The Effect of Household Permanent Income and Maternal Employment on Youth Overweight and Obesity in Australia

A Thesis submitted in partial fulfilment of the requirements of the award of the degree:

Bachelor of Commerce (Hons)

University of Wollongong

Department of Economics

By

James Bishop

Certificate of Originality

I, James Bishop, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Bachelor of Commerce (Honours), in the Department of Economics, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The work contained in this thesis has not been previously submitted for a degree or other qualification at any other higher education institution. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made.

Signed:

Date:

Acknowledgements

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). The findings and views reported in this paper, however, are those of the author and should not be attributed to either FaHCSIA or the MIAESR.

This research was conducted as part of a project titled 'An Investigation of Chronic and Transitory Poverty in Australia in the Twenty-First Century'. The project was funded by an Australian Research Council Discovery Grant. A student stipend from that grant supported the research on which this thesis is based. Financial assistance was also provided from the Reserve Bank of Australia's (RBA) Cadetship Award.

I would also like to acknowledge my Supervisors, Joan R. Rodgers and Peter Siminski, for all your encouragement, guidance and support throughout this year. I greatly appreciate all the feedback, assistance and time that you have provided me. I would like to thank the Faculty of Commerce, and, in particular, the School of Economics for its support over the past four years. I would also like to thank my parents, who have always supported me in everything that I have done. Lastly, thank you to Tanya, for your encouragement and friendship.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABS – Australian Bureau of Statistics

AUSEI06 – Australian Socioeconomic Index 2006

BMI – Body Mass Index

CNEF – Cross National Equivalent File

CPI – Consumer Price Index

GST – Goods and Services Tax

HILDA – Household, Income and Labour Dynamics in Australia Survey

LSAC – Longitudinal Survey of Australian Children

MIAESR – Melbourne Institute for Applied Economic and Social Research (University of Melbourne)

NHS – National Health Survey

OECD – Organisation for Economic Co-operation and Development

RBA – Reserve Bank of Australia

RHED Income – Real, Household Equivalised Disposable Income

SEIFA – Socio-Economic Index for Areas

SES – Socioeconomic Status

Abstract

The objective of this thesis is to analyse the effects of household permanent income and maternal employment on obesity in Australian youths. This is motivated by recent trends which suggest that household investments of both money and time may be factors contributing to the "epidemic" of obesity in Australia. Drawing on insights from household production theory and using data from the Household, Income and Labour Dynamics in Australia Survey, this thesis provides the first estimates of the *causal* effect of household income and maternal employment on the obesity status of older youths (aged over 12 years) in the Australian and international literatures. To achieve this aim, instrumental variables, sibling-differences and proxy variables are used to account for potential endogeneity in the two variables of interest. The results show that the permanent income gradient in weight is reduced to zero as confounding covariates are included in the model and after accounting for endogeneity. This suggests that income transfers are not a feasible approach to reducing the weight of low income youths. Regarding mothers' employment, part-time or full-time work is associated with decreased excess body weight in youths, relative to not working at all. This relationship is driven by youths from high socioeconomic status households but does not operate through increased earned income. This suggests a potential weakness in the underlying theory, at least in terms of its predictions for weight outcomes in adolescents.

1 Introduction

1.1 Motivation, Objectives and Background

During the past three decades, the prevalence of overweight and obesity in children and adolescents has increased throughout the world (WHO 2000). Among Australian children and adolescents the rise in overweight and obesity began in the 1980s, and had reached "epidemic proportions" by 2007-08, with approximately one quarter of 5-17 year olds classified as overweight or obese (AIHW 2004, p.1; ABS 2009c, p.8). If recent trends continue almost half of all Australian children will be overweight by the year 2025 (ASSO 2004).

Moreover, descriptive evidence for Australia has highlighted a trend toward greater prevalence of overweight and obesity among children from low income and low socioeconomic status (SES) households (O'Dea 2003; Salmon *et al.* 2005; Booth *et al.* 2001; O'Dea 2009). However, despite calls for more targeted intervention (O'Dea 2009, p.1), a *causal* relationship between family income and child and adolescent obesity has yet to be established in the Australian or in the international literature.

For this reason, the first objective of this thesis is to investigate the extent to which permanent family income is causally related to overweight and obesity in Australian youths. To achieve this aim, longitudinal data from the first seven waves of the Household, Income and Labour Dynamics in Australia (HILDA) survey are exploited to construct measures of household 'permanent income' using income data from several years. To the extent that income fluctuates between years, or is measured with error, this should provide a better measure of the economic resources available to a household than a single-year measure of household income (Blau 1999, p.265).

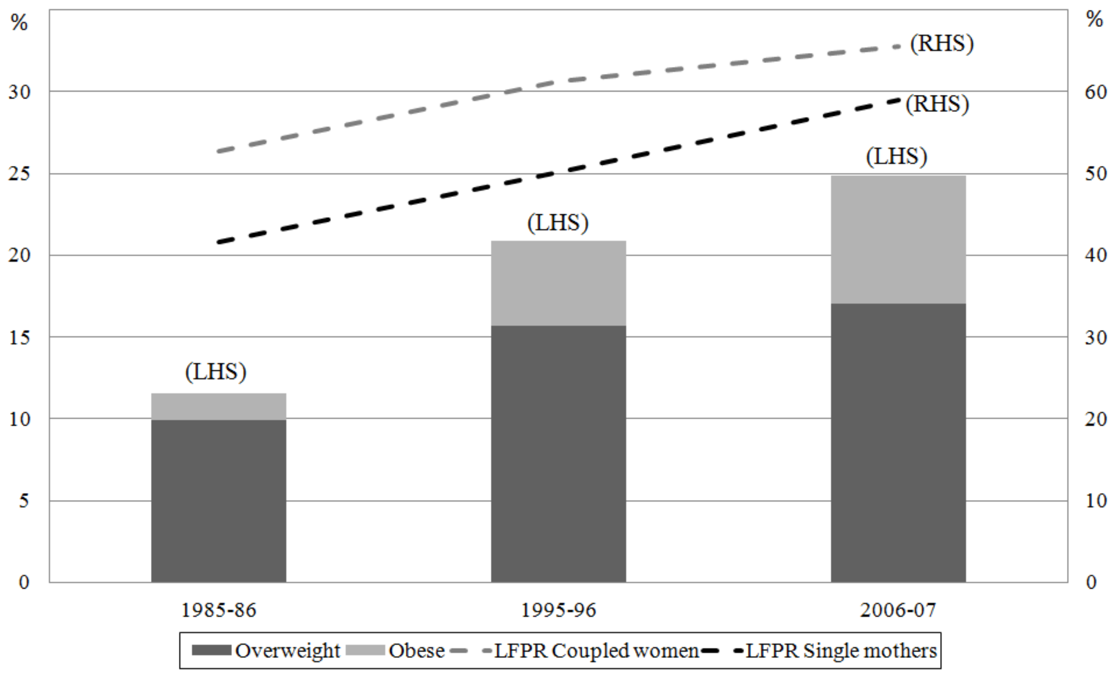
Further, since excess body weight is a 'stock', reflecting a long-term imbalance between energy intake and expenditure, low income over a number of years (i.e. low permanent income) is likely to lead to an accumulation of excess body weight over time. For this reason, obesity may better capture the accumulation of the effects of low household income and low SES than other chronic medical conditions in children.

Recognising that households also face constraints on *time*, as well as income, and that parental time allocation decisions may have important consequences for child and adolescent health outcomes, the second objective of this thesis is to explore the effects of maternal employment on youth overweight and obesity in Australia.

Studies for the US, England and Canada have argued that women's gradual reduction in domestically oriented work and accompanying increase in labour market participation has been a key factor in the rise in overweight and obese children and adolescents in these countries (Anderson *et al.* 2003, Ruhm 2004; Liu 2006; Scholder 2008). In Australia, the rise in the labour force participation of mothers with dependent children over recent decades has largely coincided with the increasing prevalence of child and youth overweight and obesity (see Figure 1.1).[1]

[1] Note for Figure 1.1: Overweight and obesity prevalence rates for 1995 and 2007 based on data from the National Health Survey (ABS 2009c), applying to children aged 5-17 years; estimates for 1985 use data from the Australian Health and Fitness Survey, made comparable by Magarey *et al.* (2001), applying to children aged 7-15 years. Labour force participation data derive from ABS 1986, 1996 and 2006 Census of Population and Housing, Available: http://www.abs.gov.au/. LHS and RHS in Figure 1.1 mean the series is plotted on the right-hand side axis and left-hand side axis, respectively.

