual distribution $F^{C}_{COT_H}$, replacing the conditional distribution of log cotinine of the HSC group with that of the LSC group replaces both θ and the conditional distribution of ε. Therefore group differences in θ will be confounded by group differences in the distribution of ε[21]. To separate the group differences in ε from the group differences in θ (and *X*), an identification restriction is imposed on the distribution of ε. Under the conditional independence/ignorability assumption, the conditional distribution of ε given *X* is the same for both groups and is independent of group membership ($M_g \perp \varepsilon|X, \ g = H, L$)[22].

In addition to the conditional independence/ignorability assumption, the overlapping support assumption is imposed to rule out cases where observable and unobservable characteristics in the cotinine structural model are different for both groups. This assumption also ensures that no single characteristic can identify membership into any one group (Fortin et al., 2011).

Under these two assumptions, the total difference in log cotinine between the two groups, Δ^{v}_{COT} (where *v* represents a distributional statistics of log cotinine such as the mean or quantiles), can be separated and identified in an aggregate decomposition as:

$$\Delta^{v}_{COT} = \Delta^{v}_{\theta} + \ \Delta^{v}_{X}$$

where $\Delta^{v}_{\theta} = v\left(F_{COT_H} - F^{C}_{COT_H}\right)$, a part explained by group differences in the log cotinine structure (structural effect) and $\Delta^{v}_{X} = v\left(F^{C}_{COT_H} - F_{COT_L}\right)$, a part explained by group differences in the distribution of the observed characteristics (composition effect).

The structural and composition effects can further be decomposed into contributions attributable to each characteristic (detailed decomposition). For the detailed decomposition, additional assumptions are required for the identification of the contribution of each characteristic. These assumptions are specific to the decomposition method and are discussed further in the estimation procedure described for each method in the following sub-section.

[21] The conditional distribution $F_{COT_g|X}(cot|X = x)$ depends on the distribution of ε as follows (Fortin et al. 2011): $F_{COT_g|X}(cot|X = x) = \Pr(\varepsilon \leq \theta_g^{-1}(X, cot)|X = x)$

[22] Where M_g denotes group membership into groups H or L, i.e. g=H or L

3.3.2 Estimation procedures

Oaxaca-Blinder (OB) decomposition method

The mean-based OB decomposition method is based on the assumption that the relationship between log cotinine and a set of characteristics is linear and additive:

$$COT_g = X_g\,\beta_g + \vartheta_g \quad , \qquad E(\vartheta_g) = 0\,, \qquad g \in \{H, L\}$$

where X is a vector of observable characteristics, β is a vector of the slope parameters including the intercept and ϑ is the error term. Given that $E(\vartheta_g) = 0$, the total difference in mean log cotinine, Δ^{μ}_{COT} or $\mu(F_{COT_H} - F_{COT_L})$, can be decomposed as follows:

$$\Delta^{\mu}_{COT} = E(COT_H) - E(COT_L)$$

$$= \underbrace{E(X_H)\beta_H - E(X_H)\beta_L}_{\mathbf{A}} + \underbrace{E(X_H)\beta_L - E(X_L)\beta_L}_{\mathbf{B}} \qquad (4)^{23}$$

where $E(X_H)\beta_L$ is the unconditional counterfactual distribution of log cotinine at the mean[24]. Equation (4) can be rearranged to obtain:

$$\Delta^{\mu}_{COT} = E(X_H)[\beta_H - \beta_L] + [E(X_H) - E(X_L)]\beta_L \qquad (5)$$

Equation (5) is a special case of a more general decomposition and can be further rearranged following Jones and Kelley (1984) to obtain:

$$\Delta^{\mu}_{COT} = E(X_L)[\beta_H - \beta_L] + [E(X_H) - E(X_L)]\,\beta_L + [\,E(X_H) - E(X_L)][\beta_H - \beta_L] \qquad (6)$$

Equation (6) is the three-fold decomposition (Jann, 2008) and is used instead of equation (5) because it separates out group differences in residuals as well as group differences in the interaction between coefficients and characteristics from the structural and composition effects. This allows direct comparisons to be made between results obtained using the three-fold OB decomposition to results obtained using the quantile regression decomposition method described below-which also allows the identification of the group differences in log cotinine attributable to group differences in residuals.

---

[23] These two terms, A and B, are analogous to components A and B described in section 3.3.1

[24] The counterfactual distribution is generated as described in equation (3) at the sample means $\mu(F^{C}_{COT_H}) \rightarrow E(COT^{C}_{H}) = E(X_H)\,\beta_L$

Replacing $E(X_H)$ and $E(X_L)$ by their sample means $\bar{X}_H$ and $\bar{X}_L$, as well as β_H and β_L by their ordinary least square (OLS) estimates, $\hat{\beta}_H$ and $\hat{\beta}_L$, equation (6) can be written as:

$$\hat{\Delta}^{\mu}_{COT} = \underbrace{\bar{X}_L\left(\hat{\beta}_H - \hat{\beta}_L\right)}_{\hat{\Delta}^{\mu}_{\theta}} + \underbrace{(\bar{X}_H - \bar{X}_L)\hat{\beta}_L}_{\hat{\Delta}^{\mu}_{X}} + \underbrace{(\bar{X}_H - \bar{X}_L)\left(\hat{\beta}_H - \hat{\beta}_L\right)}_{\hat{\Delta}^{\mu}_{I}} \qquad (7)^{25}$$

The first term, $\hat{\Delta}^{\mu}_{\theta}$, represents contributions to the total difference in log cotinine between the HSC and LSC groups attributable to group differences in the coefficients including the intercept. The second term, $\hat{\Delta}^{\mu}_{X}$, represents contributions attributable to group differences in the distribution of mean characteristics. There is no clear interpretation for the third term, $\hat{\Delta}^{\mu}_{I}$, because it represents an interaction between group differences in characteristics and coefficients as well as differences in residuals.

An attractive feature of the OB decomposition method is that it can be applied to further decompose the composition and structural effects into contributions attributable to each covariate. This is possible because of the additive linearity assumption. The total composition and structural effect is simply the sum of the contribution of individual covariates:

$$\hat{\Delta}^{\mu}_{X} = \sum_{k=1}^{P} (\bar{X}_{Hk} - \bar{X}_{Lk})\hat{\beta}_{Lk} \qquad (8)$$

and

$$\hat{\Delta}^{\mu}_{\theta} = \left(\hat{\beta}_{H0} - \hat{\beta}_{L0}\right) + \sum_{k=1}^{P} \left(\hat{\beta}_{Hk} - \hat{\beta}_{Lk}\right)\bar{X}_{Lk} \qquad (9)$$

where *k* represents the *k*th covariate and $\hat{\beta}_{H0}$ and $\hat{\beta}_{L0}$ are the estimated intercept coefficients of the HSC and LSC group respectively.

For categorical variables, the result of the detailed decomposition is not invariant to the choice of the base or omitted category. Changing the base category alters the contributions of the other categories as well as the contribution of the categorical variable as a whole[26]. This is accounted for by applying a normalization approach (Yun, 2005). This approach imposes a normalization on the coefficients of the categories by restricting the coefficients

[25] These components are estimated using the 'Oaxaca' STATA command and the three-fold option (Jann, 2008).

[26] This mainly affects results of the detailed decomposition of the structural effect. The contribution to the composition effect is unaffected by the choice of the omitted category.

of the first category to be equal to the unweighted average of the coefficients on the other categories. In addition, the sum of the coefficients are restricted to sum up to zero (Yun, 2005).

Quantile Regression (QR) decomposition method

The QR decomposition method (Melly, 2005)[27] goes beyond the mean and decomposes differences between the two groups across the entire distribution of log cotinine. It allows for the identification of the total structural and composition effect at different quantiles. The unconditional counterfactual distribution ($F^{C}_{COT_H}$) is generated, as defined in equation (3), by integrating the conditional distribution of cotinine in the LSC group ($F_{COT_L|X}$) over the marginal distribution of covariates in the HSC group (F_{X_H}). But for quantiles, the conditional distribution of cotinine in the LSC group is given as:

$$F_{COT_L|X}(cot|X=x) = \int_0^1 1\left[F^{-1}_{COT_L|X}(\tau|X=x) \leq q_\tau\right].d\tau$$

where q_τ is the τth quantile of the unconditional distribution of log cotinine. The counterfactual distribution of log cotinine can be expressed as:

$$F^{C}_{COT_H}(q_\tau) = \int \left[\int_0^1 1\left[F^{-1}_{COT_L|X}(\tau|X=x) \leq q_\tau\right].d\tau\right].dF_{X_H}(x) \qquad (10)$$

Replacing $F^{-1}_{COT_L|X}(\tau|X=x)$ in equation (10) with its consistent conditional quantile regression estimator, $X_L\hat{\beta}_L(\tau)$,[28] and inverting the distribution function $F^{C}_{COT_H}$, the unconditional quantiles of the counterfactual distribution of log cotinine can be recovered. The decomposition of the total difference at the τth quantile, Δ^{τ}_{COT} or $\tau\left(F_{COT_H} - F_{COT_L}\right)$, can then be performed as follows:

$$\hat{q}_\tau(\hat{\beta}_H, X_H) - \hat{q}_\tau(\hat{\beta}_L, X_L) = \underbrace{\left[\hat{q}_\tau(\hat{\beta}_H, X_H) - \hat{q}_\tau(\hat{\beta}_L, X_H)\right]}_{\hat{\Delta}^{\tau}_{\theta}} + \underbrace{\left[\hat{q}_\tau(\hat{\beta}_L, X_H) - \hat{q}_\tau(\hat{\beta}_L, X_L)\right]}_{\hat{\Delta}^{\tau}_{X}} \qquad (11)$$

[27] The QR decomposition method was first proposed by Mata and Machado (2005) and is similar to Melly (2005). In this chapter, Melly's (2005) STATA command (rqdeco3) is applied because it is less computationally demanding.

[28] The conditional quantile function is estimated using quantile regression $F^{-1}_{COT_L|X}(\tau|X=x) = X_L\beta_L(\tau), \forall\, \tau \in (0,1)$.

where $\hat{q}_\tau(\hat{\beta}_L, X_H)$ is the unconditional counterfactual distribution at the τth quartile, $\hat{q}_\tau(\hat{\beta}_H, X_H)$ and $\hat{q}_\tau(\hat{\beta}_L, X_L)$ are the unconditional distribution of log cotinine in the HSC and LSC groups, respectively.

Melly (2005) proposes a way of estimating the contribution of the residuals by defining an N X 1 vector, $\hat{\beta}_{HL}$, with its nth component defined as: $\hat{\beta}_{HL}(\tau_n) = \hat{\beta}_H(0.5) - \hat{\beta}_L(0.5) + \hat{\beta}_L(\tau_n)$, where $\hat{\beta}_H(0.5)$ and $\hat{\beta}_L(0.5)$ are the coefficient vectors of the median regressions for the HSC and LSC groups, respectively.

The overall decomposition in equation (11) can then be expressed as:

$$\begin{aligned}\hat{\Delta}^{\tau}_{COT} &= \hat{q}_\tau(\hat{\beta}_H, X_H) - \hat{q}_\tau(\hat{\beta}_L, X_L) \\ &= \underbrace{\left[\hat{q}_\tau(\hat{\beta}_{HL}, X_H) - \hat{q}_\tau(\hat{\beta}_L, X_H)\right]}_{\hat{\Delta}^{\tau}_{\theta}} + \underbrace{\left[\hat{q}_\tau(\hat{\beta}_L, X_H) - \hat{q}_\tau(\hat{\beta}_L, X_L)\right]}_{\hat{\Delta}^{\tau}_{X}} \\ &+ \underbrace{\left[\hat{q}_\tau(\hat{\beta}_H, X_H) - \hat{q}_\tau(\hat{\beta}_{HL}, X_H)\right]}_{\hat{\Delta}^{\tau}_{R}}\end{aligned}$$

where $\hat{q}_\tau(\hat{\beta}_{HL}, X_H)$ is the distribution of log cotinine at the τth quantile that would have prevailed if median coefficients had been similar to those of the HSC group but the residuals had been similar to those of the LSC group. $\hat{\Delta}^{\tau}_{\theta}$ represents contributions attributable to group differences in (median) coefficients at the τth quantile, $\hat{\Delta}^{\tau}_{R}$ represents contributions attributable to group differences in residuals and $\hat{\Delta}^{\tau}_{X}$ represents contributions attributable to group differences in the distribution of characteristics.

Recentered Influence Function Regression (RIFR) decomposition method

One major limitation of the QR approach is that it cannot be extended to a detailed decomposition. To assess the contributions of individual covariates at different quantiles, the RIFR decomposition method is applied (Firpo et al., 2009). The RIFR decomposition approach, which is based on an unconditional quantile estimator, is analogous to the mean-based OB decomposition method. The RIF regression[29] provides a way of estimating the marginal effect of a vector of covariates (X) on an unconditional distributional statistic of an outcome variable. The marginal effect of X is estimated by regressing a function of the outcome variable, known as the recentered influence function (RIF), on X.

[29] Firpo et al. (2009) describe the RIFR as an unconditional quantile regression, distinct from the conditional quantile regression, because it estimates the marginal effect of X on the *unconditional* quantile of log cotinine.

In this chapter, the RIF of log cotinine at each quantile is estimated directly from the data by first computing sample quantiles q and then estimating the density at each quantile using kernel density methods. An estimate of the RIF of each observation is then obtained using the following equation:

$$RIF(cot; q_\tau) = q_\tau + \frac{\tau - 1[cot \leq q_\tau]}{f_{\text{cot}}(q_\tau)} \qquad (12)$$

where q_τ is the τth quantile of log cotinine and $f_{\text{cot}}(q_\tau)$ is the unconditional density of log cotinine at the τth quantile and $1[cot \leq q_\tau]$ is an indicator function for whether the outcome variable is smaller or equal to the τth quantile. At each quantile, the coefficients on *X* for groups H and L are then estimated by regressing the RIF of log cotinine on *X*[30]:

$$q_{g,\tau} = E_X\left[E\left[\widehat{RIF}(cot_g; q_{g,\tau})|X_g\right]\right] = E[X_g]\hat{\delta}_{g,\tau} \ , \quad g \in \{H, L\} \qquad (13)$$

where $q_{g,\tau}$ is the unconditional τth quantile of log cotinine for group $g \in \{H, L\}$ and $\hat{\delta}_{g,\tau}$ is the coefficient of the RIF regression which captures the marginal effect of a change in distribution of *X* on the unconditional quantile of log cotinine. Equation (13) is analogous to the basis of the OB decomposition at the mean and the difference in log cotinine between the two groups at the τth quantile of log cotinine can be decomposed as follows:

$$\hat{\Delta}^{\tau}_{COT} = \left[\widehat{RIF}(cot_H, q_{H,\tau})\right] - \left[\widehat{RIF}(cot_L, q_{L,\tau})\right]$$

$$\hat{\Delta}^{\tau}_{COT} = \underbrace{\bar{X}_L\left(\hat{\delta}_{H,\tau} - \hat{\delta}_{L,\tau}\right)}_{\hat{\Delta}^{\tau}_{\theta}} + \underbrace{(\bar{X}_H - \bar{X}_L)\hat{\delta}_{L,\tau}}_{\hat{\Delta}^{\tau}_{X}} + \underbrace{(\bar{X}_H - \bar{X}_L)\left(\hat{\delta}_{H,\tau} - \hat{\delta}_{L,\tau}\right)}_{\hat{\Delta}^{\tau}_{I}}$$

Similarly, the composition and structural effects can be further decomposed into contributions of each covariate at the τth quantile in a detailed decomposition similar to equations (8) and (9).

Table 3.2 Descriptive statistics: Differences in log cotinine distribution

	High social class	**Low social class**	**Difference**
Mean	-0.639	0.612	-1.251**
25th percentile	-1.609	-0.693	-0.916**
50th percentile	-0.916	0.47	-1.386**
75th percentile	0	1.482	-1.482**

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Sample weights applied.

[30] This can be performed using the STATA 'rifreg ' command which is available for download as an RIF-regression STATA ado file from Firpo et al. (2009): http://faculty.arts.ubc.ca/nfortin/datahead.html.

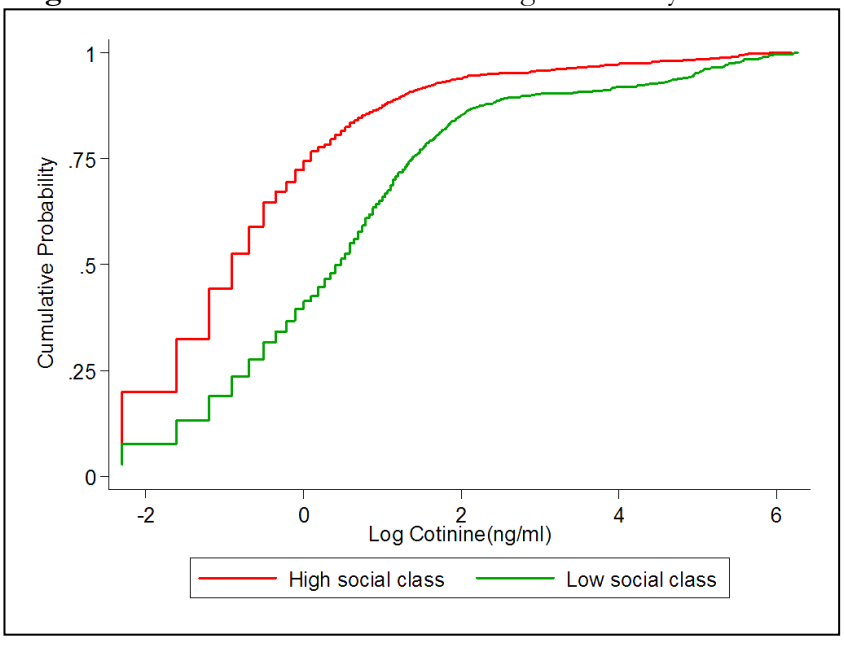
3.4 Results and Discussion

3.4.1 Aggregate decomposition

Children and adolescents in the LSC group are more likely to be exposed to social and environmental factors that either increase the probability of them becoming active smokers or increase the risk of exposure to second-hand smoke. On average, log cotinine in the LSC group is significantly higher than the HSC group (0.612 vs. -0.639; Table 3.2), by approximately 1.251 points. Similarly, at all quartiles, log cotinine is higher in the LSC group compared to the HSC group, with the gap between the two groups increasing at higher quantiles (Table 3.2). This is depicted graphically in Figure 3.3 which shows a widening of the gap between the two groups moving up the log cotinine distribution.

The results of the aggregate decomposition analysis are shown in Table 3.3. The OB decomposition shows that differences in mean characteristics account for a large proportion of the total difference between the two groups. If mean characteristics of the HSC group had been distributed similar to those of the LSC group, the total difference in average log cotinine between both groups would decrease by approximately 1.115 points (upper panel, Table 3.3). Therefore approximately 89% of the total difference in average log cotinine is explained by differences in the distribution of characteristics (composition effect).