Figure 1.1 Labour Force Participation of Mothers with Children 0-14 Years and Overweight and Obesity in Australian Children and Adolescents



Source: ABS 1986, 1996 and 2006 Census of Population and Housing, ABS (2009c), Magarey et al. (2001)

These parallel trends provide a *prima facie* case for the hypothesis that changes in maternal work patterns have played a role in the accelerating obesity epidemic.[2] A key motivation of this thesis is to determine whether these coincident trends represent a causal relation, or simply a spurious association. With the exception of one recent paper (Zhu 2007), there has been no investigation of the attendant causal relationship for Australia, and no study, Australian or otherwise, has examined the effect of mothers' employment on the weight outcomes of *older* children (aged 12 years or more). Hence, this research fills an important gap in the literature.

[2] As depicted in Figure 1.1, the labour force participation rate of both coupled and single women with dependent children increased consistently over the period 1985-86 to 2006-07, a period characterised by a two-fold increase in the rate of child and adolescent obesity.

1.2 The Social and Economic Costs of Paediatric Obesity

Though the link between obesity and disease is less clear in children than adults, due to the time it takes for the consequences of obesity to develop, child and adolescent obesity has emerged as a serious public health concern (Access Economics 2006, p.11). In particular, child and adolescent obesity has been associated with increased risk of Type 2 diabetes, insulin resistance syndrome, orthopaedic problems (Daniels 2006), and psychosocial problems such as low self-esteem, depression, discrimination and bullying (Monheit *et al.* 2007). The culmination of these adverse health consequences has led some observers to predict that the increasing prevalence of paediatric obesity could reverse the modern era's steady increase in life expectancy (Olshansky *et al.* 2005). In Australia, health problems associated with overweight also translate into substantial health care spending, estimated at around $1.5 billion a year (ASSO, 2004).[3]

Child and adolescent obesity may also develop into an important mechanism by which poverty is transmitted through the generations of a family. That is, part of the intergenerational transmission of poverty may work through the impact of parents' income on the excess body weight of their children. Studies show that obesity in childhood is a strong predictor of obesity in adulthood (Whitaker *et al.* 1997; Guo *et al.* 2002), and that obesity in adulthood is associated with lower wages and earnings (Baum and Ford 2004; Cawley 2004; Averett and Korenmann 1996), penalties in the 'marriage market' (Cawley *et al.* 2006; Gortmaker *et al.* 1993; Mukhopadhyay 2008), increased health care expenditure (Finkelstein *et al.* 2005, p247-249; Access Economics 2006) and elevated risk of health problems and mortality (Daniels 2006). Furthermore, paediatric obesity has been tenuously linked to poorer adult labour market outcomes via its effect on educational

[3] The total cost of the obesity epidemic in Australia – including both direct and indirect costs – was estimated at $21 billion in 2005 (Access Economics 2006). No data could be found on the economic costs attributing to child obesity, specifically.

outcomes in childhood (Canning and Mayer 1996; Sabia 2007; Gortmaker *et al.* 1993; Datar and Sturm 2006), however, the evidence of causality is mixed (Kaestner and Grossman 2008; Taras *et al.* 2005). This research suggests that child obesity is important not only for its own sake but because it affects children's future prospects more broadly.

1.3 Structure of the Thesis

This thesis is structured as follows. Chapter 2 reviews the theoretical models used in the literature on child health. Following this, Chapter 3 provides a review of the empirical methodologies used in the child health literature to estimate the effects of household income and mothers' employment on children's outcomes. Chapter 4 reviews the empirical findings on the relationship between household income and child and adolescent obesity. The literature examining the effects of income on child health and development is also surveyed in this chapter, to place the contribution of this thesis into context. Chapter 4 concludes with a review of the evidence on the maternal employment-child obesity linkage. The data, sample and variables used in the empirical analysis are outlined and justified in Chapter 5, followed by a brief descriptive analysis in Chapter 6 and an outline of the methodology used in this thesis in Chapter 7. Chapter 8 presents the main econometric results and Chapter 9 presents a summary of the findings and conclusions from this thesis and presents suggestions for areas for further research.

2 Economic Theory of Child Health Production

There are several economic theories that could underpin an empirical analysis of child health production. Following several recent studies (Blau 1999; Ruhm 2004; Blau and Grossberg 1992; Zhu 2007; Khanam *et al.* 2009) the theoretical model underlying this analysis derives from the theory of household production, which has origins in the work of Becker (1965) and Becker and Lewis (1973). This model, which was adapted to the analysis of health capital by Grossman (1972, 2000) and Jacobson (2000), considers the stock of child health to be the output of a multivariable production process in which the productive entity is an utility-maximising household. In recognising the role of parental time and purchased inputs in the production of child health, this theoretical model provides an encompassing yet parsimonious representation of the key phenomena investigated in this thesis.

This chapter also details several alternative theories on child health, such as the so-called 'good-parent' theories, which are grounded in sociological theory. Insights from these theories, which generally reinforce the findings of the health capital model, also inform aspects of the empirical specifications in this thesis. The chapter concludes with a discussion of the permanent income hypothesis of consumption and saving.

2.1 Child Health as a 'Commodity'

In Becker's (1965) household production model, time and market goods do not directly provide utility. Instead, time and goods are inputs into the production of "commodities", which do directly provide utility. The commodities cannot be purchased in the marketplace but are produced by households using market purchases, own time, and various environmental inputs, such as household ability and human capital (Becker 1981, pp.7-8). A Beckerian household production function can be conceptualised as follows;

$$U = U(H_1, \ldots, H_m)$$

where $H_1, \ldots, H_m$ are the various commodities consumed. These commodities are self-produced according to:

$$H_i = f_i(x_i, t_{hi}; E_i)$$

where x_i and t_{hi} represent the possibly many goods and types of time used to produce the i^{th} commodity, and E_i represents household ability, human capital, social and physical capital and other environmental variables (Becker 1981, p.9). At a basic level, "child health" is a commodity that is produced when the household purchases market goods such as medical care, nutritious foods, exercise, housing and clothing and combines these with their own time inputs. However, this simple conceptualisation ignores important aspects about the accumulation and depreciation of health capital (Grossman 1972).

2.2 The Grossman Model of the 'Demand for Health' and its Extensions

Using the key concept of home production elaborated in Becker (1965), Grossman (1972) adapted the household production model to the analysis of health capital. The Grossman model considered an intertemporal utility function where utility depends on both the flow of "healthy days" from a stock of health available in a given period and on the consumption of other commodities.[4] Individuals are assumed to inherit an initial stock of health that depreciates over time but can be increased via gross investments of own time and market goods. Furthermore,

[4] Grossman's major contribution was to incorporate an investment motive in addition to a consumption motive by proposing that health is a durable capital stock that produces an output of "healthy time". Treating health as endogenous was the major difference between the Grossman model and the health models that preceded it.

education was emphasised as the key factor that increased one's efficiency in producing health and reducing the shadow price of investment at any age.

The Grossman model was extended by Jacobson (2000) to treat the family (rather than the individuals making up the family) as the production unit. In this model, each family member is the producer of his or her own health and the health of other family members. Consequently, the time and income of all family members are used in the production of health for each member of the family. The family unit is defined as a mother, a father and a child. Parents derive utility from the health of their child, with the child being a passive participant in the production of his or her own health (Jacobson 2000, p.614).

The following section draws on aspects of the household production model and its extensions (Grossman 1972; Jacobson 2000) to detail an economic model of the determinants of child health production. The model is adapted, with small modifications from Currie (2008, pp.4-7), Ruhm (2004, p.4) and Khanam *et al.* (2009, p.3). Following previous work (Zhu 2007; Ruhm 2004), 'child health' is generalised to 'child obesity', which is assumed to enter inversely in health. For simplicity, the model is for the one-child case, and thus ignores the substitution of time and market goods between children (James-Bardumy 2005, p.179).

2.3 The Model of Child Health Production

First, it is assumed that households maximise an intertemporal utility function of the form:

$$\sum_{t=1}^{T}(1+\sigma)^{-t}\,U_t + B(A_{t+1}) \qquad (2.1)$$

where σ is the discount rate, U_t is instantaneous utility for a family, B is a bequest function and A_t is assets. It is assumed that instantaneous utility in period t is given by a strictly concave function:

$$U_t = U(Q_t, C_t, L_{Lt}; Z_{Ut}, \varepsilon_{Ut}) \qquad (2.2)$$

where Q_t is the stock of child health, C_t is other commodities consumed by the household, L_{Lt} is parental leisure time, and Z_{Ut} and ε_{Ut} are exogenous observable and unobservable factors influencing U_t, respectively. It is important to note that consumers do not necessarily value the health of their children above everything else, since child health is only one argument of the household utility function.

The health of children may contribute to parents' utility for several reasons. Grossman and Becker emphasise the increase in direct utility as the reason for parents investing in child's health. That is, parents may enjoy spending time with their healthy children or care altruistically about their future happiness. In this sense, good child health is a 'consumption commodity'. Other authors (Jacobson 2000; Guryan 2008) argue that child health is also an 'investment commodity' since increased child health reduces time spent taking care of sick children in the future.

The production function for child health describes the way that inputs can be converted into health (Grossman 1972, 2000);

$$Q_t = Q(Q_{t-1}, L_{Qt}, F_t; Z_{Qt}, \varepsilon_{Qt}) . \qquad (2.3)$$

where Q_{t-1} is the stock of child health in the previous period, L_{Qt} is the amount of parental time used in the production of child health, F_t is purchased inputs like food or medical care, and Z_{Qt} and ε_{Qt} are the other exogenous observable and unobservable variables respectively affecting Q_t. The semicolon before the latter set of variables highlights the difference between these variables and the endogenous goods and time inputs.[5]

---

[5] Grossman assumes the production functions are homogenous of degree one in both goods and time inputs (Jacobson 2000, p.617)

From the child health production function (2.3) a few initial observations can be made. First, the current health status of a child, Q_t, depends partly on prior health status, Q_{t-1}, and therefore on endowments of health at birth, Q_0, and all past investments of parents.[6] For this reason, the model is dynamic. The treatment of health as a durable capital stock that increases with investment in health and depreciates over time was an innovation introduced by Grossman (1972), and corresponds closely to economists' views of human capital accumulation.

Second, parents do not purchase child health from the market but produce it by spending time on health improving activities as well as by purchasing health inputs, such as food, clothing, sporting equipment, safe neighbourhoods and medical care for their children. Thus, the demand for market goods and services and parental non-market time is a derived demand.[7] [8]

Finally, the production function is affected by the efficiency or productivity of a given family as reflected in the personal characteristics of its members. In particular, Grossman (2000, p.351) emphasises that knowledge capital, specifically years of completed schooling, raises the efficiency of the production process in the nonmarket sector, defining efficiency as "the amount of health obtained from a given amount of health inputs". For this reason, Z_{Qt} is most often associated with the consumer's stock of knowledge or human capital exclusive of health capital.

---

[6] Recursively substituting in lags of Q, equation (2.3) can be rewritten as: $Q_t = Q(Q_0, L_Q, F; Z_Q, \varepsilon_Q)$ where L_Q, F, Z_Q and ε_Q are vectors of current and lagged values (e.g. $F = \{F_t, F_{t-1}, ..., F_{t-n}\}$) and Q_0 is the initial health stock. That is, the health stock of the child at age t (Q_t) is an additively separable function of purchased inputs, maternal work hours and other observable and unobservable determinants at ages $t-n$ through to t.

[7] In a similar respect, the child health "commodity" does not have a market price (it not purchased), but does have a shadow price equal to its cost of production.

[8] Importantly, the production function (2.3) implies a 'separability' property, in that different combinations of goods and time could be used to produce the same commodity, in this case, child health (Becker 1981, p.9).

Importantly, the family will also face a constraint on their income and time. For this reason, the model specifies two constraints; the household budget constraint and the time constraint.

First, the household budget constraint limits purchases of child inputs and other market goods to the amount of earned and unearned income:

$$Y_t = I_t + w_t L_{Wt} + rA_t = P_{Ft}F_t + P_{Ct}C_t \tag{2.4}$$

where Y_t is total family income, I_t is non-labour income, w_t is the wage rate, L_{Wt} is the time spent to earn wage income, r is the market interest rate, and P_{Ft} and P_{Ct} are the prices of F_t and C_t, respectively. The initial endowment of assets, A_0, is assumed to be given.

The household also faces a time constraint. This recognises that the parents have an endowment of time which they can allocate between either leisure, the production of child health or market work:

$$L = L_{Lt} + L_{Qt} + L_{Wt} \tag{2.5}$$

where L denotes the total fixed amount of time available in a given period (i.e. 24 hours in a day).[9]

The family's decision problem can be portrayed as

$$\max_{Q_t, F_t, Y_t, L_{lt}, L_{wt}, L_{ht}} \sum_{t=1}^{T} (1+\sigma)^{-t}\, U_t + B(A_{t+1}) \tag{2.6}$$

subject to the production function (2.4), the budget constraint (2.5), the time constraint (2.4) and the condition that the initial health stock of the child is positive $Q_0 > 0$ and exogenously determined.

[9] Jacobson (2000) included time spent taking care of a sick child in the time constraint, assuming that time required to care for a sick child at time t is a positive function of Q_t. However, treating child health endogenously complicates the model considerably.

Taking the first derivatives of the Lagrangian function with respect to Q_t, and recursively substituting in for lags of Q until the initial condition is met, yields a Marshallian demand function for Q_t;

$$Q_t^* = Q(Q_0, \omega_k; Z_{Ut}, Z_{Qt}, \varepsilon_{Qt}, \varepsilon_{Ut}) \tag{2.7}$$

where $\omega_k = \{Q, F, C, L_L, L_W, L_Q\}$ and $k = 1, 2, \ldots t-1$ (Khanam *et al.* 2009, p.3).

The Marshallian demand function (2.7) shows that the optimal stock of child health is determined by the consumption of child-related purchased inputs and other goods and services, as well as the allocation of parental time between leisure, the production of child health and time in market work (Khanam *et al.* 2009, p.3; Jacobson 2000).

This model can be used to conceptualise the two phenomena of importance to this thesis. First, it can be used to show how low parental income can translate into poor health outcomes for children. Second, it can be used to explain how changes in the intensity of maternal employment could affect the health and overweight outcomes of children and adolescents. These outcomes will be discussed in the following section.

2.3.1 Family Income

The primary mechanism by which low family income translates into poor child health is via a tighter budget constraint. That is, poor households are able to purchase fewer and lower quality market goods and services, including those involved in the production of child health, such as nutritious foods, sporting equipment and medical care. By comparison, families with higher incomes are less constrained in their potential health investment as a smaller portion of their budget must be allocated to other necessities.[10]

[10] Chia (2008, p.233) finds that family income constraints do matter in determining whether children participate regularly in sporting activities. Drewnowski and Specter (2004) find evidence

Conversely, it could also be the case that high income parents have a higher value of time in market work (relative to low income parents), and consequently invest less non-market time in the production of child health. This proposition is discussed in Section 2.3.3. Irrespectively, an exogenous increase in household incomes will unambiguously increase child health via the purchase of more, better quality health inputs.[11] The causal effect on child health of an exogenous shock to household incomes is commonly termed the "income effect" or the "income gradient in child health" in the child health literature (Blau 1999; Case *et al.* 2002; Currie and Stabile 2003).

One particularly important market good identified in the literature is housing and neighbourhood. Greater income allows a family to purchase residence in a more advantaged suburb. This is particularly important for a study of youth obesity, given that differences in the design and environmental characteristics of a neighbourhood and the SES of its residents can significantly contribute to the likelihood of overweight (Kavanagh *et al.* 2007; Monheit 2007). These neighbourhood characteristics may also have effects on child health independent of income, for instance, if the propensity for fast-food restaurants to locate in disadvantaged areas is higher (Kavanagh *et al.* 2007).[12]

2.3.2 Education and the Efficiency of Health Production

In addition to the direct effect of household income on the household budget constraint, household income may also be correlated with other factors that affect child health. As discussed, the models of Grossman (1972) and Jacobson (2000)

---

of an inverse relationship between dietary energy density (MJ/kg) and energy cost ($/MJ). They argue that because energy-dense foods allow for higher energy consumption at a lower cost, the selection of energy dense foods by low income households may represent a deliberate strategy to save money.

[11] Income effects may be small, however, if the additional income is used to purchase goods and services that are not involved in the production of child health.