Figure 3.3 Cumulative distribution of log cotinine by social class



However, the composition effect does not entirely account for the total difference and approximately 34% of the total difference is attributable to group differences in coefficients (structural effect). On the other hand, the interaction effect is positive. However, the interpretation of this effect is not unambiguous since it captures not only group differences in residuals but also the interaction between group differences in characteristics and coefficients.

Table 3.3 (middle and lower panels) shows the results of the aggregate decomposition at different quartiles. Similar to the mean, the difference in log cotinine attributable to differences in characteristics explains a larger proportion of the total difference across all three quartiles, compared to the structural effect. In addition, the composition effect is greatest at lower quartiles compared to higher quartiles. The difference in log cotinine attributable to differences in characteristics, coefficients and residuals is depicted graphically in Figure 3.4. In Figure 3.4, from the lowest quantile up to the median, the 'composition effect line' lies close to, and follows the same direction as the 'total difference line'. This implies that up to the median, differences in the distribution of characteristics explain a large proportion of the difference between the two groups. From the median, the two lines diverge implying that the composition effect explains less of the total difference between the two groups. Since lower quantiles of the log cotinine distribution are likely to comprise of passive smokers and higher quantiles, of active smokers, this result suggest that the distribution of observed characteristics explains more of the socioeconomic differences in passive smoking compared to active smoking.

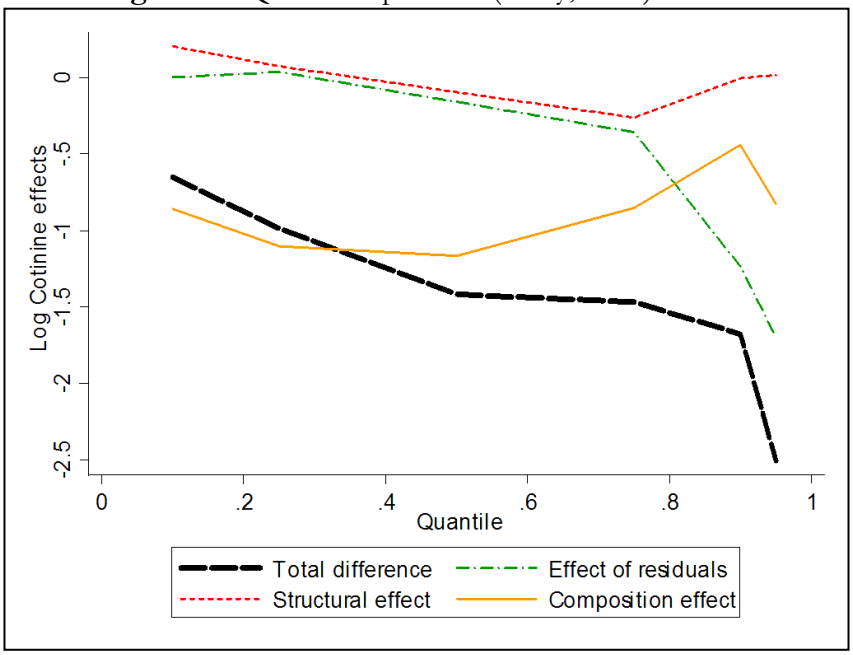
Interestingly this trend corresponds to an increasing contribution of the residuals to the total difference at higher quantiles. Unlike the interaction effect in the OB and RIFR decomposition methods, the residual effect in the QR decomposition can be interpreted as the extent to which differences in residuals contribute to the total difference in log cotinine (Melly, 2005). At the third quartile, the residuals account for approximately 24% of the total difference. Therefore, the contribution of the residuals may reflect group differences in unobserved characteristics not accounted for in the model such as attitude towards risk or rate of time preference, which may be more important in explaining socioeconomic differences in active smoking (compared to passive smoking) amongst children and adolescents.

Table 3.3 Aggregate decomposition of socioeconomic differences in log cotinine distribution

Oaxaca-Blinder Decomposition						
Difference attributable to:	**Mean**	***% Change***				
Characteristics	-1.115***	89%				
	(0.132)					
Coefficients	-0.424***	34%				
	(0.0977)					
Interaction	0.288*	-23%				
	(0.140)					
Total difference	-1.251***	*100%*				
	(0.0818)					
CQR Decomposition (Melly, 2005)						
Difference attributable to:	**Q25**	***%Change***	**Q50**	***%Change***	**Q75**	***%Change***
Characteristics	-1.101**	112%	-1.166**	82%	-0.851**	58%
	(0.193)		(0.147)		(0.127)	
Coefficients	0.076	-8%	-0.094	7%	-0.261	18%
	(0.234)		(0.199)		(0.188)	
Residuals	0.038	-4%	-0.158	11%	-0.356**	24%
	(0.125)		(0.112)		(0.142)	
Total difference	-0.987**	100%	-1.418**	100%	-1.467**	100%
	(0.103)		(0.071)		(0.095)	
RIFR Decomposition (Firpo et al., 2009)						
Difference attributable to:	**Q25**	***%Change***	**Q50**	***%Change***	**Q75**	***%Change***
Characteristics	-1.389**	*120%*	-1.379**	*106%*	-0.839**	*65%*
	(0.206)		(0.161)		(0.164)	
Coefficients	-0.644*	*56%*	-0.626*	*48%*	-0.0268	*2%*
	(0.314)		(0.248)		(0.302)	
Interaction	0.874**	*-75%*	0.703**	*-54%*	-0.424**	*33%*
	(0.219)		(0.173)		(0.198)	
Total difference	-1.158**	*100%*	-1.302**	*100%*	-1.290**	*100%*
	(0.297)		(0.229)		(0.263)	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses; Sample weights applied.

Figure 3.4 QR decomposition (Melly, 2005)



Although it has been recently disputed[31], some authors argue that the socioeconomic gradient in health risk behaviours can be explained by differences in the degree of risk aversion and rates of time preferences (Becker and Mulligan, 1997, Leigh, 1986). Adults with lower socioeconomic status are more likely to underestimate the potential health hazards of smoking (Peretti-Watel et al., 2007), are more likely to be present-oriented or have less incentive to invest in future health benefits and thus have higher discount rates of time preference in comparison to those with higher socioeconomic status (Becker and Mulligan, 1997, Leigh, 1986). The existence of a strong intergenerational transmission of the willingness to take risk (including health risk) and rates of time preferences (Breuer et al., 2011, Dohmen et al., 2008), implies that children in the LSC group may be more likely to adopt their parents' rate of time preference or attitude towards risk. In addition to reducing the perceived future cost of engaging in health risk behaviours, the social environment of a child may directly influence consumption preferences, with children emulating the consumption preference of their parents.

3.4.2 Detailed decomposition

Composition effects

The results of the detailed decomposition provide more insight into the contributions of individual covariates to the composition and structural effects. Table 3.4 shows the results

[31] Some examples include Cutler and Lleras-Muney (2010) and Khwaja et al. (2007).

of the detailed decomposition. At the mean, the composition effect is driven mainly by difference in the distribution of homes within which smoking is permitted, by household income, fathers' education and by mothers' smoking status.

Smoking within the home makes the largest contribution to the overall composition effect, accounting for approximately 44% of the total composition effect at the mean. This is unsurprising, given that on average, a larger proportion of children and adolescents in the LSC group live in non-smoke free homes in comparison to those in the HSC group (59% vs. 24%; Table 3.1). The resulting higher levels of saliva cotinine in the LSC group may either be as a direct consequence of higher exposure to second-hand smoke or indirectly through less discouragement of experimentation or of active smoking participation. The detailed decomposition of the entire log cotinine distribution sheds more light on the contributions of smoking within the home to the composition effect at different quartiles. Similar to its contribution at the mean, differences in the distribution of non-smoke free homes make the highest contribution to the composition effect at all three quartiles. The composition effect of non-smoke free homes varies at different quartiles and a distinct pattern is observed. At the first quartile, differences in the distribution of non-smoke free homes account for approximately 57% of the total composition effect. This decreases to 45% at the median and approximately 29% at the last quartile. This suggests that smoking within homes explains more of the difference in passive smoking compared to its contribution to differences in active smoking.

At the mean, differences in the distribution of household income account for approximately 21% of the total composition effect. In addition, at the first and last quartile, although statistically insignificant, differences in the distribution of household income account for approximately 12% and 19% of the total composition, respectively. At the median, the contribution of income is statistically significant and accounts for approximately 18% of the total composition effect. Interestingly, this result suggests that household income explains more of the socioeconomic difference in active smoking compared to passive smoking.

Differences in the distribution of mothers' smoking status make statistically significant contributions to the composition effect accounting for approximately 16% of the total composition effect at the mean. Differences in the distribution of mothers smoking equally makes statistically significant contributions at all three quartiles of the log cotinine distribution with the highest contribution observed at the last quartile (23% compared to 7% at the first quartile). This pattern (as well as that observed with household income) is in

direct contrast to contributions of non-smoke free homes, where smoking within homes makes the highest contribution to the composition effect at the lowest quartile, suggesting that parental smoking and household income explains more of the socioeconomic differences in active smoking compared to passive smoking. Since children/adolescents whose parents smoke are more likely to become smokers themselves (Loureiro et al., 2010, Powell and Chaloupka, 2005), and that a higher proportion of mothers smoke in the LSC group compared to the HSC group (Table 3.1), the contribution of mothers' smoking at the highest quartile suggests a higher proportion of child/adolescent active smokers within the LSC group compared to the HSC group. The contribution of mothers' smoking behaviour to the composition effect at the first quartile suggests a higher exposure of those in the LSC group to second-hand smoke.

Father's education makes statistically significant contributions to the composition effect at the mean. Similarly, the composition effect of mothers' education is negative at the mean and at all quartiles but is statistically significant only at the first quartile, accounting for 15% of the total difference in log cotinine at the first quartile. In adults, the link between years of schooling/education, health knowledge and lifestyle choices are well established (Cutler and Lleras-Muney, 2010, Kenkel, 1991, Peretti-Watel et al., 2007). Therefore, the composition effect of parental education may reflect group differences in the distribution of health knowledge or lifestyle choice, which may in turn, result in higher exposure to second-hand smoke and active smoking amongst those in the LSC group.

Structural effects

Results of the detailed decomposition provide further interesting insights into the contributions of individual covariates to the total structural effects. At the mean and median, father's education makes statistically significant contributions to the total structural effect. This suggests a differential impact of father's education across socioeconomic groups. More insight into group differences in the impact of father's education is shown in Tables A3.2-A3.5 of the Appendix. For example, in the HSC group, for children/adolescents whose fathers have no qualification, average log cotinine is approximately 11% higher compared to those whose fathers have a university degree (Table A3.2). The corresponding estimate in the LSC group is approximately 43% (Table A3.2). Similar results are observed across all categories of father's education except for the NVQ3 qualification/equivalent category (Table A3.2).

Table 3.4 Detailed decomposition of socioeconomic differences in log cotinine distribution

	OB Decomposition		RIFR Decomposition					
Difference attributable to:	Mean	%Change	Q25	%Change	Q50	%Change	Q75	%Change
Characteristics								
Log household income	-0.231*	21%	-0.168	12%	-0.255*	18%	-0.161	19%
	(0.0926)		(0.106)		(0.112)		(0.114)	
Gender	0.000683	0%	-0.000439	0%	0.000221	0%	0.00442	-1%
	(0.00189)		(0.00306)		(0.00300)		(0.00734)	
Ethnicity	0.0120	-1%	0.0113	-1%	0.0152	-1%	0.00160	0%
	(0.00747)		(0.00900)		(0.00853)		(0.00857)	
Age last birthday	0.0399	-4%	0.00650	0%	0.00143	0%	0.0335	-4%
	(0.0250)		(0.00865)		(0.00917)		(0.0215)	
Home location	-0.00801	1%	0.0507	-4%	-0.00967	1%	-0.00900	1%
	(0.0218)		(0.0270)		(0.0241)		(0.0280)	
Non-smoke-free homes	-0.496**	44%	-0.797**	57%	-0.615**	45%	-0.242**	29%
	(0.0593)		(0.0774)		(0.0666)		(0.0649)	
Mother's smoking status	-0.183**	16%	-0.0917*	7%	-0.0799*	6%	-0.190**	23%
	(0.0468)		(0.0375)		(0.0378)		(0.0465)	
Father's smoking status	-0.0405	4%	0.00506	0%	-0.0360	3%	-0.0418	5%
	(0.0668)		(0.0170)		(0.0186)		(0.0220)	
Mother's education	-0.0527	5%	-0.203*	15%	-0.136	10%	-0.0870	10%
	(0.0483)		(0.0954)		(0.0750)		(0.0717)	
Father's education	-0.388**	35%	-0.0623	4%	-0.193	14%	-0.0485	6%
	(0.147)		(0.160)		(0.118)		(0.122)	
Mother's age	0.0365	-3%	-0.0303	2%	0.00864	-1%	0.0551	-7%
	(0.0428)		(0.0388)		(0.0372)		(0.0466)	
Father's age	0.225	-20%	0.0494	-4%	-0.0203	1%	-0.0456	5%
	(0.124)		(0.0325)		(0.0299)		(0.0384)	
Marital status	-0.0315	3%	-0.159**	11%	-0.0592	4%	-0.109	13%
	(0.0649)		(0.0602)		(0.0507)		(0.0625)	
Total	-1.115**	100%	-1.389**	100%	-1.379**	100%	-0.839**	100%
	(0.132)		(0.206)		(0.161)		(0.164)	

Table 3.4 (continued) Detailed decomposition

	OB Decomposition		RIFR Decomposition					
Difference attributable to:	Mean	%Change	Q25	%Change	Q50	%Change	Q75	%Change
Coefficients								
Log household income	0.851	-201%	0.336	-52%	1.132	-181%	-0.676	2522%
	(1.023)		(1.173)		(1.169)		(1.311)	
Gender	0.00265	-1%	0.0410	-6%	-0.00441	1%	0.137	-511%
	(0.0681)		(0.0756)		(0.0727)		(0.0923)	
Ethnicity	-0.103	24%	-0.323	50%	-0.357	57%	0.350	-1306%
	(0.194)		(0.246)		(0.213)		(0.283)	
Age last birthday	0.0750	-18%	0.0323	-5%	-0.0322	5%	0.0831	-310%
	(0.0383)		(0.0400)		(0.0380)		(0.0506)	
Home location	-0.000782	0%	0.00731	-1%	-0.0882	14%	0.0140	-52%
	(0.0432)		(0.0512)		(0.0467)		(0.0575)	
Non-smoke-free homes	-0.0716	17%	-0.954**	148%	-0.312**	50%	0.854**	-3187%
	(0.113)		(0.128)		(0.115)		(0.153)	
Mother's smoking status	-0.00410	1%	-0.0474	7%	-0.0252	4%	-0.0120	45%
	(0.0525)		(0.0744)		(0.0741)		(0.0996)	
Father's smoking status	-0.00454	1%	0.0571	-9%	-0.0527	8%	-0.141*	526%
	(0.0208)		(0.0487)		(0.0480)		(0.0646)	
Mother's education	-0.0620	15%	-0.333	52%	-0.150	24%	0.121	-451%
	(0.0911)		(0.207)		(0.160)		(0.190)	
Father's education	-0.495**	117%	-0.0650	10%	-0.177*	28%	0.00579	-22%
	(0.186)		(0.109)		(0.0890)		(0.102)	
Mother's age	0.378*	-89%	-0.0530	8%	-0.105*	17%	0.0275	-103%
	(0.153)		(0.0639)		(0.0505)		(0.0598)	
Father's age	0.0845	-20%	-0.00692	1%	-0.0120	2%	0.00473	-18%
	(0.0717)		(0.0254)		(0.0219)		(0.0280)	
Marital status	0.0240	-6%	0.0151	-2%	0.00387	-1%	0.00203	-8%
	(0.0284)		(0.0127)		(0.00862)		(0.0107)	
Constant	-1.098	259%	0.650	-101%	-0.446	71%	-0.796	2970%
	(1.106)		(1.252)		(1.239)		(1.400)	
Total	-0.424**	100%	-0.644*	100%	-0.626*	100%	-0.0268	100%
	(0.0977)		(0.314)		(0.248)		(0.302)	

Table 3.4 (continued) Detailed decomposition

	OB Decomposition		RIFR Decomposition					
Difference attributable to:	Mean	%Change	Q25	%Change	Q50	%Change	Q75	%Change
Interaction								
Log household income	0.0914	32%	0.0361	4%	0.122	17%	-0.0726	17%
	(0.110)		(0.126)		(0.126)		(0.141)	
Gender	-0.000075	0%	-0.00116	0%	0.000125	0%	-0.00387	1%
	(0.00193)		(0.00411)		(0.00361)		(0.00722)	
Ethnicity	-0.00396	-1%	-0.0124	-1%	-0.0137	-2%	0.0134	-3%
	(0.00758)		(0.0110)		(0.00999)		(0.0125)	
Age last birthday	-0.0154	-5%	-0.00123	0%	0.00477	1%	-0.0167	4%
	(0.0114)		(0.00982)		(0.00901)		(0.0141)	
Home location	0.0000751	0%	-0.0179	-2%	0.00934	1%	-0.0104	2%
	(0.0257)		(0.0319)		(0.0290)		(0.0347)	
Non-smoke-free homes	0.0431	15%	0.574**	66%	0.188**	27%	-0.514**	121%
	(0.0679)		(0.0816)		(0.0700)		(0.0951)	
Mother's smoking status	0.0460	16%	0.0278	3%	0.0147	2%	0.00702	-2%
	(0.0532)		(0.0438)		(0.0436)		(0.0585)	
Father's smoking status	0.00596	2%	-0.0230	-3%	0.0212	3%	0.0569*	-13%
	(0.0837)		(0.0205)		(0.0201)		(0.0288)	
Mother's education	0.0476	17%	0.141	16%	0.0963	14%	0.124	-29%
	(0.0559)		(0.102)		(0.0808)		(0.0864)	
Father's education	0.382*	133%	0.0798	9%	0.207	29%	-0.0319	8%
	(0.150)		(0.173)		(0.131)		(0.143)	
Mother's age	-0.0342	-12%	0.0188	2%	0.0397	6%	-0.0287	7%
	(0.0547)		(0.0536)		(0.0504)		(0.0655)	
Father's age	-0.251	-87%	-0.0745	-9%	-0.0187	-3%	0.0348	-8%
	(0.131)		(0.0422)		(0.0382)		(0.0502)	
Marital status	-0.0239	-8%	0.127	15%	0.0325	5%	0.0170	-4%
	(0.0763)		(0.0719)		(0.0619)		(0.0786)	
Total	0.288*	100%	0.874**	100%	0.703***	100%	-0.424*	100%
	(0.140)		(0.219)		(0.173)		(0.198)	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses. Sample weights applied.

The result of the decomposition analysis suggests that group differences in the coefficients of father's education contribute significantly to the difference in log cotinine observed between both groups. The definition of social class in this chapter is based on parent's occupation (the RGSC classification system). Thus, for fathers with similar levels of education, those in the HSC group are likely, on average, to have higher occupational status, earn more and perhaps live in more affluent neighbourhoods compared to fathers in the LSC group. Therefore, the impact of lower parental educational attainment on log cotinine concentration in children/adolescents from the HSC group may be mitigated by other favourable social circumstances that limit exposure to second-hand smoke and/or discourage active smoking.