[12] Bradbury (2003, p.19) argues that, "in modern, geographically stratified cities perhaps the most important expenditure related to child outcomes is housing expenditure, as this provides access to locationally-specific factors such as schooling and peer groups."

assume that the education of the parents (an element of Z_{Qt}) is the most important environmental variable influencing the efficiency of the production process. Since parents from chronically poor households are, on average, less educated than parents from not-poor households (Rodgers and Rodgers 1993, pp.44-45), they can be assumed to be less efficient in producing child health. Thus, in addition to facing a tighter budget constraint, parents from chronically poor households may be less efficient in producing a given level of health from the resources at their disposal.

In a comprehensive review of the evidence on the education-health linkage, Cutler and Lleras-Muney (2006) conclude that higher levels of education do in fact lead to different decision making patterns about health behaviours and greater access to health-related information. For example, more educated persons were more likely to quit smoking after the 1964 Surgeon General Report first publicised the dangers of smoking cigarettes (Cutler and Lleras-Muney 2006, p.16).

It may also be the case that beliefs, preferences, innate ability and other unobserved factors differ across households in a way that is correlated with the household's income (Guryan *et al.* 2008, p.14; Currie 2008, p.7). For example, higher income households may derive greater utility from children's health and/or be of higher innate ability, relative to lower income households. These unobserved differences, which are reflected in Z_{Qt}, ε_{Qt}, Z_{Ut} and ε_{Ut}, confound the relationship between family income and child health, as will be discussed in Section 3.4.1.[13]

[13] This line of argument suggests that the financial means to purchase health inputs may not always translate into the actual purchase of such inputs.

2.3.3 Mothers' Employment

From the household production model, it is evident that there is a conflict in time use borne out by the fact that a mother is both a caretaker and economic provider for the child. That is, as a consequence of the time constraint, a mother who decides to work an additional hour will simultaneously increase the family's income by an amount $\$w$ and reduce her total time spent in leisure L_{Lt} and producing child health L_{Qt} by one hour. This can create opposing effects on child health, which Becker (1981) alludes to as the "income effect" and "substitution effect", respectively.

First, the increase in household income resulting from the additional hour allocated to market work (the income-effect) will affect child health positively (ceteris paribus), for the reasons discussed in Section 2.3.1.

However, the *potential* decrease in time available to invest in children may more than offset the increased ability to purchase productive inputs. As outlined previously, parental time L_{Qt} is a 'good' for children, which directly enters the production function. Put simply, a mother who works an additional hour *may* find it optimal to allocate less time to preparing healthy meals and supervising the child's activities.[14] Yet, to the extent that the household is able to substitute the mother's non-market time with market goods, such as pre-prepared foods and after-school care, any detrimental effect may be negated. This economic relationship is referred to as the "substitution effect".

It is most often argued that the substitution effect is detrimental to child health, as market goods are often inferior substitutes for the mother's time. For example, a caretaker or babysitter (or no care at all) is likely to provide inferior supervision of

[14] Using time use survey data for the US, Cawley and Liu (2007) demonstrate that a mother's employment outside the home leads to a reallocation of time away from activities at home, such as cooking and playing with children. They also find that the decreased time spent with children is only partly offset by husbands and partners.

the child, relative to the child's own mother, and convenience foods often contain higher levels of saturated fats, sugars and salt relative to foods prepared by the child's mother. In saying that, it could also be argued that caretakers and day care providers are able to provide healthier snacks and more structured exercise routines than the parents might provide (Fertig *et al.* 2006, p.3). Particularly for older children, it is unclear *a priori* how an increase in parental time tilts the balance between energy intake and expenditure. For this reason, the substitution effect is not unambiguously deleterious to health.

In addition, employed mothers may act to "protect" the most productive time by cutting back least on activities that directly engage the child as their work hours increase (Ruhm 2004, p.4; Cawley and Liu 2007, p.4). That is, a mother may cut back most on leisure L_{Lt} and least on health production L_{Qt} as market work L_{Wt} increases. Further, some of the decreases in mother's time may be offset by increases in time by the father or partner.

In sum, the existence of competing income and substitution effects suggests a potential trade-off between the positive effects of direct parental time investments in children and the positive effects of household income. *A priori*, it is unclear which effect will dominate, and hence the overall (net) effect of maternal employment on child health is an empirical question. Studies for young children (aged under 12 years) have generally uncovered adverse effects of employment on health (Anderson *et al.* 2003; Ruhm 2004; Liu 2006; Chia 2008; Phipps *et al.* 2006; Zhu 2007), suggesting that, on average, the negative effects of decreased maternal time investments outweigh the beneficial effects of any increments in income. However, this relationship has not yet been investigated for older children.

2.4 'Good Parent' Theories and the Sociological literature

The psychological and sociological literatures provide complementary mechanisms through which income and market work may affect children. The so-called 'good parent' theories, outlined in Mayer (1997, pp.48-54), focus on the *quality* of parental time to indicate alternative (but reinforcing) mechanisms by which family income can influence child outcomes.

The 'parental-stress' version of the good parent theory argues that poverty is stressful for parents and this "diminishes parent's ability to provide "supportive, consistent and involved parenting"" (Mayer 1997, p.48, quoting McLoyd 1990). This can interact with, and reinforce other "risk factors" such as poor health, poor environment and low birth weight. The 'parental-stress' theory holds that transferring income to poor families will alleviate stress, and thus improve parenting and children's outcomes.

Conversely, the 'role-model' variant of the good parent theory maintains that transferring income to poor families will *not* improve children's life chances in the short-run. In essence, this theory maintains that parents with low-incomes develop values, norms and behaviours that are "dysfunctional" for success in the wider society, and that these characteristics are transmitted to their children: and hence they represent "bad" role models for their children (Mayer 1997, p.50). Accordingly, parents' values, attitudes and behaviour must be changed, a process requiring a permanent change in the opportunity structure, not simply income transfers to poor households.[15] In fact, this theory suggests that income transfers can have negative impacts on child outcomes to the extent to which they reinforce a culture of dependence (Mayer 1997, p.51).

[15] This theory is closely related to the 'culture of poverty' hypothesis and has been politically controversial (Mayer 1997, p.51).

The sociological and psychological literatures also provides additional mechanisms through which mothers' employment may affect children, including reductions in the quality of mother-child interactions and disruption of mother-child attachments (Huston and Aronson 2005; Stifter *et al.* 1993); a weakening of social capital (Coleman 1988); and "role model" effects (Haveman and Wolfe 1995), though this literature has typically not focused on obesity outcomes.

2.5 Permanent Income Hypothesis

This section provides a brief outline of the Permanent Income Hypothesis of consumption, which is the theory underlying the specification of income in this thesis.

Friedman's (1957) Permanent Income Hypothesis suggests that annual income can be decomposed into two components: first, a stable or "permanent" component, and second an unstable or "transitory" component. In each period, an agent will consume a *constant fraction* of his or her lifetime, permanent income, which can be defined as the level of income anticipated or expected over an extended period of time.[16] If, in any given year, annual income is below the agent's level of permanent income, he or she will draw on their past savings or borrow in order to continue consuming at the level of permanent income. Conversely, if annual income is high relative to the level of permanent income, the agent will pay down their debts and/or save rather than consuming more. As a result, the time path of annual income will have no impact on an individual's consumption behaviour. It is the individual's permanent income, as determined by their level of real wealth – both human wealth and non-human wealth – that affects consumption (Romer 2006, p.348).

[16] The fraction of permanent income consumed each year will be a function of the prevailing interest rate on borrowing and saving, the ratio of the non-human wealth owned by the household to its total wealth, and other objective factors that affect the household's expectations (Romer 2006, p.349).

This hypothesis suggests that fluctuations in parental incomes will not affect children's health and well-being because they will not lead to fluctuations in consumption. Hence, two otherwise identical children, each with different time paths of parental income, but the same level of permanent income should have identical stocks of health, as the time path of consumption (including market goods used in the production of health) would be the same for both children, equalling a constant proportion of their permanent income.

However, the permanent income hypothesis is not without criticism. In particular, the assumption that agents are able to make inter-year income transfers, particularly for those with very little real wealth, has been questioned. It is more likely that some agents will experience liquidity constraints due to capital market imperfections that arise from the fact that future labour income is both uncertain and difficult to borrow against (Mayer 1997, p.72). Romer (2006, pp.374-5) shows that in the presence of binding liquidity constraints, an agent will consume less than they ordinarily would, given perfect capital markets.