At the lowest quartile and at the median, smoking within homes contributes significantly to the total structural effect. This suggests that smoking within homes also exerts a differential impact on log cotinine in children/adolescents from both groups. Although smoking within homes is associated with an increase in log cotinine in both groups, this effect is less in the HSC group[32]. This suggests possible group differences in the behaviours of parents. For example parents in the HSC group may adopt avoidance behaviours which limit the child's exposure to tobacco smoke, such as restricting smoking within the home to specific rooms or to specific times when the child is unlikely to be present. In this study sample, in households where either parent smokes, 94% of parents in the LSC group permit smoking within the home compared to 72% in the HSC group. This proportion increases to 100% and 87% respectively, when both parents smoke. The contribution of smoking within homes at the third quartile also suggests differential impact of smoking within homes on smoking behaviours of children/adolescents. However unlike at lower quartiles, the impact of smoking within homes is greater in the HSC group compared to the LSC group at the third quartile.

Further evidence of possible differences in parental behaviour is demonstrated by the differential impact of parental smoking. The structural effect of father's smoking is negative at the mean, median and last quartile (although statistically significant only at the last quartile). Similarly, the structural effect of mothers' smoking is negative at the mean and across all quartiles (although statistically insignificant). Given the strong correlation between health knowledge, health risk perception and lifestyle choices (Cutler and Lleras-Muney, 2010, Kenkel, 1991, Peretti-Watel et al., 2007), the structural effect of parental smoking may reflect group differences in parental health risk perceptions or knowledge of

[32] See Tables A3.2-A3.4 in the Appendix

the adverse health consequences of their lifestyle choices on child health, thus explaining group differences in parental attitudes towards protecting their child from the harmful effects of smoking.

Taken together, these results suggests that not only do a lower proportion of parents on the HSC group engage in health risk behaviours such as smoking and smoking within the home, parents in the HSC group who engage in these health risk behaviours adopt other compensatory or avoidance behaviours that either discourage their child from active smoking participation or limit their child's exposure to second-hand smoke.

3.5 Conclusion

Socioeconomic inequality in adolescent smoking has received very little attention in the economics literature. Chapter 3 sheds more light on the factors that contribute to socioeconomic differences in both passive and active smoking amongst children and adolescents aged 8 to 15 years. In the first instance, a mean-based decomposition method is applied to assess contributions of various characteristics or determinants of smoking to differences in average log cotinine between two groups of children and adolescents defined by parental socioeconomic status. The analysis is then extended to decompose the entire distribution of log cotinine. The lower end of the log cotinine distribution is likely to comprise of non-smokers exposed to second-hand smoke (passive smokers) while the upper end of the distribution is likely to comprise of active smokers. Therefore, the extension of the decomposition analysis to different quartiles provides useful insight into the distinct roles different determinants play in explaining socioeconomic differences in active and passive smoking amongst children and adolescents. Log cotinine is modelled as a function of a set of characteristics, which are considered to explain within- and between-group variations in log cotinine. Differences in log cotinine between the two groups are decomposed into a part explained by differences in the distribution of characteristics (composition effect) and a part explained by differences in the impact of these characteristics (structural effect).

The results show that the composition effect accounts for a large proportion of the total difference between the two groups both at the mean and at different quartiles of the log cotinine distribution. At lower quartiles (indicative of passive smoking), the composition effect explains more of the difference in log cotinine and less of the difference at higher quartiles (indicative of active smoking). Conversely, group differences in the distribution of

residuals explain more of the socioeconomic variations in active smoking compared to passive smoking.

Characteristics making the largest contribution to the total composition effect are smoking within the home, father's education, household income and mother's smoking. These characteristics make different contributions at different quartiles of the log cotinine distribution, suggesting distinct roles in explaining socioeconomic differences in passive versus active smoking. Smoking within the home explains more of socioeconomic differences in passive smoking compared to socioeconomic differences in active smoking while household income and parental smoking behaviour explain more of socioeconomic differences in active smoking compared to passive smoking. Given that smoking in adults follows a socioeconomic gradient, children and adolescents in the LSC group are more likely to have parents who smoke or live in homes where smoking is permitted within the home. These factors are likely to increase the probability of passive or active smoking in children/adolescents, thus explaining, at least in part, the higher levels of log cotinine observed in the LSC group.

Although the composition effect explains a large part of the difference between the two groups, it does not tell the whole story. Conditional on having similar characteristics, the differential impact of these characteristics (structural effect) contributes (albeit to a smaller extent) to differences in log cotinine between the two groups. The structural effect attributable to smoking within homes and parental smoking suggest group differences in parental health risk behaviours or attitudes which may in turn limit or increase the impact of parental smoking on child/adolescent passive or active tobacco consumption. In tackling socioeconomic inequalities in smoking amongst children and adolescents, these findings highlight the importance of not only reducing the distribution of factors that are associated with an increased probability of passive and active smoking, but by also encouraging positive behavioural changes in parents through promoting greater awareness of the negative externalities of parental tobacco consumption, particularly amongst those from lower socioeconomic backgrounds.

3.6 Appendix

Table A3.1 Description of Variables

Variable name	Variable Label
Male	1 if male
White	1 if white
8-10 years old (base group)	1 if aged between 8 and 10 years
11-12 years old	1 if aged between 11 and 12 years
13 years old	1 if aged 13 years
14-15 years old	1 if aged between 14 and 15 years
Household characteristics	
Log household income	Log household income
Urban	1 if lives in inner city/other dense urban or city centre
Suburb	1 if lives in a suburb residential (city/large town outskirts)
Rural	1 if lives in rural residential/village centre/rural isolated dwelling
Non-smoke free homes	1 if no one smokes inside the house/flat on most days
Mother(M) smokes	1 if mother is an smoker
Father (F) smokes	1 if father is a smoker
Father's highest qualification	
F Degree	1 if NVQ4/NVQ5/Degree or equivalent
F Below degree	1 if higher education below degree
F NVQ3	1 if NVQ3/GCE A level equivalent
F NVQ2/NVQ1	1 if NVQ2/GCE O level/NVQ1/CSE or equivalent/foreign
F No qualification	1 if no qualification
Mother's highest qualification	
M Degree	1 if NVQ4/NVQ5/Degree or equivalent
M Below degree	1 if higher education but below degree
M NVQ3	1 if NVQ3/GCE A level equivalent
M NVQ2/NVQ1	1 if NVQ2/GCE O level/NVQ1/CSE or equivalent/foreign
M No qualification	1 if no qualification
Parent's age	
F/M age ≤35 years	1 if father/mother's age is less than or equal to 35 years
F/M age 36-45 years	1 if father/mother is aged between 36 and 45 years
F/M age≥46 years	1 if father/mother's age is greater or equal to 46 years
Parents marital status	
Parent married	1 if parents are married/cohabiting
Parent single	1 if parent is single/divorced/widowed/separated
Parents missing variable indicator	
F missing	1 if father information is missing because mother is single/father not home
M missing	1 if mother information is missing because father is single/mother not home
Boost sample	1 if information on both parents are missing (boost sample)

Table A3.2 OLS regression of Log cotinine

	High social class		**Low social class**	
Variables	Coeff.	Std. error	Coeff.	Std. error
Log household income	-0.150*	(0.0613)	-0.244*	(0.0965)
Male	-0.0385	(0.0767)	-0.0574	(0.108)
White	0.229	(0.133)	0.353*	(0.173)
8-10 years old (base group)				
11-12 years old	-0.00630	(0.0697)	-0.00767	(0.0945)
13 years old	0.213	(0.111)	0.627**	(0.225)
14-15 years old	0.876**	(0.131)	1.337**	(0.196)
Urban (base group)				
Suburb	0.0225	(0.126)	0.0409	(0.144)
Rural	-0.0456	(0.132)	-0.0354	(0.183)
Non-smoke free homes	1.253**	(0.120)	1.363**	(0.144)
Mother(M) smokes	0.635**	(0.148)	0.817**	(0.194)
Father (F) smokes	0.0737	(0.144)	0.166	(0.236)
M Degree (base group)				
M Below degree	-0.0662	(0.160)	0.236	(0.438)
M NVQ3	0.00323	(0.184)	0.806	(0.422)
M NVQ2/NVQ1	-0.0470	(0.158)	0.447	(0.350)
M No qualification	0.00673	(0.201)	0.622	(0.368)
F Degree (base group)				
F Below degree	0.145	(0.166)	0.261	(0.552)
F NVQ3	-0.0198	(0.145)	-0.584	(0.455)
F NVQ2/NVQ1	0.247	(0.181)	0.345	(0.428)
F No qualification	0.109	(0.240)	0.426	(0.433)
M age ≤35 years(base group)				
M age 36-45 years	0.0216	(0.139)	0.182	(0.185)
M age≥46 years	0.0383	(0.206)	0.107	(0.360)
F age ≤35 years(base group)				
F age 36-45 years	0.0241	(0.157)	0.0332	(0.228)
F age≥46 years	-0.199	(0.195)	-0.360	(0.332)
Parent single	0.429*	(0.178)	0.468*	(0.213)
M missing	0.109	(0.345)	1.993**	(0.716)
F missing	0.0638	(0.216)	0.0179	(0.488)
Boost sample	0.448*	(0.185)	1.362**	(0.498)
Constant	-0.182	(0.671)	-0.0879	(1.032)
Observations	1397		958	
R^2	0.275		0.380	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses.
Sample weights applied.

Table A3.3 Q25 RIF regression of Log cotinine

	High social class		Low social class	
Variable	Coeff.	Std. error	Coeff.	Std. error
Log household income	-0.140	(0.0719)	-0.177	(0.110)
Male	0.110	(0.0887)	0.0291	(0.117)
White	-0.0346	(0.167)	0.330	(0.221)
8-10 years old (base group)				
11-12 years old	0.0910	(0.112)	-0.0731	(0.150)
13 years old	0.105	(0.155)	0.335	(0.210)
14-15 years old	0.162	(0.120)	0.124	(0.167)
Urban (base group)				
Suburb	0.103	(0.160)	0.161	(0.173)
Rural	0.276	(0.173)	0.428	(0.223)
Non-smoke free homes	0.618***	(0.109)	2.206***	(0.178)
Mother(M) smokes	0.363**	(0.140)	0.528*	(0.210)
Father (F) smokes	0.268	(0.148)	-0.0712	(0.239)
M Degree (base group)				
M Below degree	0.137	(0.220)	0.489	(0.701)
M NVQ3	0.0760	(0.235)	1.196	(0.766)
M NVQ2/NVQ1	0.276	(0.187)	0.940	(0.582)
M No qualification	0.185	(0.232)	1.072	(0.584)
F Degree (base group)				
F Below degree	0.106	(0.186)	-0.155	(0.798)
F NVQ3	0.0392	(0.223)	-0.994	(0.779)
F NVQ2/NVQ1	0.0644	(0.215)	0.0673	(0.696)
F No qualification	-0.188	(0.279)	0.524	(0.682)
M age ≤35 years(base group)				
M age 36-45 years	-0.114	(0.190)	-0.112	(0.202)
M age≥46 years	0.00388	(0.271)	-0.344	(0.395)
F age ≤35 years(base group)				
F age 36-45 years	-0.173	(0.239)	0.405	(0.284)
F age≥46 years	-0.243	(0.281)	0.522	(0.376)
Parent single	0.190	(0.208)	0.824**	(0.305)
M missing	0.513	(0.382)	0.535	(0.782)
F missing	0.0937	(0.319)	-0.315	(0.770)
Boost sample	0.510	(0.267)	1.785*	(0.783)
Constant	-0.736	(0.791)	-2.552*	(1.257)
Observations	1397		958	
R^2	0.106		0.372	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses.
Sample weights applied.

Table A3.4 Q50 RIF regression of Log cotinine

	High social class		Low social class	
	Coeff.	Std. error	Coeff.	Std. error
Log household income	-0.140*	(0.0596)	-0.269*	(0.117)
Male	-0.0219	(0.0794)	-0.0171	(0.117)
White	0.0420	(0.158)	0.446*	(0.180)
8-10 years old (base group)				
11-12 years old	-0.0686	(0.100)	-0.231	(0.137)
13 years old	0.0320	(0.132)	-0.174	(0.207)
14-15 years old	0.234*	(0.107)	0.143	(0.174)
Urban (base group)				
Suburb	-0.127	(0.138)	0.198	(0.148)
Rural	-0.0531	(0.149)	0.00988	(0.195)
Non-smoke free homes	1.184**	(0.106)	1.702**	(0.159)
Mother(M) smokes	0.376**	(0.132)	0.464*	(0.214)
Father (F) smokes	0.218	(0.138)	0.521*	(0.241)
M Degree (base group)				
M Below degree	0.101	(0.183)	0.108	(0.550)
M NVQ3	-0.0231	(0.184)	0.954	(0.520)
M NVQ2/NVQ1	0.0904	(0.156)	0.348	(0.387)
M No qualification	0.120	(0.193)	0.712	(0.389)
F Degree (base group)				
F Below degree	0.000485	(0.159)	0.0950	(0.586)
F NVQ3	0.110	(0.189)	-0.503	(0.550)
F NVQ2/NVQ1	0.224	(0.184)	0.472	(0.492)
F No qualification	-0.00473	(0.212)	0.637	(0.484)
M age ≤35 years(base group)				
M age 36-45 years	0.272	(0.177)	0.205	(0.193)
M age≥46 years	0.219	(0.232)	-0.541	(0.316)
F age ≤35 years(base group)				
F age 36-45 years	-0.314	(0.219)	-0.0689	(0.258)
F age≥46 years	-0.419	(0.248)	-0.271	(0.347)
Parent single	0.144	(0.183)	0.305	(0.261)
M missing	0.632	(0.357)	1.419**	(0.548)
F missing	-0.102	(0.279)	0.539	(0.591)
Boost sample	0.379	(0.233)	1.473*	(0.579)
Constant	0.0909	(0.675)	-0.0874	(1.228)
Observations	1397		958	
R^2	0.198		0.349	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses.
Sample weights applied.

Table A3.5 Q75 RIF regression of Log cotinine

	High social class		Low social class	
	Coeff.	Std. error	Coeff.	Std. error
Log household income	-0.246**	(0.0868)	-0.173	(0.119)
Male	-0.0395	(0.108)	-0.307*	(0.143)
White	0.440*	(0.213)	0.0474	(0.235)
8-10 years old (base group)				
11-12 years old	-0.0941	(0.130)	-0.0602	(0.166)
13 years old	0.116	(0.191)	0.577*	(0.265)
14-15 years old	0.635**	(0.151)	1.092**	(0.222)
Urban (base group)				
Suburb	0.0183	(0.183)	0.0137	(0.195)
Rural	-0.132	(0.197)	-0.0606	(0.240)
Non-smoke free homes	2.090**	(0.180)	0.670**	(0.176)
Mother(M) smokes	1.051**	(0.231)	1.093**	(0.245)
Father (F) smokes	-0.213	(0.236)	0.610*	(0.285)
M Degree (base group)				
M Below degree	0.00955	(0.215)	0.106	(0.472)
M NVQ3	0.0532	(0.242)	0.865	(0.520)
M NVQ2/NVQ1	-0.138	(0.196)	0.122	(0.342)
M No qualification	-0.361	(0.261)	0.550	(0.373)
F Degree (base group)				
F Below degree	0.130	(0.203)	0.0116	(0.573)
F NVQ3	-0.0348	(0.216)	-0.336	(0.537)
F NVQ2/NVQ1	0.607*	(0.251)	0.474	(0.499)
F No qualification	0.482	(0.354)	0.312	(0.492)
M age ≤35 years(base group)				
M age 36-45 years	0.121	(0.242)	0.275	(0.246)
M age≥46 years	0.0368	(0.297)	0.305	(0.382)
F age ≤35 years(base group)				
F age 36-45 years	0.102	(0.302)	-0.209	(0.314)
F age≥46 years	-0.185	(0.339)	-0.564	(0.454)
Parent single	0.496*	(0.247)	0.581	(0.319)
M missing	-0.260	(0.480)	1.906**	(0.718)
F missing	0.0809	(0.361)	-0.0691	(0.587)
Boost sample	0.478	(0.339)	1.003	(0.577)
Constant	1.088	(1.011)	1.181	(1.251)
Observations	1397		958	
R^2	0.283		0.217	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Standard errors in parentheses.
Sample weights applied.

Chapter 4

Weather Shocks, Household Consumption and Nutritional Status of Disadvantaged Children in Vietnam

4.1 Introduction

The increasing frequency of occurrence and the devastating impact of weather shocks represent a growing concern globally, particularly in developing countries where the impact is further exacerbated by the lack of adequate infrastructures and facilities capable of mitigating the immediate impact or aftermaths of weather shocks (Kahn, 2005, UNISDR, 2011b). The enormous human and welfare losses associated with weather shocks are widely documented. For example, in 2011 alone, approximately 332 weather shocks where reported worldwide, affecting 244.7 million and killing over 30,000 with a total economic cost estimated at approximately 366.1 billion US dollars (Guha-Sapir et al., 2012).

In addition to the immediate impact, weather shocks often result in huge secondary public health crises resulting from the outbreak of diseases, the disruption of safe drinking water supply and sanitation, the displacement of families and the relocation of survivors into crowded rescue centres, exposing survivors to further health hazards (Watson et al., 2007).

Children are particularly vulnerable and approximately 30-50% of fatalities resulting from the immediate repercussions of weather shocks are reported to be children (UNISDR, 2011a). Furthermore, weather shocks have been implicated in long-term child health outcomes including higher morbidity and mortality amongst children long after they survive the immediate impact. For example, following extreme drought in Zimbabwe, exposed children experienced slower growth rates (Hoddinott and Kinsey, 2001), the 1997 forest fire in Southeast Asia resulted in higher infant and child mortality in Indonesia (Jayachandran, 2009), while the 1998 Hurricane Mitch affecting large parts of Central America was associated with an increase in the prevalence of wasting and malnutrition amongst affected children in Honduras and Nicaragua (Barrios et al., 2000).

There is a growing body of evidence showing links between child stature and future labour market achievements (Case and Paxson (2008), and references therein). Therefore, shocks which affect child physical development and growth are likely to have long-term economic consequences. For example, Alderman et al (2006) showed that in addition to childhood stunting, exposure to drought and civil war in early childhood resulted in lower educational

attainment in adulthood. Other studies have equally highlighted the long-term health and economic consequences of other forms of early childhood shock. Some examples include higher mortality rates amongst adults born during an economic downturn compared to those born during an economic boom (van den Berg et al., 2006); shorter height at age 20 amongst cohorts whose parents experienced income shocks resulting from a widespread destruction of vineyards in mid-19th century France (Banerjee et al., 2010); lower educational attainment and occupational status amongst adults born during the food crisis in Germany following World war II, compared to those born shortly before or after the crisis (Jürges, forthcoming).