Irrespective of the shortcomings, the specification of 'permanent income' in empirical models of child health production is common practice in the literature (Taylor *et al.*, 2004; Phipps *et al.* 2006; Case *et al.* 2002; Blau 1999; Mayer 1997) and in the general poverty literature (Rodgers and Rodgers 2008). The methodological advantages for this approach are discussed in the following chapter.

3 Review of Methodologies

This chapter reviews the methodological approaches used in the literature to estimate the effects of household income and maternal employment on child outcomes. Section 3.1 reviews the methodological conventions in the literature, including least-squares and binary dependent variable regression models and the measurement of excess body weight. The measurement of household income and maternal employment is outlined in Section 3.2, and the issues associated with jointly estimating these variables are discussed in Section 3.3. Section 3.4 outlines the causes and consequences of endogeneity in the variables of interest. Section 3.5 concludes with a review of methods used to address the potential endogeneity of household income and maternal employment.

3.1 Methodological Conventions

3.1.1 Multivariate Regression for BMI

In the literature, the simplest approach to measuring the association between household income and/or maternal employment and child weight outcomes is a multiple linear regression of the form

$$y_i = \alpha + \delta \mathrm{I}_i + \gamma \mathrm{E}_i + \mathbf{x}_i \boldsymbol{\beta} + u_i \qquad i = 1, \ldots, n \qquad (3.1)$$

where y_i is a continuous dependent variable measuring the body mass index (BMI) of the i^{th} child, I_i is current or lagged-average household income, E_i is current and/or lagged maternal employment, $\mathbf{x}_i$ is a vector of exogenous child and household characteristics and u_i is a disturbance.[17] BMI is defined as weight (in kilograms) divided by height (in metres), squared (WHO 2000, p.8), with the estimated δ and γ giving the effect of household income and mothers'

[17] This type of specification has been used in Anderson *et al.* (2007), Baum and Ruhm (2007), Monheit *et al.* (2007), Wake *et al.* (2007) and Chia (2008).

employment on children's BMI, respectively. The complications in estimating these parameters consistently are discussed in Section 3.4.

3.1.2 Binary Overweight and Obese

Since children and adolescents experience gradual changes in body composition that depend on their age and gender (i.e. during puberty), BMI is less appropriate as a measure of excess weight than in adults (WHO 2000, p.9). For this reason, studies using children also specify the dependent variable as a binary outcome set equal to one if the i^{th} child is overweight (or obese); and zero otherwise (Chia 2008; Baum and Ruhm 2007; Zhu 2007; Booth *et al.* 2001). The advantage of a binary-outcome specification is that the thresholds used to classify a child as overweight or obese (based on their BMI) can be constructed to account for age and gender-specific sensitivities in body composition. Notably, Cole *et al.* (2000) developed age- and gender-specific cut-off points for children aged 2-18 corresponding to the adult overweight threshold of 25kg/m^2 and adult obesity threshold of 30kg/m^2. These thresholds, which are presented in Table 3.1 for adolescents aged 15-18 years, have been accepted as the international standard for measuring obesity in children and adolescents and have been widely used in Australian studies of child and youth obesity (Wake *et al.* 2007; Zhu 2007; Booth *et al.* 2001; Access Economics 2006; Li *et al.* 2007).[18]

[18] These definitions have been adopted by the International Obesity Task Force (IOTF), the Australian Department of Health and Ageing and the Longitudinal Survey of Australian Children (LSAC). These definitions have not yet been adopted in the HILDA survey, which continues to categorise overweight and obesity in respondents 15-18 years according to the adult cut-offs of 25kg/m^2 and 30kg/m^2, respectively.

Table 3.1. Australian Reference Standard BMI Cut-off Points for Overweight and Obesity: Males and Females, 15 to 18 years

	BMI Equivalent to 25 in Adults		BMI Equivalent to 30 in Adults	
Age (years)	*Males*	*Females*	*Males*	*Females*
15.0	23.29	23.94	28.30	29.11
15.5	23.60	24.17	28.60	29.29
16.0	23.90	24.37	28.88	29.43
16.5	24.19	24.54	29.14	29.56
17.0	24.46	24.70	29.41	29.69
17.5	24.73	24.85	29.70	29.84
18.0+	25.00	25.00	30.00	30.00

Source: Cole *et al.* (2000)

3.1.3 Binary Dependent Variable Specifications

There are three main approaches to developing a probability model for a binary response variable: the linear probability model (LPM), the logit model and the probit model. The simplest is the LPM, which is represented by Equation (3.1) (with a binary outcome), where δ and γ give the effect of household income and mothers' employment on the probability of child overweight (or obese), respectively. However, despite being used often in the literature (Chia 2008; Baum and Ruhm 2007; Zhu 2007) there are several known limitations to using the LPM. These include the fact that the predicted probabilities can nonsensically lie out of [0,1] interval, the limited value of the R-squared statistic as a measure of goodness-of-fit, an error term that is inherently heteroskedastic, and the assumption that the marginal effect of any explanatory variable will remain constant irrespective of the level of that variable (Gujarati 2003, pp.584-8).[19]

Consequently, most studies have used either probit (Scholder 2008; Ruhm 2004; Anderson *et al.* 2003; Monheit *et al.* 2007) or logit (Booth *et al.* 2001; Hofferth and Curtin 2005; Wake *et al.* 2007) regression models, which use s-shaped cumulative probability density functions to overcome the problems associated with the LPM. In these models, the predicted values of the dependent variable are forced to lie

[19] The 'linearity' assumption implies that a given change in household income will affect a child living in a poor household to exactly the same extent as a child living in a rich household, in terms of the probability of overweight/obesity.

between zero and one and the probability of 'overweight' will be non-linearly related to the independent variables, such as permanent income and maternal employment (Stock and Watson 2007, p.389).[20] The probit model used in this thesis is developed in Section 7.2.

3.1.4 Measurement Error in the Dependent Variable

It is widely recognised that BMI is a less than optimal measure of ideal body weight. In particular, it does not distinguish between muscle-mass and fat-mass and can be affected by the distribution of body fat, the body build and certain health conditions, such as pregnancy and osteoporosis (WHO 2000, p.8).[21]

Data on BMI, particularly when it is self-reported, can also be measured with error. Research has shown that self-reported height and weight data can be affected by social desirability biases, in which people tend to overstate their height and understate their weight (Wooden *et al.* 2008). Wang *et al.* (2002b) demonstrated that this pattern of reporting bias led to an incorrect classification of overweight or obesity in around 30 per cent of Australian adolescents aged 15-19 in the 1995 National Health Survey (NHS), and that the extent of the bias was larger in overweight/obese adolescents than normal/underweight adolescents.[22]

The height and weight data in HILDA are also self-reported, and hence a similar level of error is expected in this analysis.[23] However, provided the errors of measurement in BMI are uncorrelated with the explanatory variables and the

[20] The key differences between the logit and probit model relates to differences in the assumption of the error variance, with the errors assumed to be logistically distributed under the logit model and normally distributed under the probit model.

[21] Fat-distribution measures, such as waist circumference or waist-to-hip ratio, are better predictors of mortality risk than BMI (Access Economics 2009, p.9).

[22] Wang *et al.* (2002b) compared self-reported data from the 1995 NHS with measurements collected by trained nutritionists in the 1995 NNS (conducted in conjunction with the NHS).

[23] However, as Wooden *et al.* (2008) argue, the fact that the relevant questions are collected in a self-administered questionnaire, rather than during a personal interview, as is the case in the National Health Survey (NHS), may give respondents time to actually measure themselves and lead to lower rates of item refusal (Wooden *et al.* 2008).

error term in Equation (3.1), the estimates of δ and γ should still be unbiased, albeit with larger estimated standard errors than in the case where there are no such errors of measurement (Gujarati 2003, p.525).