Previous research on weather shocks and child health has focused mainly on the impact of single large-scale weather shocks on child health with fewer studies on the impact of smaller-scale weather shocks. Although the human and economic costs of smaller-scale weather shocks are likely to be lower compared to large-scale shocks, recurrent exposure to small-scale weather shocks is likely to have significant impacts on household welfare as well as on children's short- and long-term health outcomes. To the best of my knowledge only two studies have investigated the impact of small-scale weather shocks on child health. Pörtner (2010) showed using three rounds of the Guatemala Demographic and Health Surveys (DHS), that exposure to hydro-metrological disasters (storms, flooding, heavy rainfall, hurricanes and frost) has a negative impact on child's health. After controlling for area and time fixed effects, exposure to small-scale weather shocks in the past year was associated with lower nutritional status in children under 5 years of age (Pörtner, 2010). Similar findings were reported by Datar et al. (2011) in rural India. Using repeated cross-sections of the National Family Health Surveys (NFHS), Datar et al. (2011) showed that exposure to different small-scale weather shocks in the previous year reduced child height-for-age Z-score (HAZ-score) by approximately 0.12-0.15 standard deviations and increased the probability of reporting symptoms of acute illnesses by 9-18% (Datar et al., 2011). HAZ-scores are regarded as a long-run indicator of child nutritional status and are estimated by standardising child height using the median height of a well-nourished child of the same age and gender in a reference population (where the United States National Centre for Health Statistics (US NCHS) sample is used as the reference population). Low HAZ-scores are indicative of past disruptions to child nutritional status resulting from inadequate food nutrient intake and/or recurrent infections and illnesses. The HAZ-score is widely used as a proxy for child health and is an important determinant of child's future health outcomes. For example, childhood malnutrition and wasting (HAZ-score less that -2) is associated with higher morbidity and mortality in adulthood (Victora et al., 2008).

In addition to fatalities and injuries resulting from the direct repercussions of weather shocks, shock to household income and changes in parental behaviour such as investment decisions in child health represent possible mechanisms through which weather shocks affect child health. In developing countries, the immediate and long-term impact of weather shocks on household welfare is well documented. Significant reductions in both agricultural and non-agricultural wages have been reported several years after the occurrence of a natural disaster (Mueller and Quisumbing, 2010, Mueller and Osgood, 2009, Jayachandran, 2006, Thomas et al., 2010). Since child health is a function of a set of inputs such as food nutrients, time and resources invested in caring for the child (Behrman and Deolalikar, 1988, Grossman, 1972, Rosenzweig and Schultz, 1983), shocks to household income are likely to reduce the demand for these inputs, potentially making child health vulnerable. In addition, shocks to household income may increase the opportunity cost of parents time in caring for the child when the need to replenish lost income and for day-to-day subsistence supersedes the need to investment in child health. For example, Datar et al. (2011) showed that in addition to the impact on child's nutritional status, children exposed to small-scale weather shocks are less likely to have full age-appropriate immunization coverage. Similar findings are reported by Miller and Urdinola (2010) who show an association between weather-induced increases in coffee prices and a decline in the use of preventative care and vaccination services during the first year of a child's life. Furthermore, the need to generate extra income may result in children having to contribute to household income and an increase in the supply of child labour, further compromising child health outcomes (O'Donnell et al., 2002, Roggero et al., 2007).

Chapter 4 contributes to this literature by estimating the impact of small-scale weather shocks on both child health and household income[33]. This chapter differs from previous studies which have either estimated the impact of weather shocks on child health or on household income, by estimating the impact of small-scale weather shocks on both child health and on household income using the same sample. Thus, chapter 4 is able to explicitly demonstrate that the adverse impact of weather shocks on child health is mediated through a reduction in household income. It uses the 2006 and 2009 panels of the Vietnam Young Lives Surveys (VYLS), which consist of a pro-poor sample of children aged 4 and 12 years in 2006.

[33] Household per capita consumption and expenditure (PCCE) on all goods including food and non-food goods (excluding medical care expenditures) is used as a proxy for household income.

Consistent with other studies, a negative association is observed between small-scale weather shocks and child HAZ-scores as well as between household total (log) per capita consumption and expenditure (PCCE). The analysis is extended to assess the impact on the quantity (household total PCCE on food) and the quality (household PCCE on high-nutrient and energy-rich food) of dietary intake. No statistically significant difference is observed in total food consumption between exposed and unexposed households. However, the results suggest that exposed households are able to smooth consumption of total food by decreasing the consumption of high-nutrient food (fish, meat, fruits and vegetables) by approximately the same magnitude as their increase in the consumption of low-nutrient, high calorie food (rice and tubers). This is indicative of a fall in quality of households' food intake, thus, providing an explanation for the negative impact of small-scale weather shocks on child nutritional status.

Disadvantaged groups such as children living in poorer households have been shown to be more vulnerable to weather shocks (Datar et al., 2011, Hoddinott and Kinsey, 2001), therefore chapter 4 also investigates the extent to which differential impact of small-scale weather shocks on household PCCE explains differential impact on child HAZ-score.

The rest of chapter 4 is organized as follows: the conceptual framework and econometric models are outlined in sections 4.2 and 4.3, respectively. Section 4.4 provides a description the VYLS and variables included in econometric models. The results are presented in section 4.5 and section 4.6 concludes by summarizing the key findings of the chapter.

4.2 Conceptual Framework

Following the literature on the demand for child health[34], the conceptual framework adopted in this chapter relies on a model of child health production in which child health is embedded in a household utility function. Households are assumed to maximise a utility function at time t given as:

$$U_t = U(C_t, H_t; K_t) \qquad (1)$$

where C is a vector of goods consumed (health and non-health goods), H is a vector of home-produced commodities such as child health, and K is a vector of household characteristics which may affect utility. Households face two constraints in the production of commodities: a constraint imposed by the technology through which it combines goods to produce commodities (technological constraint) and a budget constraint which

[34] Some examples include Pitt and Rosenzweig (1985), Thomas et al. (1990), Alderman and Garcia (1994) Hoddinott and Kinsey (2001) and Behrman and Skoufias (2004)

determines the bundle of goods it can afford. Thus, the maximization problem facing households is subject to a budget constraint, households' technology and a child health production function.

The child health production function can be described as a function of a set of inputs which can be combined to produce child health. These inputs such as food nutrients, time and resources invested in caring for the child are demanded by parents because they affect parents' utility indirectly through their impact on child health. In this chapter, child HAZ-score or nutritional status is used as an indicator of child health. Child nutritional status is described as a function of a set of material and environmental inputs which affect child stature:

$$H_t = f(X_t, Z_t, M_t, P_t, K_t, \epsilon_i, \epsilon_h, \epsilon_c) \qquad (2)$$

where H is the child's HAZ-score, X is a vector of observable child characteristics such as age and gender which may affect growth rate, Z is household consumption and expenditure on food nutrients which captures food nutrient input, M is a vector of non-material inputs such as time invested in caring for the child, P is a vector of parental characteristics such as education and age which may affect the technology through which health inputs are combined, K is a vector of household characteristics capturing the health environment facing each child such as good sanitation and availability of safe drinking water, ϵ_i captures time-invariant unobserved child characteristics such as genetic predispositions which are uninfluenced by parental behaviours or preferences but which may affect child health and ϵ_h and ϵ_c are unobserved time-invariant household and community characteristics, respectively, which could also affect child health.

Weather shocks are often associated with economic and welfare losses, particularly in poor households already facing budget constraints and limited abilities to smooth consumption. This may in turn affect child's nutritional status through a reduction in household food consumption and expenditure. Household food consumption and expenditure, Z, is described as follows:

$$Z_t = f(D_{t_{-1}}, I_t, K_t, \epsilon_z) \qquad (3)$$

where $D_{t_{-1}}$ represents households' exposure to small-scale weather shocks between two time periods, t and t -1, I is household income (or total PCCE on all food and non-food goods) and ϵ_z are unobserved time-invariant household and community characteristics that could affect household food consumption and expenditure. Substituting equation (3) into

(2), yields a child health production function that includes households' exposure to small-scale weather shocks:

$$H_t = f\left(X_t, D_{t_{-1}}, I_t, M_t, P_t, K_t, \epsilon_i, \epsilon_h , \epsilon_c \right) \qquad (4)$$

4.3 Data and Variables

This chapter uses data from the Vietnam Young Lives Survey (VYLS), an ongoing longitudinal survey of children and households in Vietnam. The first survey was conducted in 2002 and has since followed children and their households for two further rounds in 2006 and 2009. The original sample consists of 2,000 children aged 6-18 months (the younger cohort) and 1,000 children aged between 7.5-8.5 years (the older cohort). Children were selected from 31 communities[35] within five provinces representative of five socioeconomic regions in Vietnam: Lao Cai (North-East region), Hung Yen (Red River Delta), Da Nang (City), Phu Yen (South Central Coast) and Ben Tre (Mekong River Delta). In line with the main aim of the VYLS, which is to track the dynamics of childhood poverty, an over-poor sampling strategy[36] was adopted in the selection of communities, resulting in a purposive over-sampling of poor communities. In each selected community, 150 children were randomly selected from a list of eligible households[37]. In households with more than one eligible child, one child was randomly selected. This chapter uses only the last two rounds of the survey (2006 and 2009) including both the younger and older cohorts. Round 1 was excluded because information on household food and non-food consumption and expenditure was not collected in 2002.

Self-reported household exposure to small-scale weather shocks were obtained from responses to a series of questions on exposure to small-scale hydro-meteorological weather events including droughts, excessive rainfall or floods, erosions, landslides, frosts and storms. For example, for exposure to droughts prior to round 2 (2006), households were asked 'Have you experienced drought in the last four years?' and in round 3 (2009), 'Have you experienced droughts since we last came to see you?'. Similar questions were asked on exposure to excessive rainfall/floods, erosions, landslides, frosts and storms. In this chapter, households are classed as 'exposed' if they responded 'yes' to these questions.

[35] A community is defined as having a local government, primary school, commune health centre, post office and market.
[36] Tuan et al. (2003) provides a detailed description of the sampling strategy.
[37] Eligibility of households was based on the presence of a child born between January 2001 and May 2002 (for the younger cohort) and between January 1994 and June 1995 (for the older cohort)

In each round, objective measures of child height was collected and age-standardized to a HAZ-score using the World Health Organisation (WHO) recommended US NCHS sample as the reference population. HAZ-scores above 3 and below – 5 are recoded as missing following WHO recommendations which consider HAZ scores outside this range implausible and likely to be due to measurement errors (WHO, 1995).

In rounds 2006 and 2009, the VYLS collected detailed information on household consumption and expenditure on a wide range of food and non-food goods. Household food consumption and expenditure (estimated at 2006 prices) comprise the sum of the value of all food goods bought or obtained from own stock /harvest or received as gift or food aid. Household consumption and expenditure on micronutrient-rich food is estimated as the sum of all high-nutrient food (including fish, meat, eggs, milk, fruits, vegetables, legumes, lentils and beans) consumed within households in the past two weeks[38]. Similarly, household consumption and expenditure on energy-rich food is estimated as the sum of all high-calorie food (including rice, pasta, bread, wheat, cereal, tubers and potato) consumed within households in the past two weeks.

The VYLS collects a wide range of child, parent and household characteristics which are used as controls for observable characteristics that may affect the probability of exposure to weather shocks and child nutritional status. Child characteristics include child's gender (male/female), age categories (younger/older cohort) and ethnicity (ethnic majority group (kinh)/ethnic minority groups); parents' (fathers' and mothers') characteristics include education categories (no education/primary/secondary/high school/degree), age group (≤35 years/ >35 years of age) mothers' religion (religion/no religion) and mothers' height (in centimetres). Controls for household characteristics include (log) household size, proportion of children bellow the age of 6 and access to safe drinking water and good sanitation. For equations (7) and (8), in addition to household characteristics, characteristics of the head of the household (age group, education categories and gender) are included as controls. The final sample across both rounds consists of a total of 4,772 children (2,639 from round 2 and 2,133 round 3).

[38] In this chapter, household monthly consumption and expenditure is estimated by multiplying the two weeks consumption values by 2.

4.4 Empirical Models

In the first instance, the empirical analysis adopted in this chapter investigates the impact of small-scale weather shocks on child nutritional status. The second part of the empirical analysis investigates the impact of small-scale weather shocks on household total PCCE. The aim of the second part is to explicitly demonstrate that weather-induced negative shocks to household consumption mediate the impact of small-scale weather shocks on child nutritional status. Finally, the analysis is extended to investigate possible differences in the impact of weather shocks between two groups of children defined by their household socioeconomic status: children living in households below and above the sample median household total PCCE.

4.4.1 The impact of small-scale weather shocks on child HAZ-score

To estimate the impact on child nutritional status, an estimable version of equation (4) is specified to allow the comparison of HAZ-scores of children exposed to small-scale weather shocks to those of unexposed children:

$$H_{ict} = \beta_0 + \delta_1 D_{ict_{-1}} + \beta_1 X_{ict} + \beta_2 P_{ict} + \beta_3 K_{ict} + \beta_4 I_{ict} + \gamma Y_t + u_{ict} \qquad (5)$$

$$u_{ict} \sim (0, \sigma_u^2)$$

where H_{ict} is the HAZ-score of the *i*th child living in community c observed at time t (2006 and 2009). $D_{ict_{-1}}$ indicates whether a household was exposed to any small-scale weather shock prior to rounds 2 (2006) and 3 (2009), i.e. $D_{ict_{-1}}$ equals 1 when a household reports experiencing any weather shock between 2002 (t-1) and 2006 (t) or between 2006 (t-1) and 2009 (t) and 0, otherwise.

X, *P* and *K* are vectors of child, parent and household characteristics respectively, *I* is household's monthly (log) total PCCE on all food and non-food goods, *Y* is a vector of survey year dummies which captures general time trends in child HAZ-score, and u_{ict} is the random error term.

The impact of weather shocks on child HAZ-score (captured by δ_1) estimated from equation (5) will be valid if exposure to weather shocks are randomly assigned. However, communities with higher incidence of small-scale weather shocks are likely to experience less economic growth/development and wealthier households are more likely to migrate from these communities. In addition, households residing in high-risk communities may

over time, adopt less risky work or labour strategies in order to minimise the potential impact of weather shocks. This may in turn result in lower average income or returns within these communities. Thus, households living within high-risk communities are likely to face greater constraints in investing in child health, resulting in lower child health outcomes. Failure to control for this will result in an overestimation of the impact of small-scale weather shocks on child health outcomes. A community fixed effect model is specified by decomposing the random error term in equation (5) into two components: $u_{ict} = \alpha_c + \varepsilon_{ict}$. This model controls for community time-invariant characteristics that may be associated with both child health and the probability of exposure to weather shocks:

$$H_{ict} = \beta_0 + \delta_2 D_{ict_{-1}} + \beta_1 X_{ict} + \beta_2 P_{ict} + \beta_3 K_{ict} + \beta_4 I_{ict} + \gamma Y_t + \alpha_c + \varepsilon_{ict} \qquad (6)$$

$$\varepsilon_{ict} \sim (0, \sigma_\varepsilon^2)$$

where α_c represents time-invariant community environment common to all children living within the same community and ε_{ict} is the random error term. Due to serial correlation in the random error term, standard errors are estimated to allow for arbitrary variance-covariance structure within communities.

The parameter, δ_2 is estimated using variations in exposure within communities and across time. In other words, equation (6) compares the HAZ-scores of children exposed to small-scale weather shocks to unexposed children within the same community. Identification of δ_2 relies on the assumption that amongst households with similar characteristics living within the same community, exposure to small-scale weather shock is uncorrelated with unobservable household characteristics that could affect child nutritional status. Failure to control for time invariant unobserved household characteristics that are correlated with both the probability of exposure and child nutritional status may result in biased estimates of the impact of small-scale weather shocks on child nutritional status. For example, households may report exposure to weather shocks depending on the extent to which they perceive a fall in household economic welfare, ex-post (Dercon, 2002). Thus, differences in exposure to weather shocks may reflect differences in households' level of preparedness or ability. Lower 'ability' households, for example, may possess lower adaptive or coping strategies, resulting in 'exposure' to weather shocks and lower 'ability' may also be associated with lower technical efficiency in the combination of child health inputs, resulting in lower child health outcomes. Failure to account for differences in household

'ability' could therefore, result in an overestimation of the impact of small-scale weather shocks.

Due to limitations imposed by the data[39], household fixed effects cannot be explicitly accounted for in this model. Nonetheless, the validity of the assumption that exposure to small-scale weather shocks is uncorrelated with unobservable household characteristics is verified by estimating δ_2, using an alternative specification of equation (6) which excludes parents' (*P*) and household (*K*) characteristics from equation (6). If small-scale weather shocks randomly affect households, inclusion of parents' and households' characteristics should not change the estimated effect of small-scale weather shocks on child nutritional status.

4.4.2 The impact of small-scale weather shocks on household consumption and expenditure

A fall in household total consumption and expenditure, particularly in food consumption is likely to explain the impact of small-scale weather shock on child nutritional status. To investigate this further, the impact of small-scale weather shocks on household total consumption and expenditure, on household food consumption and expenditure and on household food budget shares, are estimated. Similar to section 4.4.1, a series of community fixed effects models are specified. First, the impact on household (log) total PCCE is modelled controlling for household characteristics and characteristics of the head of household:

$$I_{ict} = \beta_0 + \rho D_{ict_{-1}} + \beta_1 Q_{ict} + \beta_2 K_{ict} + \gamma Y_t + \alpha_c + \varepsilon_{ict} \quad (7)$$

$$\varepsilon_{ict} \sim (0, \sigma_\varepsilon^2)$$

Second, the impact on household (log) food PCCE (and household food budget share[40]) is modelled, controlling for household total PCCE, household characteristics and characteristics of the head of household:

$$Z_{ict} = \beta_0 + \rho D_{ict_{-1}} + \varphi I_{ict} + \beta_1 Q_{ict} + \beta_2 K_{ict} + \gamma Y_t + \alpha_c + \varepsilon_{ict} \quad (8)$$

$$\varepsilon_{ict} \sim (0, \sigma_\varepsilon^2)$$

[39] The VYLS collects data on one child per household.
[40] Food budget share is estimated as the sum of households' consumption and expenditure on all food items divided by household total consumption and expenditure on food and non-food items.

where I_{ict} is the *i*th household's monthly (log) total PCCE on all food and non-food goods, Z_{ict} is household's monthly (log) food PCCE (or household budget share on food). Household monthly food consumption and expenditure constitute all food items obtained from three sources: either bought by the household or obtained from own stock/harvest or received as gift/food aid within the past four weeks. *K* is a vector of household characteristics including (log) household size, proportion of children less than 6 years old, access to safe drinking water and good sanitation (flush toilet/septic tanks), and Q is vector of the characteristics of the head of household including education, gender and age.

To investigate the impact of weather shocks on the quality of household dietary intake, (log) food PCCE is disaggregated into household consumption and expenditure on micronutrient-rich and energy-rich food. Micronutrient-rich foods are high-nutrient food, rich in trace minerals and vitamins but very low in calories. They are needed by the body in small quantities and are vital for maintaining healthy body functions and in reducing the risk of chronic infections. On the other hand, energy-rich food (carbohydrates, fat and proteins) constitute the major part of a standard diet and are high in calories but have very little micronutrient content. Equation (8) is estimated separately using (log) PCCE on micronutrient-rich food and energy-rich food as dependent variables.