3.2 Measurement of Key Explanatory Variables

3.2.1 Household Income

Household income is generally measured contemporaneously with child outcomes (Baum and Ruhm 2007; Anderson *et al.* 2007; Monheit *et al.* 2007; Wake *et al.* 2007), particularly when the data are cross-sectional. However, as argued by Mayer (1997, p.63) and Blau (1999, p.265), estimates of income-effects relying on a single year of parental income are likely to understate the impact of permanent income on children, since annual income also includes a 'transitory' component and may be measured with error.[24] For this reason, studies using longitudinal data typically use a measure of permanent household income, calculated as the arithmetic average of real household income over the sample considered (Taylor *et al.* 2004; Phipps *et al.* 2006; Case *et al.* 2002; Blau 1999; Mayer 1997).[25]

In addition, household income is often 'equivalised' to account for differences in living needs between families of different size and composition, using an appropriate equivalence scale (Phipps *et al.* 2006; Wake *et al.* 2007; Currie *et al.* 2007).[26] Likewise, US studies generally divide household income by some poverty

---

[24] Even if current consumption tracks current income (i.e. the permanent income model is not valid) a measure of permanent income is still justified on the grounds that all current and lagged inputs belong in the structural equation (see Section 2.3) and that children's health may respond with a lag to short-term income fluctuations (Case *et al.* 2002, p.1323).

[25] Notably, Mayer (1997) uses average family income over three, five and ten years prior to assessment, Anderson *et al.* (2003) and Case *at al.* (2002) use average family income since the child's birth, Korenman *et al.* (1995) use a 13-year average of income and Currie *et al.* (2007) used a six year average income.

[26] To accommodate a non-linear effect of income, studies have specified income in logarithmic form (Doyle *et al.* 2007; Shea 2000; Duncan *et al.* 1998; Mayer 1997; Maurin 2002; Khanam *et al.* 2009); including quadratic terms in the regression model (Doyle *et al.* 2007; Bradbury 2007); using separate dummy variables for income quartile (Duncan *et al.* 1998; Classen and Hoyakem 2005); interacting income with a poverty dummy variable (Taylor *et al.* 2004); estimating models by

threshold that depends on household size and composition – giving what is known as the "poverty-ratio" or the "income-to-needs ratio" (Korenman *et al.* 1995; Anderson *et al.* 2007; Monheit *et al.* 2007; Hofferth and Curtin 2005). A discussion of equivalence scales is given in Section 5.4.2.

3.2.2 Maternal Employment

Like household income, longitudinal studies typically measure maternal employment as an average of its current and lagged levels (Anderson *et al.* 2003, Ruhm 2004; Chia 2008; Scholder 2008; Phipps *et al.* 2006). As Ruhm (2004, p.7) asserts, if the mother's work *history* is not controlled for, and if maternal employment is positively serially correlated, then the coefficient on contemporaneous employment can be biased as it may pick-up part of the effects of previous employment. This relates to the theoretical model developed in Section 2.3, which shows that all current and lagged time inputs in the structural equation; hence excluding lagged employment may constitute omitted variable bias to the extent that lagged employment is correlated with current employment.[27]

Maternal employment itself is generally measured in terms of hours of paid employment per week or month (Anderson *et al.* 2003; Ruhm 2004; Chia 2006; Phipps *et al.* 2006), or using dummy variables for part-time and full-time employment; with non-working mothers as the omitted category (Liu 2006; Scholder 2008; Classen and Hokayem 2005; Garcia *et al.* 2006; Takahashi *et al.* 1999; Zhu 2007).

---

sub-groups defined on the basis of SES (Taylor *et al.* 2004; Blau 1999; Hofferth and Curtin 2005); and estimating spline functions that allow separate slopes for children in families with household incomes under and above a certain level, such as the poverty threshold (Duncan *et al.* 1998).

[27] Anderson *et al.* (2003, p.482) argue that contemporaneous employment may fluctuate, and with it the flows of calories expended and consumed, but it is actually the 'stock' of net calories that will determine a child's weight status. This 'stock' concept can be better approximated using the child's lifetime exposure to maternal employment.

3.3 Joint Estimation of Household Income and Maternal Employment

An important issue in estimating Equation (3.1) is that household income and maternal employment are potentially jointly chosen variables in the child-health production model (Taylor *et al.* 2004; Blau 1999; Zhu 2007; Ruhm 2004; Cawley and Liu 2007; Blau and Grossberg 1992). However, most studies ignore this issue and persist with the joint estimation of household income and maternal employment (Anderson *et al.* 2003; Phipps *et al.* 2006; Korenman *et al.* 1995).

As Blau (1999, p.263) argues, if the objective is to measure the "full" or "policy-relevant" effect of income, the specification should exclude variables that are jointly chosen with, or caused by, income, such as maternal employment. The reason is that such variables will likely change in *response* to changes in income and hence should not be held fixed. In recognition of this, Blau (1999) and Taylor *et al.* (2004) exclude maternal employment from most specifications.

For similar reasons, holding household income fixed will not yield a measure of the "full" effect of maternal employment, since part of the labour supply effect may work through increased income, as discussed in Chapter 2 (Cawley and Liu 2007, p.8; Zhu 2007, p.16).

To estimate the full effect of mothers' employment, Blau and Grossberg (1992, p.479) control for household income *excluding* the mother's wages and salary income. Here, the coefficient on mothers' employment will measure the sum of the direct effect of employment, controlling for any increment in income (the "substitution effect"), and the indirect effect due to the increase in income (the "income effect").

The issue, however, is that a mother's earned income is likely to be reflective of her genetic endowments, human capital, motivation, occupation and other factors potentially correlated with employment and child health outcomes. Hence, excluding this information from the model may exacerbate the problem of omitted variable bias. For this reason, Blau (1999, p.268) and James-Bardumy (2005, p.185) include the mother's *wage rate* as a separate explanatory variable, in an attempt to capture some of the information lost after excluding maternal wages and salary from the model. However, as non-working mothers do not have an observable wage, this specification can only be estimated over the sub-sample of women in paid employment, and these women may not be representative of *all* mothers.

3.4 Establishing Causality: the Problem of Endogeneity

The objective of this study, and much of the empirical literature, is to identify a causal relationship between family income and/or maternal employment and the outcome of interest. However, establishing causality is complicated by potential endogeneity in the treatment. That is, one or more of the explanatory variables may be correlated with the disturbance term, leading to biased and inconsistent parameter estimates (Wooldridge 2002, p.50).

Concern about the potential endogeneity of household income (Blau 1999, Taylor *et al.* 2004, Doyle *et al.* 2007; Duncan *et al.* 1998; Dahl and Lochner 2005) and maternal employment (Anderson *et al.* 2003, Ruhm 2004, Zhu 2007; Liu 2006; Chia 2008; Scholder 2008) has been expressed in the literature. In particular, three potential sources of endogeneity have been identified: omitted variables (unobserved heterogeneity), simultaneous causality and "errors-in-variables". These sources of bias are discussed in the following section, while Section 3.5 outlines the econometric approaches that have been used in the literature to account for endogeneity.

3.4.1 Unobserved Heterogeneity

Unobserved heterogeneity is a problem if household income and/or mothers' employment are correlated with unobserved factors that directly affect child BMI and obesity. For instance, unobserved maternal ability may be positively correlated with mothers' hours in paid employment and household income (via wages and hours) and negatively correlated with children's BMI (via increased productivity in the home). In this case, the estimated coefficient on household income and mothers' employment will be biased downward, since the confounding effects of maternal ability will be (incorrectly) attributed to household income and maternal employment.[28] Conversely, market ability and health production efficiency may be negatively correlated, if for instance, women who work have less interest or skill in raising children, *ceteris paribus*. For this reason, and since other confounding factors are also unobserved (parental altruism, preferences, rate of time preference etc.), the direction of the bias is unknown *a priori*.

3.4.2 Reverse Causality

Household income may also be endogenous if there is a simultaneous determination of family income and health. For studies using adults, this problem is particularly acute, since income improves access to health inputs, such as medical care, and health improves a person's ability to participate in the labour market and earn a high wage. In contrast, children in developed countries do not generally contribute substantially to household income, and hence any observed correlation between household income and child outcomes cannot be attributed to the earnings of children. As this largely mitigates the channel running from health to household income, studies using children have been more liberal in interpreting the association as causal (Case *et al.* 2002; Currie and Stabile 2003).

[28] In this discussion, if the outcome was 'health' more generally, the direction of the bias would be upward.

Notwithstanding, reverse causation will present a problem if mothers reduce their work hours to provide better care for overweight/obese children in poor health. In this case, the effect of household income (Baum and Ruhm 2007, p4; Ruhm 2004; Currie *et al.* 2007; Doyle *et al.* 2007; Khanam *et al.* 2009) and maternal employment (Ruhm 2004; Scholder 2008) on BMI and obesity will be biased *downward*.[29]

3.4.3 Measurement Error

Finally, household income may be endogenous due to "errors-in-variables", particularly when single-year measures of income are used. If an imperfect measure of income is observed, the least squares estimates of income effects will be biased toward zero, a problem known as 'attenuation bias' (Wooldridge 2002, p.75).[30] These errors of measurement can be reduced by averaging household income over a number of years, as this yields a less noisy measure of income and a closer approximation of the 'permanent' component of household income.