4.4.3 Differential impact of small-scale weather shocks

The impact of small-scale weather shocks on child health may vary depending of households' capacity to cope with the shock ex post. For example, wealthier households exposed to weather shocks are less likely to experience reductions in absolute consumption if they have access to credit markets or possess assets which can be used to smooth consumption. On the other hand, poorer households often live in more risky environments and children from these households already experience very low levels of consumption and poorer health status, such that exposure to weather shocks may have little impact on child nutritional status. To assess the differential impact of weather shocks across socioeconomic groups, equation (6) – (8) is estimated separately for children living in households above and below the sample median household (log) total PCCE and the coefficients on $D_{ict_{-1}}$ for the two groups are compared. Furthermore, the extent to which differences in the impact of weather shocks on household consumption explains differences in the impact on child HAZ-scores is investigated.

4.5 Results and Discussion

Table 4.1a shows the summary statistics of the full sample and separately for children exposed and unexposed to small-scale weather shocks. Approximately 27% of children were exposed to small-scale weather shocks across both rounds with 40% of these shocks occurring between 2002 and 2006 and 60% between 2006 and 2009. On average children are 1.28 standard deviations shorter than children of the same age and gender within the US reference population, with those exposed to weather shocks statistically significantly shorter than children unexposed by approximately 0.2 standard deviations. Mean household (log) total PCCE, (log) food PCCE and food budget shares are lower in exposed households compared to unexposed households. Table 4.1a also shows differences in the quality of food consumed between exposed and unexposed households. Household (log) PCCE on micronutrient-rich food is lower while (log) PCCE on energy-rich food is higher in exposed household compared to unexposed households. Similar patterns are observed in households' allocation of the food budget to micronutrient-rich and energy-rich food[41]. The proportion of the food budget allocated to micronutrient-rich food is lower in exposed households compared to unexposed households (51% vs. 49%), while the proportion of the food budget allocated to energy-rich food is higher in exposed households compared to unexposed households (31% vs. 28%).

In Vietnam, agriculture constitutes a major, or in some households the only, source of income particularly for poorer households where approximately 75% of households in the lowest income quintile rely solely on agricultural income (Vietnam General Statistics Office, 2010). Therefore weather shocks such as droughts, heavy rainfalls or floods which adversely affect agricultural production are likely to have a negative impact on household income, particularly in poorer regions. A fall in household income will in turn affect the quality of food exposed households can afford to purchase, resulting in a shift from high-nutrient food to more affordable, but less quality food. This can be seen in Table 4.1b which shows households' PCCE on food (in Vietnamese Dong) obtained from three different sources: food bought/purchased, food consumed from own stock/harvest or food received as gifts. Purchased food constitutes the major part of households' total PCCE on food compared to food obtained from own stock/ harvest or received as gift.

[41] Budget shares for micronutrient-rich and energy-rich food are calculated as the sum of household consumption and expenditure on all individual micronutrient-rich and energy-rich food respectively, divided by total food consumption and expenditure.

Table 4.1a Summary statistics

	Full sample	Unexposed	Exposed	Difference
Child's Characteristics				
HAZ-score	-1.278	-1.232	-1.404	0.172**
Male	0.505	0.508	0.497	0.011
Kinh (Ethnic majority)	0.891	0.909	0.841	0.067**
Older cohort	0.340	0.328	0.371	-0.042**
Survey year				
2006	0.553	0.608	0.401	0.208**
2009	0.447	0.392	0.599	-0.208**
Father's (F) and Mother's (M) Characteristics				
F age≤35 years	0.393	0.408	0.351	0.057**
F age >35 years	0.607	0.592	0.649	-0.057**
M age ≤35 years	0.548	0.570	0.486	0.084**
M age>35 years	0.452	0.430	0.514	-0.084**
F No education	0.052	0.047	0.064	-0.017*
F Primary	0.231	0.209	0.290	-0.081**
F Secondary	0.461	0.460	0.462	-0.002
F High School	0.178	0.195	0.130	0.065**
F Degree	0.079	0.088	0.054	0.035**
M No education	0.079	0.070	0.102	-0.032**
M Primary	0.265	0.249	0.308	-0.058**
M Secondary	0.475	0.478	0.468	0.011
M High School	0.120	0.133	0.084	0.049**
M Degree	0.061	0.069	0.039	0.030**
M No religion	0.926	0.929	0.916	0.013
M Height (cm)	152.34	152.47	151.97	0.506**
Household Characteristics				
Urban residence	0.203	0.211	0.181	0.030*
Safe drinking water	0.694	0.712	0.644	0.069**
Good sanitation	0.385	0.403	0.336	0.067**
Log household size	1.627	1.622	1.640	-0.017+
Prop. of children≤6 years	0.132	0.141	0.106	0.035**
Log total PCCE	5.958	5.964	5.942	0.022
Log food PCCE	5.378	5.394	5.334	0.059**
Log energy-rich food PCCE	4.007	3.992	4.047	-0.055**
Log nutrient-rich food PCCE	4.642	4.669	4.567	0.101**
Budget share of food	0.590	0.595	0.576	0.018**
Energy-rich food share	0.289	0.281	0.310	-0.029**
Nutrient-rich food share	0.500	0.505	0.487	0.018**
Household Head (H) Characteristics				
H No education	0.058	0.052	0.074	-0.022**
H Primary	0.247	0.227	0.304	-0.076**
H Secondary	0.449	0.450	0.446	0.004
H High School	0.170	0.187	0.122	0.064**
H Degree	0.076	0.084	0.054	0.031**
H age ≤35 years	0.339	0.354	0.297	0.056**
H age >35 years	0.661	0.646	0.703	-0.056**
H Gender (Female)	0.092	0.096	0.080	0.015
PY Observations	4772	3504	1268	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.
PY: Person-years.

Table 4.1b Mean household food PCCE from three sources (in 1000 VND)

Sources of food consumed:	Full sample	Unexposed	Exposed	Difference
Total food				
Bought	207.35	212.66	192.69	19.97**
Own stock/harvest	38.40	37.09	40.67	-3.58*
Gift	5.65	5.92	4.90	1.02
Energy-rich food				
Bought	38.84	37.58	42.31	-4.74**
Own stock/harvest	21.73	21.72	21.75	-0.03
Gift	0.90	0.74	1.34	-0.60**
Nutrient-rich food				
Bought	111.10	115.27	99.57	15.69**
Own stock/harvest	15.61	14.67	18.22	-3.55**
Gift	2.77	2.97	2.19	0.78*
PY Observations	4772	3504	1268	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

PY: Person-years. VND: Vietnamese Dong

However, compared to unexposed households, exposed households consume more from own stock and purchase less food goods, suggesting higher budgetary constraints amongst exposed households and a higher reliance on own stock or harvest to meet dietary needs. Disaggregating total food PCCE into PCCE on energy- and micronutrient-rich food shows lower expenditure on high-nutrient food and higher expenditure on energy-rich food in exposed households compared to unexposed households. This suggests a shift from purchasing high-nutrient food to perhaps more affordable energy-rich food, by exposed households. Although consumption of high-nutrient food from own stock/harvest is higher in exposed households, this is not high enough to offset the lower expenditure on high-nutrient food.

Figures 4.1-4.3 presents a series of nonparametric locally weighted regressions (Cleveland, 1979), showing the impact of small-scale weather shocks on child HAZ-scores, on (log) food PCCE and on household food budget shares[42]. Figures 4.1-4.3 also plots densities of (log) total PCCE (Figures 4.1A, 4.2A-C and 4.3A-C) as well as densities of (log) food PCCE (Figures 4.1B-D). Figure 4.1A and 4.1B plots child HAZ-score as a function of (log) total PCCE and (log) food PCCE, splitting the sample by exposed and unexposed households. Both graphs show a positive relationship between household PCCE and child nutritional status, with child HAZ-scores increasing as household total PCCE and total food PCCE increases. However, across the entire PCCE distribution, the HAZ-scores of children exposed to small-scale weather shocks are lower than the HAZ-scores of

[42] These plots are obtained using the pooled sample across both years.

unexposed children. The gap between the 'exposed' and 'unexposed' lines is indicative of the magnitude of the impact of weather shocks on child nutritional status and a widening of the gap between the two lines going up the (log) total PCCE distribution is indicative of a differential impact of small-scale weather shocks at different quantiles. The gap between the two lines is greatest at higher quantiles, suggesting a greater impact of weather shocks on wealthier households[43]. A similar positive relationship is observed between child HAZ-score and the consumption of micronutrient-rich food (Figure 4.1D), while no clear relationship is observed between child nutritional status and PCCE on energy-rich food (Figure 4.1C), suggesting a more important role of micronutrient-rich food in predicting nutritional status.

Given the positive relationship between the quality of food and child HAZ-score, lower HAZ-scores in exposed children is likely to be due to a reduction in household food PCCE, particularly PCCE on high-nutrient food. Figure 4.2 and 4.3 shows differences between exposed and unexposed households in terms of household budget share on food (Figure 4.2A-C) and household (log) food PCCE (Figure 4.3A-C). Figure 4.2A-C plots food budget share as a function of (log) total PCCE, splitting the sample by exposed and unexposed households. A negative relationship is observed between (log) total PCCE and household food budget share (Figure 4.2A) suggesting that food is a necessity for both groups[44]. However, food budget share in exposed households is lower across the entire distribution of (log) total PCCE compared to the food budget shares of unexposed households.

Similarly, energy-rich food is a necessity for both groups (Figure 4.2B), but exposed household allocate a higher proportion of their food budget to energy-rich food compared to unexposed households. On the other hand, the share of the food budget allocated to high-nutrient food is lower in the exposed households compared to unexposed households (Figure 4.2C) suggesting a reduction in the quality of dietary intake. Similar results are observed when household (log) food PCCE is plotted as a function of (log) total PCCE (Figure 4.3A-C). Compared to unexposed households, exposed households consume less high-nutrient food and more energy-rich food.

[43] This effect is discussed in more detail in section 4.5.3 which examines the differential impact of small-scale weather shocks using parametric modelling.
[44] This implies that household PCCE on food grows more slowly than household total PCCE.

Figure 4.1 HAZ-scores and household consumption: exposed vs. unexposed households

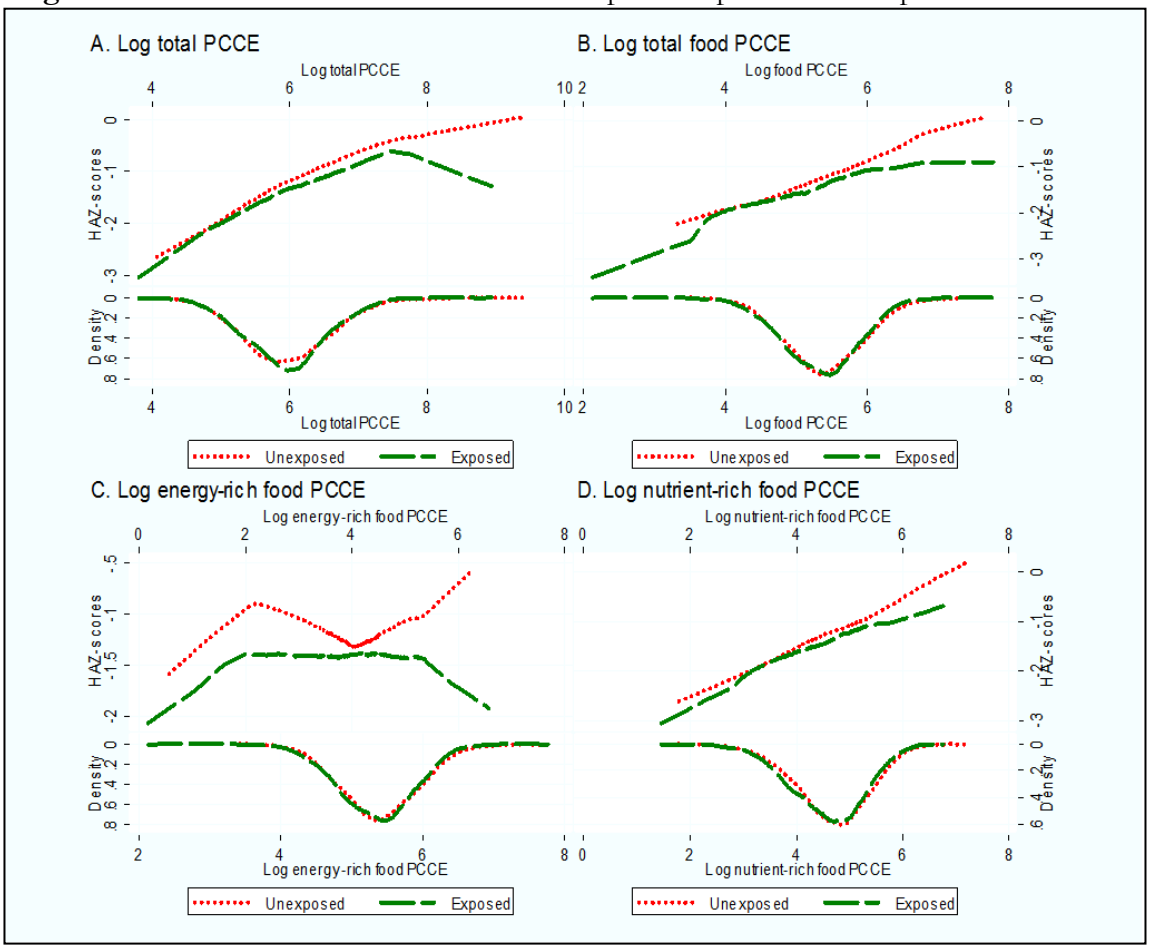
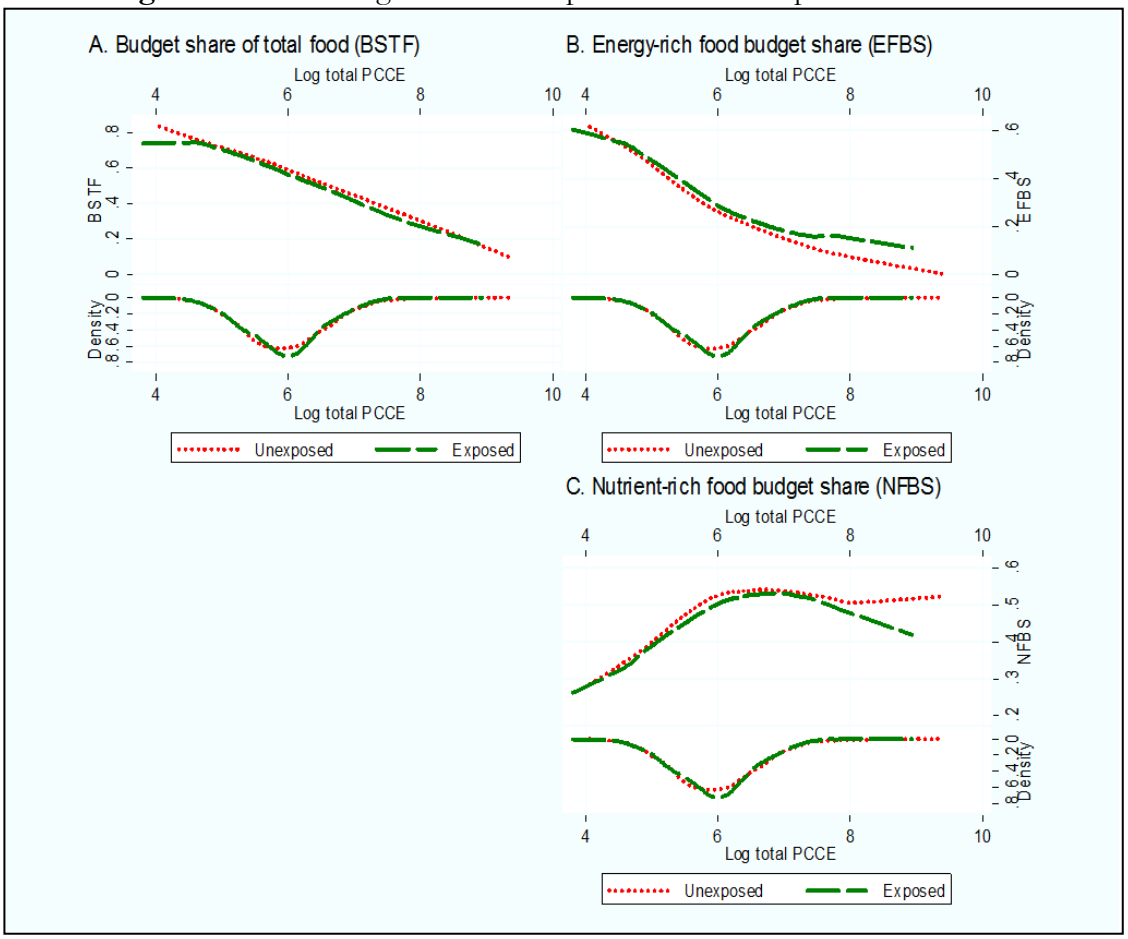


Figure 4.2 and 4.3 provides some explanation for the observed differential impact of small-scale weather shocks on child HAZ-scores at different quantiles of the (log) total PCCE distribution. The gap in household budget share and (log) PCCE on energy-rich food (Figure 4.2B and 4.3B, respectively) observed between the 'exposed' and 'unexposed' lines is widest at higher quantiles of the (log) total PCCE distribution compared to lower quantiles. Similarly, the gap in household budget share and (log) PCCE on micronutrient-rich food (Figure 4.2C and 4.3C), although smaller than the gap observed with energy-rich food, appears to be larger at the higher quantiles compared to lower quantiles. This suggests that the differential impact of small-scale weather shocks on child HAZ-scores may be mediated by the differential impact on the quality of dietary intake.

Figure 4.2 Food budget shares in exposed versus unexposed households

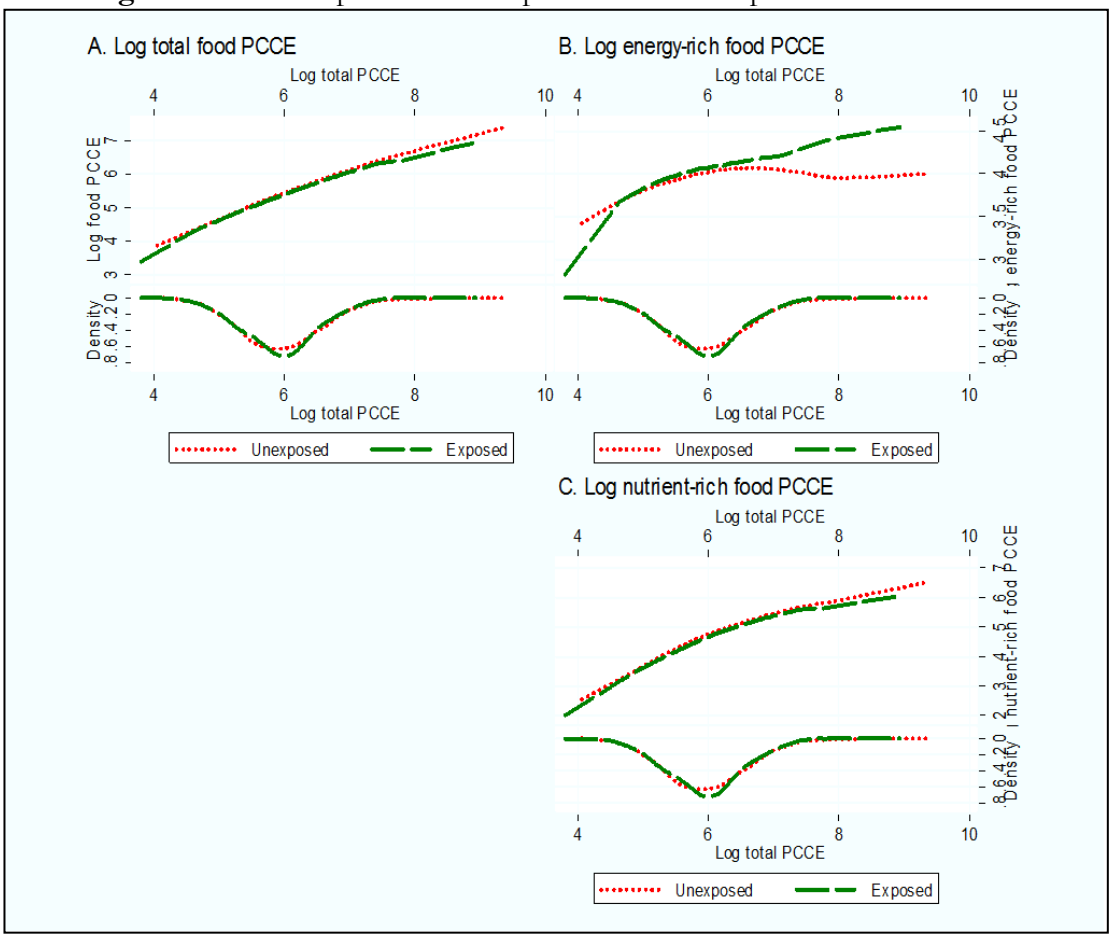


Differences in parent and household characteristics observed between exposed and unexposed children (Table 4.1a) suggests that exposure to small-scale weather shocks are not randomly distributed. On average, households exposed to small-scale weather shocks are of lower socioeconomic status compared to unexposed households (Table 4.1a). For example, compared to unexposed household, parents in exposed households are less educated, fewer proportions of households exposed to small-scale weather shocks have access to safe drinking water (64% versus 71%), flush toilet/septic tank (34% versus 40%) and have larger (log) household sizes (1.64 versus 1.62). Parental education, access to safe drinking water, and good sanitary conditions has been shown to be important determinants of child health and nutritional status in developing countries (Behrman and Deolalikar, 1988, Fewtrell et al., 2005). For example, higher education is associated with higher income and children with more educated parents are likely to have better health outcomes compared to children with less educated parents[45]. The following sub-sections present

[45] Currie (2009) provides an excellent review of the literature on child health and parental socioeconomic status.

results of the parametric analyses that accounts for differences in observable characteristics between exposed and unexposed children.

Figure 4.3 Food expenditure in exposed versus unexposed households



4.5.1 Impact of small-scale weather shocks on child nutritional status

The parametric strategy adopted here estimates the impact of small-scale weather shocks on child HAZ-scores controlling for observable characteristics correlated with both exposure to small-scale weather shocks and child HAZ-scores. Exposure to small-scale weather shocks is assumed to be randomly distributed amongst households living within the same community, conditional on observable parent and household characteristics. Two community fixed effect models are specified to test this assumption; the first excludes parent and household characteristics and the second accounts for parent and household characteristics. Table 4.2a shows results of the parametric estimation of the impact of small-scale weather shocks on child HAZ-scores. In the first specification (first column, Table 4.2a), after controlling for only child characteristics, HAZ-score of children exposed to small-scale weather shocks are on average approximately 0.15 standard deviations lower

than those of unexposed children. This finding is consistent with other studies that have reported similar estimates of the impact of small-scale weather shocks on child nutritional status (Datar et al., 2011, Pörtner, 2010).

An indication of the magnitude of this impact can be deduced using the World Health Organisation (WHO) growth reference charts[46] which shows corresponding height differences (comparable to the US NCHS sample) for a given age and gender for a one standard deviation in HAZ-scores. For example, a one standard deviation in the HAZ-score of a 4 year old male child is equivalent to a 4.25cm difference in height. Similarly, for a male child aged 8, 12 and 16 year old, the equivalent height difference is approximately 6.25cm, 6.75cm and 7.5cm, respectively. Equivalent approximations for girls of similar ages (i.e. 4-16years old) are 4cm, 6cm, 6.5cm and 7cm, respectively. Therefore a reduction in HAZ-scores by 0.15 standard deviation is equivalent to a reduction in height by approximately 0.6 -1.125cm. Given the potential for future catch-up growth in children who experience temporary growth retardation in childhood (Adair, 1999)[47], the functionally small height differences estimated here may disappear in late childhood/adolescence. However, repeated exposure to shocks, such as small-scale weather shocks, throughout childhood is likely to impede any catch-up growth that may have occurred in late childhood or early adolescence (Martorell et al., 1994).

In the second specification (second column, Table 4.2a), a set of observable household characteristics are included while the third specification includes the full set of child, household (excluding household log total PCCE) and parent characteristics (third column, Table 4.2a). If small-scale weather shocks randomly affect households, the inclusion of observable household and parent characteristics should have little or no effect on the magnitude of the estimated impact of small-scale weather shocks on child HAZ-score. However, after controlling for household characteristics-household size, access to safe drinking water and good sanitation, the magnitude of the estimated effect of small-scale weather shocks on child HAZ-scores reduces slightly to approximately 0.13 standard deviations. Similarly, the inclusion of parent characteristics further reduces the magnitude of the estimated impact.

[46] These charts are downloadable from the WHO webpage:
http://www.who.int/growthref/who2007_height_for_age/en/index.html
[47] The potential for future catch-up growth in children has been disputed by some authors. Some examples include Martorell et al. (1994) and Hoddinott and Kinsey (2001).

This reduction in the magnitude of the impact of small-scale weather shocks suggest that small-scale weather shocks are disproportionately distributed amongst those whose observed characteristics are correlated with a higher probability of malnourishment. For example, consistent with the literature on child health and parental education, higher parental education is associated with higher child HAZ-scores (Table 4.2a). Since exposed children, on average, have parents with lower levels of education (Table 4.1a), the reduction in the magnitude of the estimated impact of small-scale weather shocks after controlling for parents education suggests that part of this impact can be explained by the impact of parent education on child HAZ-scores.

Table 4.2a Impact of small-scale weather shocks on child HAZ-score

	(1)		(2)		(3)	
Exposed	-0.146**	(0.0439)	-0.125**	(0.0393)	-0.0875*	(0.0364)
Male	-0.0352	(0.0332)	-0.0407	(0.0346)	-0.0278	(0.0352)
Kinh (Ethnic majority)	0.576**	(0.107)	0.515**	(0.102)	0.277**	(0.0813)
Older cohort	-0.208**	(0.0439)	-0.248**	(0.0470)	-0.207**	(0.0407)
Year=2009	0.189**	(0.0433)	0.138**	(0.0410)	0.162**	(0.0397)
Safe drinking water			-0.0224	(0.0421)	-0.0193	(0.0365)
Good sanitation			0.247**	(0.0514)	0.145**	(0.0416)
Log household size			-0.116	(0.0737)	-0.0461	(0.0709)
Prop. of children≤6 years			-0.202	(0.125)	-0.181+	(0.105)
Urban residence			0.304	(0.225)	0.194	(0.321)
M No religion					0.00868	(0.0657)
M Height (cm)					0.0535**	(0.00267)
F age >35 years					-0.0462	(0.0559)
F Primary					0.0700	(0.0931)
F Secondary					0.162	(0.102)
F High School					0.165	(0.128)
F Degree					0.371**	(0.132)
M age>35 years					-0.0488	(0.0438)
M Primary					0.110	(0.0736)
M Secondary					0.149*	(0.0722)
M High School					0.270**	(0.0791)
M Degree					0.392**	(0.0981)
Constant	-1.749**	(0.104)	-1.586**	(0.167)	-9.880**	(0.443)
PY Observations	4772		4772		4772	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Cluster-robust standard errors in parentheses.

Base categories: F and M ≤35years for parents' age, F and M with no education for parents' education. PY: Person-years. All models control for community fixed effects.

Table 4.2b Impact of small-scale weather shocks on child HAZ scores, controlling for Total PCCE and Food PCCE

	(1)		(2)		(3)		(4)	
Exposed	-0.0775*	(0.0387)	-0.0793*	(0.0359)	-0.0891*	(0.0364)	-0.0761*	(0.0357)
Log total PCCE	0.174**		-		-		-	
Log food PCCE	-		0.143**	(0.0324)	-		-	
Log energy-rich food PCCE	-		-		0.0483+	(0.0282)	-	
Log nutrient-rich food PCCE	-		-		-		0.117**	(0.0219)
Male	-0.0311	(0.0347)	-0.0281	(0.0351)	-0.0271	(0.0353)	-0.0272	(0.0352)
Kinh (Ethnic majority)	0.242**	(0.0818)	0.260**	(0.0829)	0.275**	(0.0824)	0.256**	(0.0808)
Older cohort	-0.201**	(0.0399)	-0.209**	(0.0404)	-0.210**	(0.0412)	-0.208**	(0.0404)
Year=2009	0.0997*	(0.0397)	0.138**	(0.0405)	0.164**	(0.0401)	0.137**	(0.0405)
Safe drinking water	-0.0303	(0.0371)	-0.0302	(0.0367)	-0.0194	(0.0367)	-0.0297	(0.0361)
Good sanitation	0.106*	(0.0427)	0.120**	(0.0411)	0.144**	(0.0418)	0.115**	(0.0417)
Log household size	0.00287	(0.0687)	0.0495	(0.0715)	-0.0170	(0.0733)	0.0341	(0.0690)
Prop. of children≤6 years	-0.0988	(0.112)	-0.168	(0.104)	-0.163	(0.106)	-0.198+	(0.102)
Urban residence	0.169	(0.324)	0.192	(0.329)	0.197	(0.323)	0.189	(0.328)
M No religion	0.0110	(0.0658)	0.0115	(0.0663)	0.00993	(0.0663)	0.00833	(0.0652)
M Height (cm)	0.0525**	(0.00264)	0.0530**	(0.00259)	0.0533**	(0.00267)	0.0531**	(0.00265)
F age >35 years	-0.0334	(0.0559)	-0.0458	(0.0561)	-0.0487	(0.0558)	-0.0455	(0.0561)
F Primary	0.0425	(0.0915)	0.0514	(0.0929)	0.0672	(0.0932)	0.0475	(0.0914)
F Secondary	0.120	(0.0996)	0.135	(0.101)	0.159	(0.102)	0.129	(0.100)
F High School	0.113	(0.126)	0.133	(0.128)	0.161	(0.128)	0.125	(0.127)
F Degree	0.290*	(0.131)	0.314*	(0.131)	0.366**	(0.131)	0.307*	(0.133)
M age>35 years	-0.0467	(0.0436)	-0.0482	(0.0432)	-0.0520	(0.0438)	-0.0441	(0.0434)
M Primary	0.100	(0.0731)	0.110	(0.0735)	0.112	(0.0742)	0.106	(0.0736)
M Secondary	0.123	(0.0742)	0.136+	(0.0722)	0.151*	(0.0725)	0.133+	(0.0723)
M High School	0.224**	(0.0781)	0.248**	(0.0776)	0.274**	(0.0790)	0.240**	(0.0782)
M Degree	0.325**	(0.100)	0.358**	(0.0984)	0.395**	(0.0975)	0.348**	(0.0978)
Constant	-10.72**	(0.488)	-10.66**	(0.479)	-10.10**	(0.454)	-10.40**	(0.411)
PY Observations	4772		4772		4772		4772	

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Cluster-robust standard errors in parentheses. All models control for community fixed effects.

Table 4.2b extends from Table 4.2a and shows results of the impact of weather shocks on child HAZ-scores that is mediated through household PCCE. After controlling for (log) total PCCE, the magnitude of the impact of weather shocks on child HAZ-scores reduces from approximately 0.09 (Table 4.2a, third column) to 0.08 (Table 4.2b, first column). This suggests that difference in (log) total PCCE between exposed and unexposed households explains approximately 11% of the impact of weather shocks on child HAZ-score. As shown in sub-section 4.5.2 below (Table 4.3), weather shocks negatively affects (log) total PCCE, resulting in lower (log) total PCCE amongst exposed households. Therefore, the impact of weather shocks on child HAZ-score can be explained, at least in part, by the negative impact of weather shocks on household log total PCCE.

Similar to household (log) total PCCE, after controlling for households' (log) food PCCE, the magnitude of the impact of weather shocks on child HAZ-scores reduces from approximately 0.09 (in Table 4.2a, third column) to 0.08 (Table 4.2b, second column).

Disaggregating total food PCCE into PCCE on micronutrient- and energy-rich food suggest that the mediating effect of weather shocks on child HAZ-score is mainly driven by lower micronutrient-rich PCCE in exposed households. After controlling for energy-rich food PCCE, the magnitude of the impact of weather shocks on child HAZ-score remain largely unchanged (Table 4.2a, third column vs. Table 4.2b, third column), while the magnitude of the impact on HAZ-score reduces by approximately 11% after controlling for micronutrient-rich food PCCE (Table 4.2a, third column vs. Table 4.2b, fourth column). This is unsurprising given that the magnitude of the direct effect of micronutrient-rich food on child HAZ-score is much higher than the magnitude of the effect of energy-rich food (0.12 vs. 0.05), suggesting that micronutrient food is a stronger determinant of child HAZ-scores compared to energy-rich food PCCE. Therefore lower PCCE on micronutrient food in exposed households is likely to explain part of the lower HAZ-scores of children living in exposed households.

Overall, after controlling for the full set of child, household and parent characteristics, the HAZ-scores of children exposed to weather shocks are statistically significantly lower (by approximately 0.08) compared to HAZ-scores of unexposed children (Table 4.2b, first column).

4.5.2 Impact of small-scale weather shocks on household consumption

The second part of the empirical analysis estimates the impact of small-scale weather shocks on household consumption and expenditure. Consistent with the literature on child health and parental socioeconomic status (Currie, 2009, Cameron and Williams, 2009, Case et al., 2002) a significant positive correlation is observed between child HAZ-scores and household (log) total PCCE. An increase in household (log) total PCCE is associated with an increase in child HAZ-scores (Table 4.2b). Similar effects on child HAZ-scores are observed with (log) food PCCE and (log) micronutrient-rich food PCCE (Table 4.2b). Although energy-rich food has a positive effect on child HAZ-scores, the magnitude of this effect is considerably less and estimated with less precision, compared to the effect of micronutrient-rich food on child HAZ-scores (Table 4.2b, columns 3 and 4). The greater effect of high-nutrient food on child nutritional status is unsurprising given that micronutrients are more important for maintaining normal body physiological functions and micronutrient deficiencies could result in higher rates of infection and stunting as well as higher mortality rates in children (Black et al., 2008, Dewey and Begum, 2011).

Table 4.3 shows that exposure to small-scale weather shocks is associated with a statistically significant reduction in (log) total PCCE as well as a reduction in (log) food PCCE (although statistically insignificant). Differences in the magnitudes of the impact on household (log) total PCCE and on (log) food PCCE (9 vs. 2 percent) suggest that although small-scale weather shocks are associated with a reduction in the overall consumption, exposed households appear to protect food consumption.

Disaggregating food consumption into consumption of micronutrient- and energy-rich foods shows that exposed households are able to protect total food consumption by reducing the consumption of high-nutrient food and increasing the consumption of energy-rich food. Similar results are observed with the impact of weather shocks on household allocation of the food budget to high-nutrient and energy-rich food (Table 4.3, column 5-7). The concomitant decrease and increase in high-nutrient and energy-rich food respectively, is indicative of poorer dietary intake, thus providing a strong explanation for lower HAZ-scores in children exposed to small-scale weather shocks.

Table 4.3 Impact of weather shocks on household consumption and budget share on food

	Log Per Capita Consumption& Expenditure on:				Budget Share on food:		
Explanatory Variables	Total PCCE	Total Food	Energy-rich	Nutrient-rich	Total Food	Energy-rich	Nutrient-rich
Exposed	-0.0907**	-0.0161	0.0549*	-0.0569**	-0.00311	0.0213**	-0.0137**
	(0.0194)	(0.0154)	(0.0232)	(0.0203)	(0.00746)	(0.00602)	(0.00498)
Log total PCCE	-	0.753**	0.235**	0.880**	-0.107**	-0.127**	0.0652**
	-	(0.0195)	(0.0273)	(0.0292)	(0.00735)	(0.00693)	(0.00865)
H Primary[‡‡]	0.247**	0.0217	0.0286	0.0641+	0.00223	-0.0131	0.0110
	(0.0316)	(0.0218)	(0.0432)	(0.0330)	(0.0119)	(0.0103)	(0.00876)
H Secondary	0.419**	0.00859	0.00245	0.0778*	-0.00728	-0.0184+	0.0220*
	(0.0355)	(0.0209)	(0.0422)	(0.0316)	(0.0117)	(0.0106)	(0.00947)
H High School	0.536**	-0.00735	-0.0384	0.0961**	-0.0207	-0.0265*	0.0352**
	(0.0351)	(0.0229)	(0.0461)	(0.0306)	(0.0124)	(0.0118)	(0.00981)
H Degree	0.790**	0.0158	-0.0788	0.149**	-0.0123	-0.0310*	0.0523**
	(0.0503)	(0.0278)	(0.0479)	(0.0390)	(0.0135)	(0.0116)	(0.0119)
H age >35 years[‡]	-0.0466+	0.105**	0.139**	0.1000**	0.0450**	0.00584+	-0.00150
	(0.0233)	(0.0158)	(0.0217)	(0.0166)	(0.00676)	(0.00332)	(0.00312)
H Gender (Female)	0.0364	-0.0144	-0.0552+	-0.0318	-0.00778	-0.00350	-0.00248
	(0.0363)	(0.0181)	(0.0278)	(0.0239)	(0.00871)	(0.00600)	(0.00675)
Prop. of children≤6years	-0.373**	0.262**	-0.393**	0.574**	0.145**	-0.160**	0.153**
	(0.0537)	(0.0446)	(0.0562)	(0.0613)	(0.0191)	(0.0123)	(0.0174)
Log household size	-0.302**	-0.473**	-0.555**	-0.462**	-0.214**	-0.0137*	0.00654
	(0.0335)	(0.0227)	(0.0292)	(0.0236)	(0.00992)	(0.00579)	(0.00618)
Urban residence	0.260	-0.0874+	-0.0978	-0.0546	-0.0550**	-0.000794	0.00941
	(0.160)	(0.0454)	(0.0696)	(0.0750)	(0.0156)	(0.0220)	(0.0199)
Year=2009	0.394**	-0.104**	-0.141**	-0.100**	-0.0705**	-0.0131*	-0.00684
	(0.0231)	(0.0160)	(0.0242)	(0.0211)	(0.00835)	(0.00605)	(0.00874)
Constant	5.922**	1.620**	3.546**	0.00283	1.580**	1.103**	0.0644
	(0.0670)	(0.130)	(0.180)	(0.180)	(0.0502)	(0.0434)	(0.0511)
PY Observations	4772	4772	4772	4772	4772	4772	4772

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Cluster-robust standard errors in parentheses. Base categories: [‡]H ≤35years, [‡‡] H no education. PY: Person-years.