3.5 Methods for Addressing Endogeneity

There have been several approaches used in the literature to estimate causal effects, in the presence of endogeneity in the treatment. These include experimental designs, proxy variables, instrumental variables and fixed-effects. Each of these is discussed in the following section.

[29] Yet, the resulting bias is likely to be substantially smaller than using a sample of adults, and this simultaneity problem is largely, but not entirely, avoided using lagged measures of income.

[30] This will be the case under the 'classical errors-in-variables assumption', in which the random errors of measurement are uncorrelated with 'true' income and the disturbance term of the equation of interest (Wooldridge 2002, p.75).

3.5.1 Random Assignment to the Treatment

In a randomised controlled experiment, both household income and maternal work intensity can be randomly assigned, allowing the causal effect of these variables to be measured.[31]

However, very few studies have been able to examine the effects of exogenous changes in household income on child outcomes, due to the difficulty in identifying exogenous sources of income variation that do not also (directly) affect other conditions in the household (Doyle 2007, p.9).[32] The only study examining the effect of exogenous changes in *lump-sum* income on child outcomes in the context of a natural experiment is that of Costello *et al.* (2003), who measured the mental health and developmental outcomes of Native American Indian children before and after the opening of a casino which resulted in substantial lump-sum income transfers being made to their parents.[33] This study and others that have examined the effect of experimental and non-experimental welfare reforms in the United States (Chase-Lansdale *et al.* 2003; Morris and Gennetian 2003; Dahl and Lochner 2005) have mostly found modest beneficial effects of income on child outcomes.

3.5.2 Controls for Unobservable Heterogeneity: Proxy Variables

In the absence of a random assignment, a comparison of the treated with the untreated groups must control for differences in observable and unobservable characteristics. To account for the latter, several studies have relied on proxy

---

[31] This would involve, for instance, dropping money on the doorstep of randomly selected parents and then tracking the subsequent health outcomes of their children, along with a randomly selected control group which does not receive the money.

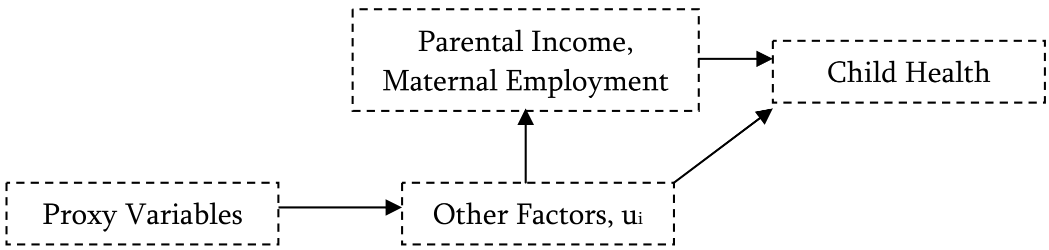
[32] One possibility mentioned is welfare-to-work experiments. However, these programs increase both income and the incentive for mothers to work, thus do not provide information on the causal effects of lump-sum income transfers (Doyle 2007, p.9).

[33] In Costello *et al.* (2003), both the treatment and control groups (the children of other non-Native American poor parents in the district) benefited from the improvement in labour market conditions associated with the opening of the casino. This potentially contaminates the experiment.

variables for the unobservables (Ruhm 2004; Gregg *et al.* 2005; Scholder 2008; Mayer 1997).

Valid proxy variables have no direct effect on child health, but are correlated with unobservable characteristics of the mother and child that do directly influence the health of the child (Figure 3.1). Including proxy variables in the model can reduce (or even eliminate) the conditional correlation between I_i and/or E_i and the unobserved factors (Gregg *et al.* 2005, p.F57).

Figure 3.1. Path Diagram for a Proxy Variable



A proxy variable used in the literature is the mother's pre-pregnancy occupation and hours of work (Gregg *et al.* 2005; Ruhm 2004). As Ruhm (2004, p.7) argues, pre-birth employment characteristics can supply information on the mother's "tastes for employment" and "opportunity costs of not working" which may be correlated with the unobserved influences on child health outcomes.[34] [35]

In addition, Mayer (1997) uses *future* family income (measured after child assessment) as a proxy for stable unobserved family characteristics, such as

[34] As proxies for maternal attitudes towards parenting and the quality of the mother's time investments, Gregg *et al.* and Ruhm also use characteristics of the mother's background and home environment. These include the presence of the mother's own mother during her childhood, the degree of stimulation in her childhood home (i.e. the presence of magazines, newspapers), and whether her own mother was in paid employment at age 14.

[35] The proxy variables used in the literature generally measure characteristics of the mother *prior* to the birth of the child, as these measures are necessarily predetermined and hence unaffected by the mother's current employment patterns and income (Ruhm 2004; Gregg *et al.* 2005). The notable exception is the mother's Armed Forces Qualification Test (AFQT) score, which is the most common proxy for maternal ability in studies for the US, and is itself, arguably predetermined (Anderson *et al.* 2003; Blau 1999; Taylor *et al.* 2004; Ruhm 2004).

parental intelligence, competence and motivation; arguing that future earned income should not have a direct influence on children's outcomes but should be correlated with these stable parental characteristics.[36] By the same logic, Ruhm (2004, p.4) and Scholder (2008, p.902) control for maternal employment in the period *after* child assessment, asserting that future employment (which is unlikely to have casual effects on child outcomes in a prior period) can serve as a proxy for the mother's unobserved "tastes" with respect to her working status. This specification also provides a test for the presence of reverse causality, since large or significant parameter estimates on future employment may provide some evidence for simultaneous causality, and hence, model misspecification.[37]

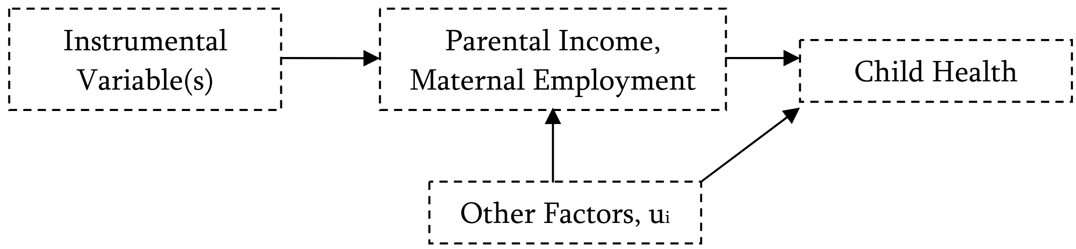
3.5.3 Instrumental Variables (IV)

IV methods have been used to account for endogeneity in both household income (Doyle *et al.* 2007; Maurin 2002; Shea 2000; Case *et al.* 2002; Currie *et al.* 2007) and maternal employment (Anderson *et al.* 2003; James-Bardumy 2005; Cawley and Liu 2007; Zhu 2007). In theory, IV estimation can be used to control for unobserved heterogeneity (fixed or variable), and for measurement error bias (Anderson *et al.* 2003, p.484). This approach requires the availability of at least one instrumental variable for each endogenous regressor – a variable that is correlated with household income and/or mothers' employment ("relevant") and uncorrelated with the error term in the child-outcome equation ("exogenous") (Verbeek 2008, p142; Stock and Watson 2007, p423). Figure 3.2 depicts the path diagram for a valid instrumental variable.

[36] However, as Mayer (1997) points out, if families anticipate future income changes and adjust their consumption accordingly, then future income can indeed have a causal effect.

[37] For example, if child health problems cause mothers to cut back work hours in future periods, a positive coefficient would be expected.

Figure 3.2. Path Diagram of an Instrumental Variable



If variables satisfying the relevance and exogeneity conditions can be identified, the exogenous variation in these variables can be used to isolate that part of the endogenous regressor(s) which is uncorrelated with the error term, and hence obtain consistent estimates of the population coefficients (Stock and Watson 2007, p.423). The instruments used in the literature are outlined below. Full treatment of IV methods, particularly two-stage least squares (2SLS) and IV-probit, is left until Section 7.3 and 7.4.