4.5.3 Differential impact of small-scale weather shocks

The nonparametric analyses discussed above suggest that the impact of small-scale weather shocks is greater amongst children living in wealthier households. In this section, this effect is further investigated by modelling the impact of small-scale weather shocks using two sub-groups defined by household (log) total PCCE: children living in households below the sample median and children living above the sample median (log) total PCCE (where sample median total (log) PCCE equals 5.94)

Similar to the results seen with nonparametric modelling, small-scale weather shocks has a higher impact on the stature of children living in wealthier households compared to children from poorer households (Table 4.4, column 1). Table 4.4 (columns 2-6) also provides some explanation for this observed differential impact and shows that the impact of weather shocks on household (log) total PCCE is greater in households above the median compared to household below the median (approximately 8 percent vs. 2 percent). Although the impact on the quantity of food consumed (log food PCCE) is approximately similar for both sub-groups, the impact on the quality of food is greater in households above the sample median. In both sub-groups, exposure to weather shocks is associated with approximately similar magnitudes of reduction in the consumption of high-nutrient food. However, compared to households below the median, in households above the median, exposure to weather shocks is associated with a higher increase in the consumption of energy-rich food (7 percent vs. 3 percent).

As alluded to in section 4.3.3, the impact of small-scale weather shocks may depend on households coping strategies or the availability of credit or savings to smooth consumption. This means that wealthier households may be better equipped to cope with the aftermath of small-scale weather shocks, thereby, resulting in a lower impact on child nutritional status. For example, Hoddinott and Kinsey (2001) showed that exposure to droughts adversely affects the height of only children living in poorer households possessing fewer livestock holdings, with no significant effect on the height of children living in wealthier households. However our results suggest that wealthier households are not better able to compensate the adverse effects of small-scale weather shocks and experience greater reductions in household (log) total PCCE as well as in the quality of dietary intake and hence, lower child HAZ-scores.

This result should be interpreted with caution given the pro-poor sampling strategy adopted by the VYLS resulting in a sample largely consisting of poorer households

compared to a nationally representative sample[48]. For example, in a nationally representative sample such as the Vietnam Household Living Standard Survey (VHLSS), average household monthly total PCCE in 2006, 2008 and 2010[49] was estimated at 511, 792 and 1,211 thousand Vietnamese Dong (VND), respectively (Vietnam General Statistics Office, 2010). Estimates from this study sample are 546 and 817 thousand VND in 2006 and 2009 respectively for household above the sample median (log) PCCE and 207 and 327 thousand VND, respectively for household below the median. Thus, 'wealthier' households within the VYLS are unlikely to be true representatives of an average (in terms of a national average) rich household and therefore may not truly reflect the ability of wealthier households to smooth consumption following weather-induced income shocks.

On the other hand, failure to observe a significant impact of small-scale weather shocks on the nutritional status of children from poorer households within this study, may be a reflection of the wider adverse environmental and living conditions to which children from poorer households are exposed, which may in turn, pose more risks to child health.

Taken together, this may explain the higher impact of small-scale weather shocks on children living in 'wealthier' households compared to those living in poorer households. Although direct extrapolations on the heterogeneity of the impact of small-scale weather shocks cannot be made to a nationally representative sample, this result further strengthens the main findings of this chapter on the mechanisms through which small-scale weather shocks affect child HAZ-scores.

[48] Nguyen (2008) compares the VYLS to two nationally representative samples: the VHLSS and the Demographic Health Survey (DHS). Average household wealth index is significantly lower in the VYLS compared to the VHLSS and the DHS (Nguyen, 2008).
[49] No survey was conducted in 2009.

Table 4.4 Differential Impact of small-scale weather shocks on child HAZ-scores, household consumption and budget share on food

		Log Per capita Consumption& Expenditure on:				Budget share on:		
	HAZ-scores	Total PCCE	Total Food	Energy-rich Food	Nutrient-rich Food	Total Food	Energy-rich Food	Nutrient-rich Food
Below sample median log total PCCE (*N=2283*)	-0.0324	-0.0214	-0.0186	0.034	-0.0677**	-0.00584	0.0217**	-0.0166*
	(0.0431)	(0.0144)	(0.0137)	(0.0232)	(0.0212)	(0.00765)	(0.00688)	(0.00684)
Above sample median log total PCCE (*N=2486*)	-0.126*	-0.0800**	-0.0255	0.0653+	-0.0590+	-0.00397	0.0230**	-0.0111
	(0.0527)	(0.0127)	(0.0238)	(0.0337)	(0.0293)	(0.0104)	(0.00791)	(0.00665)

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$; Cluster-robust standard errors in parentheses. Sample median log total PCCE=5.94

Column 1 includes the full set of child, parent and household characteristics and community fixed effects. Columns 2-8 include the full set of household characteristics, characteristics of the head of household and community fixed effects.

4.6 Conclusion

Small-scale weather shocks are unlikely to be accompanied by an immediate or extensive loss of lives or wide-spread destruction of social infrastructures often seen in cases of large-scale natural disasters. However, repeated exposure to small-scale weather shocks is likely to have important long-term adverse economic consequences on households, potentially making child health vulnerable.

This chapter uses the Vietnam Young Lives Survey to investigate the impact of small-scale weather shocks on household welfare and on the nutritional status of children living in households within disadvantaged communities of Vietnam. Similar to previous studies, chapter 4 confirms that both child health and household welfare are compromised by weather shocks. After controlling for a wide range of observed characteristics and community fixed effects, children living in households exposed to weather shocks are statistically significantly more malnourished than children living in unexposed households. Chapter 4 focuses mainly on investigating the role weather-induced shocks to household income (using household total PCCE as a proxy for household income) plays in mediating the adverse impact of small-scale weather shocks on child health. This is because shocks to household income is likely to represent a major mechanism through which small-scale weather shocks affects child health, either directly (for example, through a reduction in household consumption of high-nutrient food which are crucial for child physical development), or indirectly (for example, through inducing changes in parental behaviour such as reducing the time spent caring for the child or increasing the supply of child labour).

Chapter 4 shows that small-scale weather shocks are associated with a reduction in total household consumption and expenditure on all food and non-food goods as well as a reduction in the quality of dietary intake. Since child nutritional status is a function of the quality of dietary intake, this result shows that weather shocks impacts on child nutritional status through a reduction in household economic welfare which in turn, results in a fall in the quality of child dietary intake.

Finally, chapter 4 shows evidence of a differential impact of weather shocks in children from different socioeconomic backgrounds. However, contrary to previous studies, a greater impact of small-scale weather shock is observed in children from wealthier backgrounds due to a larger reduction in quality of household dietary intake. However, this

result should be interpreted with caution given that this study sample consists predominantly of disadvantaged households. Therefore 'wealthier' households are likely to still face high budget constraints with limited capacities to smooth consumption in response to household consumption shocks. On the other hand, for poorer households, weather shocks may be one of a host of other risky environmental conditions which are perhaps more important in explaining the nutritional status of children living within these environments. This may explain the failure to observe a significant impact of weather shocks in children from poorer backgrounds.

Chapter 5

Conclusion

This thesis contributes broadly to the empirical literature on the socioeconomic determinants of child health outcomes and child/adolescent smoking behaviours.

Chapters 2 and 3 focus on socioeconomic determinants of child and adolescent smoking behaviours. They contribute to the literature on socioeconomic inequalities in smoking amongst children and adolescents, an area that has received very little attention in the economics literature. Chapter 2 highlights the implications of misclassification errors in adolescent smoking participation models. Divergent conclusions have often been reported in previous empirical studies on the impact of different determinants of adolescent smoking. Variations in the extent of misclassification errors in adolescent self-reported smoking participation may partly be responsible. Using saliva cotinine (an objective measure of smoking participation) and a cut-point of 12ng/ml to define cotinine-validated smokers, a 'true' or cotinine-validated model of adolescent smoking participation is specified.

A comparison of estimates of the impact of parental income from the cotinine-validated smoking model to estimates from a self-reported smoking participation model shows an underestimation of the impact of parental income on the probability of smoking participation due to under-reporting of adolescent smoking participation. After controlling for a full set of adolescent and parent's observable characteristics, a 1% increase in parental income is associated with a statistically significant lower probability of being a smoker in the cotinine-validated smoking model. This effect is functionally smaller and statistically insignificant in the self-reported smoking model. Finally, the results show that under-reporting of adolescent smoking participation also results in an underestimation of concentration indices measuring income-related inequalities in adolescent smoking.

Although the main aim of this chapter is to highlight the possible role the under-reporting of smoking participation may have in explaining previous divergent findings on the relationship between parent's socioeconomic status and adolescent smoking, the findings of this chapter could be extrapolated to other determinants of adolescent smoking including tobacco price increases. For example, misclassification errors may play a role in explaining some of the conflicting conclusions often reported on the price elasticity of

adolescent smoking. This chapter highlights the importance of accounting for misclassification errors in empirical studies of adolescent self-reported smoking participation.

Chapter 3 builds on the theme of the preceding chapter by estimating the contributions of observable characteristics to socioeconomic differences in child/adolescent smoking but extends the analysis to both passive and active smoking. It takes advantage of the entire distribution of saliva cotinine and decomposes differences in log cotinine between children/adolescents from high and low socioeconomic backgrounds (the two study groups). Three decomposition methods are applied: a mean-based decomposition method and two methods that allow the decomposition of quantiles. Lower quantiles of the log cotinine distribution are likely to comprise of passive smokers while upper quantiles of active smokers. Therefore, by extending the decomposition analysis to different quantiles, chapter 3 provides useful insight into the contributions of observable characteristics to socioeconomic differences in both active and passive smoking. Differences in saliva log cotinine is decomposed into a part attributable to group differences in the distribution of characteristics (composition effect) and a part attributable to group differences in the impact of these characteristics (structural effects).

The results show that the composition effect accounts for a larger proportion of the total difference between the two groups both at the mean and at different quantiles of the log cotinine distribution. At lower quantiles, the composition effect explains more of the difference in log cotinine and less of the difference at higher quartiles. Observable characteristics making the largest contribution to the total composition effect include smoking within the home, father's education, household income and mother's smoking. These characteristics make different contributions at different quantiles, suggesting distinct roles in explaining socioeconomic differences in passive and/or active smoking. For example, smoking within the home explains more of the socioeconomic difference in passive smoking compared to its contribution to active smoking while household income and parental smoking explains more of the socioeconomic difference in active smoking compared to passive smoking.

Structural effects account for a proportion of the total difference between the two groups, albeit to a smaller extent compared to the composition effect. The structural effect attributable to smoking within homes at lower quartiles and parental smoking at higher quartiles suggest behavioural differences amongst parents from different social classes which perhaps, limit the impact of these characteristics on child passive and active smoking

participation. These finding suggests that promoting greater awareness of the negative externalities of parental tobacco consumption, particularly amongst parents from lower socioeconomic backgrounds, is likely to play an important role in tackling childhood and adolescent smoking.

The final part of the empirical analysis of this thesis (chapter 4) uses the Vietnam Young Lives Survey to investigate the mechanism through which small-scale weather shocks affect child nutritional status. Similar to previous studies, the results show that exposure to small-scale weather shocks is associated with a negative impact on child nutritional status. After controlling for a wide range of observed characteristics and community fixed effects, children living in households exposed to weather shocks are statistically significantly more malnourished than children living in unexposed households.

The results also show that small-scale weather shocks are associated with a reduction in household economic welfare (measured using total household consumption and expenditure on all food and non-food goods) as well as a reduction in the quality of dietary intake. Since child nutritional status is a function of the quality of dietary intake, this result suggests that the impact of weather shocks on child nutritional status is mediated through weather-induced shocks to household economic welfare which compromises the quality of child dietary intake.

Finally, chapter 4 shows evidence of a differential impact of weather shocks in children from different socioeconomic backgrounds. However, contrary to previous studies, a greater impact is observed on the nutritional status of children from wealthier backgrounds compared to children from poorer backgrounds. Given the pro-poor sampling strategy adopted by the VYLS, the study sample consists predominantly of disadvantaged households. Therefore 'wealthier' households are likely to be poor (at least in comparison to a nationally representative sample) and thus may still possess limited capacities to smooth consumption in response to weather-induced household consumption shocks. For poorer households, weather shocks may represent a small part of a host of other risky environmental conditions facing children living within these environments. This may explain the failure to observe a significant impact of weather shocks in children from poorer backgrounds.

Malnutrition in children remains a leading public health concern globally particularly in developing countries, not least because it has been implicated in 53% of the 10.6 million preventable deaths in children each year (Black et al., 2008, Hosseinpoor et al., 2006). In

addition, childhood malnutrition has been linked to adverse health, lower education and labour market outcomes in adulthood (Case and Paxson, 2008, Currie, 2009). Given the increasing frequency of natural disasters and the vulnerability of children when exposed, there is the need to investigate the extent and mechanisms through which weather shocks impact on child health as well as to identify particular groups of children who are most at risk. This will in turn inform policy responses aimed at protecting vulnerable households and children against the impact of weather shocks. Overall, chapter 4 highlights the importance of insuring household consumption as a means of protecting long- and short-term health outcomes of vulnerable children.

References

ACTION ON SMOKING AND HEALTH (ASH) 2011. Passive smoking: the impact on children. Available at: http://www.ash.org.uk/information/secondhand-smoke.

ADAIR, L. S. 1999. Filipino children exhibit catch-up growth from age 2 to 12 years. *Journal of Nutrition,* 129**,** 1140-1148.

ADAMS, J., PARKINSON, L., SANSON-FISHER, R. W. & WALSH, R. A. 2008. Enhancing self-report of adolescent smoking: The effects of bogus pipeline and anonymity. *Addictive Behaviors,* 33**,** 1291-1296.

ALDERMAN, H. & GARCIA, M. 1994. Food Security and Health Security: Explaining the Levels of Nutritional Status in Pakistan. *Economic Development and Cultural Change,* 42**,** 485-507.

ALDERMAN, H., HODDINOTT, J. & KINSEY, B. 2006. Long term consequences of early childhood malnutrition. *Oxford Economic Papers,* 58**,** 450-474.

ARTÍS, M., AYUSO, M. & GUILLÉN, M. 2002. Detection of Automobile Insurance Fraud With Discrete Choice Models and Misclassified Claims. *Journal of Risk and Insurance,* 69**,** 325-340.

BALIA, S. & JONES, A. M. 2008. Mortality, lifestyle and socio-economic status. *Journal of Health Economics,* 27**,** 1-26.

BANERJEE, A., DUFLO, E., POSTEL-VINAY, G. & WATTS, T. 2010. Long-Run Health Impacts of Income Shocks: Wine and Phylloxera in Nineteenth-Century France. *Review of Economics and Statistics,* 92**,** 714-728.

BARRIOS, R. E., STANSBURY, J. P., PALENCIA, R. & MEDINA, M. T. 2000. Nutritional status of children under 5 years of age in three hurricane-affected areas of Honduras. *Revista Panamericana de Salud Pública,* 8**,** 380-384.

BAUMAN, K. E. & ENNETT, S. E. 1994. Tobacco use by black and white adolescents: the validity of self-reports. *American Journal of Public Health,* 84**,** 394-398.

BECKER, G. S. & MULLIGAN, C. B. 1997. The Endogenous Determination of Time Preference. *The Quarterly Journal of Economics,* 112**,** 729-758.

BEHRMAN, J. R. & DEOLALIKAR, A. B. 1988. Chapter 14 Health and Nutrition. *In:* HOLLIS, C. & SRINIVASAN, T. N. (eds.) *Handbook of Development Economics.* Elsevier.

BEHRMAN, J. R. & SKOUFIAS, E. 2004. Correlates and determinants of child anthropometrics in Latin America: background and overview of the symposium. *Economics & Human Biology,* 2**,** 335-351.

BENOWITZ, N. L. 1996. Cotinine as a Biomarker of Environmental Tobacco Smoke Exposure. *Epidemiologic Reviews,* 18**,** 188-204.

BLACK, R. E., ALLEN, L. H., BHUTTA, Z. A., CAULFIELD, L. E., DE ONIS, M., EZZATI, M., MATHERS, C. & RIVERA, J. 2008. Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet,* 371**,** 243-260.

BLINDER, A. 1973. Wage discrimination: reduced form and structural estimates. *Journal of Human Resources,* 8**,** 436 - 455.

BLOW, L., LEICESTER, A. & WINDMEIJER, F. 2005. Parental Income and Children's Smoking Behaviour: Evidence from the British Household Panel Survey. *Institute of Fiscal Studies Working Paper Series*. London: Institute of Fiscal Studies

BOMMIER, A. & STECKLOV, G. 2002. Defining health inequality: why Rawls succeeds where social welfare theory fails. *Journal of Health Economics,* 21**,** 497-513.

BOUND, J., BROWN, C. & MATHIOWETZ, N. 2001. Chapter 59 Measurement error in survey data. *In:* HECKMAN, J. J. & LEAMER, E. (eds.) *Handbook of Econometrics*. Elsevier.

BREUER, W., HENS, T., SALZMANN, A. J. & WANG, M. 2011. Time Preferences, Culture, and Household Debt Maturity Choice. *NCCR Working Paper 674*. Available at: http://www.nccr-finrisk.uzh.ch/media/pdf/wp/WP674_A1.pdf.

CAMERON, L. & WILLIAMS, J. 2009. Is the relationship between socioeconomic status and health stronger for older children in developing countries? *Demography,* 46**,** 303-324.

CASE, A., LUBOTSKY, D. & PAXSON, C. 2002. Economic Status and Health in Childhood: The Origins of the Gradient *American Economic Review,* 92**,** 1308-1334.

CASE, A. & PAXSON, C. 2008. Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy,* 116**,** 499-532.

CENTER FOR DISEASE CONTROL AND PREVENTION 2004. The Health Consequences of Smoking: A report of the Surgeon General's Executive Summary. Available at: http://www.cdc.gov/tobacco/data_statistics/sgr/2004/pdfs/executivesummary.pdf.

CHALOUPKA, F. J. & GROSSMAN 1996. Price, Tobacco Control Policies and Youth Smoking. *Working Paper 5740*. Chicago: Department of Economics, University of Illinois, chicago.

CHALOUPKA, F. J. & PACULA, R. L. 1998. Limiting Youth Access to Tobacco: The Early Impact of the Synar Amendment on Youth Smoking *Department of Economics, University of Illinois at Chicago Working Paper*. Chicago, IL.

CHARLTON, A. 1996. Children and smoking: the family circle. *British Medical Bulletin,* 52**,** 90-107.

CLEVELAND, W. S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association,* 74**,** 829-836.

CONDLIFFE, S. & LINK, C. R. 2008. The Relationship between Economic Status and Child Health: Evidence from the United States. *American Economic Review,* 98**,** 1605-18.

CONTOYANNIS, P. & JONES, A. M. 2004. Socio-economic status, health and lifestyle. *Journal of Health Economics,* 23**,** 965-995.

CRAIG, R. & MINDELL, J. 2008. Health Survey for England 2006. Vol 2: Obesity and other risk factors in children. The Information Centre.

CURRIE, A., SHIELDS, M. A. & PRICE, S. W. 2007. The child health/family income gradient: Evidence from England. *Journal of Health Economics,* 26**,** 213-232.

CURRIE, C., MOLCHO, M., BOYCE, W., HOLSTEIN, B., TORSHEIM, T. & RICHTER, M. 2008. Researching health inequalities in adolescents: The development of the Health Behaviour in School-Aged Children (HBSC) Family Affluence Scale. *Social Science & Medicine,* 66**,** 1429-1436.

CURRIE, J. 2009. Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature,* 47**,** 87-122.

CUTLER, D. M. & LLERAS-MUNEY, A. 2010. Understanding differences in health behaviors by education. *Journal of Health Economics,* 29**,** 1-28.

DATAR, A., LIU, J., LINNEMAYR, S. & STECHER, C. 2011. The Impact of Natural Disasters on Child Health and Investments in Rural India. *RAND Working Paper Series WR-886*. Available at SSRN: http://ssrn.com/abstract=1980401.

DECICCA, P., KENKEL, D. & MATHIOS, A. 2002. Putting out the Fires: Will Higher Taxes Reduce the Onset of Youth Smoking? *The Journal of Political Economy,* 110**,** 144-169.

DEPARTMENT OF HEALTH 2010. A Smokefree Future: A Comprehensive Tobacco Control Strategy for England. London.

DERCON, S. 2002. Income Risk, Coping Strategies, and Safety Nets. *The World Bank Research Observer,* 17**,** 141-166.

DEWEY, K. G. & BEGUM, K. 2011. Long-term consequences of stunting in early life. *Maternal & Child Nutrition,* 7**,** 5-18.

DOHMEN, T. J., FALK, A., HUFFMAN, D. & SUNDE, U. 2008. The Intergenerational Transmission of Risk and Trust Attitudes. *CEPR Discussion Paper no. 6844*. London: Centre for Economic Policy Research

DOLCINI, M., ADLER, N. & GINSBERG, D. 1996. Factors influencing agreement between self-reports and biological measures of smoking among adolescents. *Journal of Research on Adolescence,* 6**,** 515–542.

DOUGLAS, S. & HARIHARAN, G. 1994. The hazard of starting smoking- estimates from a split population duration model. *Journal of Health Economics,* 13**,** 213-230.

EDOKA, I. P. 2011. Parental Income and Smoking Participation in Adolescents: Implications of misclassification error in empirical studies of adolescent smoking participation. *Health Econometrics and Data Group working paper WP 11/29*. York: University of York.

EMERY, S., WHITE, M. M. & PIERCE, J. P. 2001. Does cigarette price influence adolescent experimentation? *Journal of Health Economics,* 20**,** 261-270.

ENVIRONMENTAL PROTECTION AGENCY 1997. Health Effects of Exposure to Environmental Tobacco Smoke. California Environmental Protection Agency: Sacramento. Available at: http://oehha.ca.gov/pdf/chapter2.pdf.

ERIKSEN, W. 2004 Do people who were passive smokers during childhood have increased risk of long-term work disability? *European Journal of Public Health,* 14**,** 296–300.

ERREYGERS, G. 2009. Correcting the Concentration Index. *Journal of Health Economics,* 28**,** 504-515.

FALARIS, E. M. 2011. Misclassification of the dependent variable in binary choice models: evidence from five Latin American countries. *Applied Economics,* 43**,** 1315 - 1327.

FEWTRELL, L., KAUFMANN, R. B., KAY, D., ENANORIA, W., HALLER, L. & COLFORD, J. M. 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *The Lancet infectious diseases,* 5**,** 42-52.

FIRPO, S., FORTIN, N. M. & LEMIEUX, T. 2009. Unconditional Quantile Regressions. *Econometrica,* 77**,** 953-973.

FORTIN, N., LEMIEUX, T. & FIRPO, S. 2011. Chapter 1 - Decomposition Methods in Economics. *In:* ORLEY, A. & DAVID, C. (eds.) *Handbook of Labor Economics*. Elsevier.

FRIEDMAN, L., JOHNSON, B. & BRETT, A. 1990. Evaluation of substance-abusing adolescents by primary care physicians. *Journal of Adolescent Health Care,* 11**,** 227-230.

FRIJTERS, P., SHIELDS, M. A., WHEATLEY PRICE, S. & WILLIAMS, J. 2011. Quantifying the cost of passive smoking on child health: evidence from children's cotinine samples. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 174**,** 195-212.

FULLER, E. (ed.) 2009. *Smoking, drinking and drug use among young people in England in 2008*: The NHS Information Centre for Health and Social Care.

GILLILAND, F. D., BERHANE, K., ISLAM, T., WENTEN, M., RAPPAPORT, E., AVOL, E., GAUDERMAN, W. J., MCCONNELL, R. & PETERS, J. M. 2003. Environmental Tobacco Smoke and Absenteeism Related to Respiratory Illness in Schoolchildren. *American Journal of Epidemiology,* 157**,** 861-869.

GLENDINNING, A., SHUCKSMITH, J. & HENDRY, L. 1994. Social class and adolescent smoking behaviour. *Social Science & Medicine,* 38**,** 1449-1460.

GRAHAM, H. & POWER, C. 2004. Childhood Disadvantage and Adult Health: a Lifecourse Framework. National Health Service Health Development Agency, London.

GRIESLER, P. C., KANDEL, D. B., SCHAFFRAN, C., HU, M.-C. & DAVIES, M. 2008. Adolescents' Inconsistency in Self-Reported Smoking. *Public Opinion Quarterly,* 72**,** 260-290.

GROSSMAN, M. 1972. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy,* 80**,** 223-255.

GROSSMAN, M. 2000. Chapter 7 The human capital model. *In:* ANTHONY, J. C. & JOSEPH, P. N. (eds.) *Handbook of Health Economics.* Elsevier.

GRUBER, J. & ZINMAN, J. 2001. Youth Smoking in the United States: Evidence and Implications. *In:* GRUBER, J. (ed.) *Risky Behavior Among Youths: An Economic Analysis.* The University of Chicago Press.

GUHA-SAPIR, D., VOS, F., BELOW, R. & PONSERRE, S. 2012. Annual Disaster Statistical Review 2011: The Numbers and Trends. Brussels: Centre for Research on the Epidemiology of Disasters (CRED).

HARRIS, J. E. & CHAN, S. W. 1999. The continuum-of-addiction: cigarette smoking in relation to price among Americans aged 15–29. *Health Economics,* 8**,** 81-86.

HAUSMAN, J. 2001. Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left. *Journal of Economic Perspectives,* 15**,** 57-67.

HAUSMAN, J. A., ABREVAYA, J. & SCOTT-MORTON, F. M. 1998. Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics,* 87**,** 239-269.

HODDINOTT, J. & KINSEY, B. 2001. Child Growth in the Time of Drought. *Oxford Bulletin of Economics and Statistics,* 63**,** 409-436.

HÖFLER, M. 2005. The effect of misclassification on the estimation of association: a review. *International Journal of Methods in Psychiatric Research* 14**,** 92-101.

HOSSEINPOOR, A., VAN DOORSLAER, E., SPEYBROECK, N., NAGHAVI, M., MOHAMMAD, K. & MAJDZADEH, R. 2006. Decomposing socioeconomic inequality in infant mortality in Iran. *Int J Epidemiol,* 35**,** 1211 - 9.

JANN, B. 2008. The Blinder-Oaxaca decomposition for linear regression models. *Stata Journal,* 8**,** 453-479.

JARVIS, M. J., FIDLER, J., MINDELL, J., FEYERABEND, C. & WEST, R. 2008. Assessing smoking status in children, adolescents and adults: cotinine cut-points revisited. *Addiction,* 103**,** 1553-1561.

JAYACHANDRAN, S. 2006. Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy, University of Chicago Press,* 114**,** 538-575.

JAYACHANDRAN, S. 2009. Air Quality and Early-Life Mortality. *Journal of Human Resources,* 44**,** 916-954.

JONES, F. L. & KELLEY, J. 1984. Decomposing Differences between Groups. *Sociological Methods & Research,* 12**,** 323-343.

JÜRGES, H. forthcoming. Collateral damage: The German food crisis, educational attainment and labor market outcomes of German post-war cohorts. *Journal of Health Economics*.

KAHN, M. E. 2005. The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions. *Review of Economics and Statistics,* 87**,** 271-284.

KAKWANI, N., WAGSTAFF, A. & VAN DOORSLAER, E. 1997. Socioeconomic inequalities in health: measurement, computation, and statistical inference. *Journal of Econometrics,* 77**,** 87 - 103.

KANDEL, D. B., SCHAFFRAN, C., GRIESLER, P. C., HU, M.-C., DAVIES, M. & BENOWITZ, N. 2006. Salivary Cotinine Concentration Versus Self-Reported Cigarette Smoking: Three Patterns of Inconsistency in Adolescence. *Nicotine & Tobacco Research,* 8**,** 525-537.

KENKEL, D. S. 1991. Health Behavior, Health Knowledge, and Schooling. *Journal of Political Economy,* 99**,** 287-305.

KENKEL, D. S., LILLARD, D. R. & MATHIOS, A. D. 2004. Accounting for misclassification error in retrospective smoking data. *Health Economics,* 13**,** 1031-1044.

KHWAJA, A., SILVERMAN, D. & SLOAN, F. 2007. Time preference, time discounting, and smoking decisions. *Journal of Health Economics,* 26**,** 927-949.

LEIGH, J. P. 1986. Accounting for Tastes: Correlates of Risk and Time Preferences. *Journal of Post Keynesian Economics,* 9**,** 17-31.

LOPEZ, A. 2005. The evolution of the Global Burden of Disease framework for disease, injury and risk factor quantification: developing the evidence base for national, regional and global public health action. *Globalization and Health,* 1**,** 5.

LOUREIRO, M. L., SANZ-DE-GALDEANO, A. & VURI, D. 2010. Smoking Habits: Like Father, Like Son, Like Mother, Like Daughter? *Oxford Bulletin of Economics and Statistics,* 72**,** 717-743.

MACHADO, J. A. F. & MATA, J. 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics,* 20**,** 445-465.

MAGDER, L. S. & HUGHES, J. P. 1997. Logistic Regression When the Outcome Is Measured with Uncertainty. *American Journal of Epidemiology,* 146**,** 195-203.

MARTORELL, R., KHAN, L. K. & SCHROEDER, D. G. 1994. Reversibility of Stunting-Epidemiologic finding of children from developing countries. *European Journal of Clinical Nutrition,* 48**,** S45-S57.

MCCLEMENTS, D. 1977. Equivalence scales for children. *Journal of Public Economics* 8**,** 191-210.

MELLY, B. 2005. Decomposition of differences in distribution using quantile regression. *Labour Economics,* 12**,** 577-590.

MILLER, G. & URDINOLA, B. P. 2010. Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia. *The journal of political economy,* 118**,** 113-155.

MUELLER, V. & QUISUMBING, A. 2010. Short- and long-term effects of the 1998 Bangladesh flood on rural wages. *IFPRI discussion papers 956*. International Food Policy Research Institute (IFPRI).

MUELLER, V. A. & OSGOOD, D. E. 2009. Long-term Impacts of Droughts on Labour Markets in Developing Countries: Evidence from Brazil. *The Journal of Development Studies,* 45**,** 1651-1662.

NGUYEN, N. P. 2008. An Assessment of the Young Lives Sampling Approach in Vietnam. Young Lives Technical Note number 4.

O'DONNELL, O., VAN DOORSLAER, E. & ROSATI, F. C. 2002. Child Labour and Health: Evidence and Research Issues. Understanding Children's Work Programme Working Paper. Available at SSRN: http://ssrn.com/abstract=1780320

O'DONOGHUE, T. & RABIN, M. 2001. Risky Behaviour Among Youth: Some Issues from Behavioral Economics. *In:* GRUBER, J. (ed.) *Risky Behaviour Among Youth: An Economic Analysis*. The University of Chicago Press.

OAXACA, R. 1973. Male-female wage differentials in urban labour markets. *International Economic Review,* 9**,** 693 - 709.

PERETTI-WATEL, P., CONSTANCE, J., GUILBERT, P., GAUTIER, A., BECK, F. & MOATTI, J.-P. 2007. Smoking too few cigarettes to be at risk? Smokers' perceptions of risk and risk denial, a French survey. *Tobacco Control,* 16**,** 351-356.

PITT, M. M. & ROSENZWEIG, M. R. 1985. Health and Nutrient Consumption Across and Within Farm Households. *The Review of Economics and Statistics,* 67**,** 212-223.

PÖRTNER, C. C. 2010. Natural Hazards and Child Health University of Washington, Seattle. Available at SSRN: http://ssrn.com/abstract=1599432

POWELL, L. M. & CHALOUPKA, F. J. 2005. Parents, public policy, and youth smoking. *Journal of Policy Analysis and Management,* 24**,** 93-112.

POWELL, L. M., TAURAS, J. A. & ROSS, H. 2005. The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. *Journal of Health Economics,* 24**,** 950-968.

PRESCOTT-CLARKE, P. & PRIMATESTA, P. (eds.) 1998. *Health Survey for England 1996*: The Stationery Office.

REINHOLD, S. & JÜRGES, H. 2011. Parental income and child health in Germany. *Health Economics (Available from: DOI: 10.1002/hec.1732).*

ROGGERO, P., MANGIATERRA, V., BUSTREO, F. & ROSATI, F. 2007. The Health Impact of Child Labor in Developing Countries: Evidence From Cross-Country Data. *American Journal of Public Health,* 97**,** 271-275.

ROSENZWEIG, M. R. & SCHULTZ, T. P. 1983. Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight. *Journal of Political Economy,* 91**,** 723-746.

ROSS, H. & CHALOUPKA, F. J. 2003. The effect of cigarette prices on youth smoking. *Health Economics,* 12**,** 217-230.

SIMS, M., TOMKINS, S., JUDGE, K., TAYLOR, G., JARVIS, M. J. & GILMORE, A. 2010. Trends in and predictors of second-hand smoke exposure indexed by cotinine in children in England from 1996 to 2006. *Addiction,* 105**,** 543-553.

SOTERIADES, E. S. & DIFRANZA, J. R. 2003. Parent's Socioeconomic Status, Adolescents' Disposable Income, and Adolescents' Smoking Status in Massachusetts. *American Journal of Public Health,* 93**,** 1155-1160.

STEIN, L., COLBY, S., O'LEARY, T., MONTI, P., ROHSENOW, D., SPIRITO, A., RIGGS, S. & BARNETT, N. 2002. Response Distortion in Adolescents Who Smoke: a Pilot Study. *Journal of Drug Education,* 32**,** 271–286

THE NHS INFORMATION CENTRE 2010. Statistics on Smoking: England. Available at: http://www.ic.nhs.uk/pubs/smoking10.

THOMAS, D., STRAUSS, J. & HENRIQUES, M.-H. 1990. Child survival, height for age and household characteristics in Brazil. *Journal of Development Economics,* 33**,** 197-234.

THOMAS, T., CHRISTIAENSEN, L., DO, Q. T. & TRUNG, L. D. 2010. Natural disasters and household welfare: evidence from Vietnam. *Policy Research WPS 5491*. The World Bank Development Research Group.

TOBACCO ADVISORY GROUP OF THE ROYAL COLLEGE OF PHYSICIANS 2010. Report on passive smoking and children.

TUAN, T., LAN, P. T., HARPHAM, T., HUONG, N. T., THACH, T. D., TOD, B., DUA, T. T. & HA, N. T. 2003. Young Lives Preliminary Country Report: Vietnam. London.

TUINSTRA, J., GROOTHOFF, J. W., VAN DEN HEUVEL, W. J. A. & POST, D. 1998. Socio-economic differences in health risk behavior in adolescence: Do they exist? *Social Science & Medicine,* 47**,** 67-74.

TYAS, S. L. & PEDERSON, L. L. 1998. Psychosocial factors related to adolescent smoking: a critical review of the literature. *Tobacco Control,* 7**,** 409-420.

UNISDR 2011a. Disaster Risk Management for Health: Child Health. Geneva: United Nations International Strategy for Disaster Reduction Secretariat.

UNISDR 2011b. Global Assessment Report on Disaster Risk Reduction 2011 Geneva: United Nations International Strategy for Disaster Reduction Secretariat.

US DEPARTMENT OF HEALTH AND HUMAN SERVICES 2006. The health consequences of involuntary exposure to tobacco smoke: A report of the surgeon general. Available at: www.surgeongeneral.gov/library/secondhandsmoke/.

VALLEJO-TORRES, L. & MORRIS, S. 2010. The contribution of smoking and obesity to income-related inequalities in health in England. *Social Science & Medicine,* 71**,** 1189-1198.

VAN DEN BERG, G. J., LINDEBOOM, M. & PORTRAIT, F. 2006. Economic Conditions Early in Life and Individual Mortality. *American Economic Review,* 96**,** 290-302.

VICTORA, C. G., ADAIR, L., FALL, C., HALLAL, P. C., MARTORELL, R., RICHTER, L. & SACHDEV, H. S. 2008. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet,* 371**,** 340-357.

VIETNAM GENERAL STATISTICS OFFICE 2010. Result of the Survey on Household Living Standards 2010. : Statistical Publishing House. Hanoi, Vietnam.

WAGENKNECHT, L. E., BURKE, G. L., PERKINS, L. L., HALEY, N. J. & FRIEDMAN, G. D. 1992. Misclassification of smoking status in the CARDIA study: a comparison of self-report with serum cotinine levels. *American Journal of Public Health,* 82**,** 33-36.

WAGSTAFF, A. 2005. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics,* 14**,** 429 - 32.

WAGSTAFF, A., PACI, P. & VAN DOORSLAER, E. 1991. On the measurement of inequalities in health. *Social Science and Medicine,* 33**,** 545 - 57.

WASSERMAN, J., MANNING, W. G., NEWHOUSE, J. P. & WINKLER, J. D. 1991. The effects of excise taxes and regulations on cigarette-smoking *Journal of Health Economics,* 10**,** 43-64.

WATSON, J. T., GAYER, M. & CONNOLLY, M. A. 2007. Epidemics after Natural Disasters. *Emerging Infectious Diseases* 13**,** 1-5.

WHO 1995. Physical Status: Use and Interpretation of Anthropometry. World Health Organisation, Geneva.

YOLTON, K., DIETRICH, K., AUINGER, P., LANPHEAR, B. P. & HORNUNG, R. 2004. Exposure to Environmental Tobacco Smoke and Cognitive Abilities among U.S. Children and Adolescents. *Environmental Health Perspective,* 113.

YUN, M.-S. 2005. A simple solution to the identification problem in detailed wage decompositions *Economic Inquiry,* 43**,** 766-772.