3.5.3.1 Instruments for Household Income

Table 3.2 summarises the instruments for family income used in the child health literature. In most cases, these relate to the SES of the child's grandparents and/or parents.[38] For example, Doyle *et al.* (2007, p.17) used indicator variables for whether grandparents smoked cigarettes, assuming that grandparental smoking is associated with lower parental income and that this does not affect child health outcomes once other factors that affect child health are controlled for.[39] However, as Currie (2008, p.20) points out, the assumption that grandparents' smoking (an indicator of grandparent SES) affects grandchildren's outcomes only indirectly through parental income is a strong one.

[38] Note that lagged values of income are not valid instruments because they belong in the structural equation.

[39] Doyle *et al.* (2007) also include a set of dummy variables accounting for whether the parents smoked before the age of 16, between the age 16 and 19, or after age 19. Again, they assume that these do not affect health directly once they control for other child health determinants, which include parental smoking and the number of years the child has been exposed to parental smoking.

Table 3.2. Instrumental Variables for Household Income

Study	Endogenous Regressor	Instrument	Findings
Doyle *et al.* (2007)	Log income (and squared)	● grandparents smoked ● parents smoked before age 16; between 16 and 19; after 19 years ● increase in compulsory schooling age	- IV estimates > estimates assuming exogeneity. - Instruments found to be relevant and exogenous.
Maurin (2002)	Log income	● grandfathers' occupational status score ● fathers' sector and employment status	- IV (GMM) estimates 3 times > OLS estimates. - Instruments found to be exogenous.
Shea (2000)	Log permanent income	● fathers' union status ● fathers' average industry wage premium[1] ● father lost job due to plant closure	- IV (2SLS) estimates < OLS estimates and insignificant. - Instruments found to be relevant and exogenous. - Income found to be endogenous.
Case *et al.* (2002)	Log income + log permanent income	● industry of each worker in household ● occupation of each worker in household ● class of each worker in household	- IV estimates 25-50% > estimates assuming exogeneity. - First-stage results not reported.
Currie *et al.* (2007)	Log income + log permanent income	● industry of each worker in household ● occupation of each worker in household ● class of each worker in household	- IV estimates > estimates assuming exogeneity. - First-stage results not reported.

More questionable instruments used in the literature include the sector and employment status of the child's father (Maurin 2002), the father's industry, union membership, involuntary job loss from plant closure (Shea 2000), and the industry, occupation and class of worker in the household (Case *et al.* 2002; Currie *et al.* 2007).[40] It is harder to argue that these instruments are uncorrelated with the unobserved determinants of child health, particularly parents' unobserved ability, and hence, the "income effects" in these studies may still be biased (Wooldridge 2002, pp.101-2).[41]

3.5.3.2 Instruments for Maternal Labour Supply

The maternal labour supply literature has tended to use local economic conditions as instruments (see Table 3.3). For instance, Anderson *et al.* (2003, p.485) use the variation over time and between states in the unemployment rate, child care regulations, wages of child care workers, welfare benefit levels and the status of

[40] Maurin (2002, p.308 footnote 12) cites several studies finding that inter-industry wage differentials only reflect inter-industry differences in workers' unobserved abilities. Similarly, Shea (2000, pp.2-3) cites several studies finding that industry and union wage premia mostly reflect rents rather than unobserved ability differences. In addition, Shea makes the case that plant closures (which lead to large and persistent negative effects on earnings) are likely to be exogenous with respect to employees' unobservable skills.

[41] Currie *et al.* (2007, p.222) expressed concern about the validity (exogeneity) of these instruments.

welfare reform.[42] However, as Anderson *et al.* point out, these instruments exhibit little variation, and hence, are inefficient. Similarly, weak instruments were an issue in James-Burdumy's (2005) IV analysis, which used county per capita income, the county unemployment rate and the percentage of the county labour force employed in services as instruments. Only one study has found the state unemployment rate to be strongly predictive of mothers' employment (Cawley and Liu 2007, p.7).

Table 3.3. Instrumental Variables for Maternal Employment

Study	Endogenous Regressor	Instrument	Findings
Anderson *et al.* (2003)	Average hours per week + average weeks per year	● state unemployment rate ● child care regulations in the state[1] ● wages of child care workers in state ● state welfare benefit levels ● status of welfare reform in the state	- IV (2SLS) estimates similar to OLS and fixed-effects estimates but insignificant. - Instruments found to be weak.
James-Bardumy (2005)	Average hours per week + average weeks per year.	● county per capita income ● county unemployment rate ● percentage of county labour force employed in services	- IV-Fixed-effects estimates had larger standard errors than FE estimates, and thus rejected. - Unemployment rate the only relevant instrument. - Instruments not exogenous.
Cawley and Liu (2007)	Employed (dummy) + hours per week (if employed)	● state unemployment rate	- IV estimates had larger standard errors than OLS estimates. - Instruments relevant. - Employment found to be exogenous (Hausman test).
Zhu (2007)	Average hours per week + non-working, part-time, full-time	● English was mother's first language ● mother does volunteering work	- IV (2SLS) estimates 6 times > OLS estimates. - Employment found to be endogenous. - Instruments relevant and exogenous.

Note 1: Child care regulations include requirements for liability insurance, staff training, number of facility inspections and the caregiver/child ratio.

Instead of using local economic conditions, Zhu (2007, p.14) used two binary instruments: whether English was the first language the mother was exposed to and whether the mother participates in any type of volunteering work. Unlike the geographic-based variables used in previous research, these maternal and familial variables are more likely to be correlated with the error term in the structural equation, as they are not external to the mother. In particular, volunteer work is a choice of the mother which cannot be separated from the mother's employment decision, whilst English as a first language may influence the mother's information

[42] Higher unemployment makes it more difficult for mothers to find work; child care regulations and higher child care wages may reflect higher costs and less utilization; and women in states with more restrictive welfare rules may be more likely to work (Anderson *et al.* 2003, p.485).

on the benefits of exercise and nutrition and/or be correlated with the unobserved determinants of child health.[43]

3.5.3.3 Findings Using IV

IV estimates of household income are typically found to be larger than estimates assuming exogeneity (Doyle *et al.* 2007; Maurin 2002; Case *et al.* 2002; Currie *et al.* 2007), albeit less precise. For maternal employment, studies have generally failed to reject the hypothesis that maternal employment is exogenous (Anderson *et al.* 2003, p.491; Cawley and Liu 2007; Garcia 2006, p.15), implying little, if any bias due to endogeneity. The exception was Zhu (2007, p.18), for which the IV estimates were six times larger than the corresponding OLS estimates.

3.5.4 Fixed-Effects Specifications

A common approach to accounting for unobserved heterogeneity is fixed-effects (Blau 1999; Sabia 2007; Anderson *et al.* 2003; Ruhm 2004; Scholder 2008; Chia 2008; Duncan *et al.* 1998; Dahl and Lochner 2005).[44]

Individual (child) fixed-effects have been used in the literature to account for unobservable, time-invariant characteristics of the child and family (Blau 1999; Dahl and Lochner 2005; Anderson *et al.* 2003; Scholder 2008). However, to difference out the fixed-effect this approach requires two or more observations on each child over time, and hence, this approach is not feasible in this analysis. The

[43] Other researchers have failed to devise adequate instruments (Ruhm 2004, Blau *et al.* 1996, Chia 2008, Scholder 2008; Khanam *et al.* 2009). In the case of Blau *et al.* (1996) and Khanam *et al.* (2009), the data they used was not rich enough to locate enough excluded exogenous variables to identify IV estimates of the health production function. Other studies have failed to find instruments that satisfy the 'relevance' criteria, such as Chia (2008, p.227), who experiments with variables including indicators for provincial childcare policy, provincial unemployment rates, and actual and predicted maternal wages, finding that these variables were not strong predictors of maternal work hours.

[44] Fixed-effects specifications are robust to endogeneity as long as the fixed-effect is the only source of correlation between the explanatory variables and the error term (Blau 1999, p.266). Fixed-effects cannot account for bias resulting from simultaneity or measurement error. In fact, fixed-effects will magnify any bias due to measurement error in permanent income or maternal labour supply, as growth rates are more noisily measured than levels (Dahl and Lochner 2005, p.6).

reason is that the panel is too short to construct two measures of permanent income with sufficient variation to obtain precise estimates of the effect of permanent income.[45]

Sibling fixed-effects have also been used to control for fixed unobserved genetic and/or environmental influences specific to the mother and household (Anderson *et al.* 2003; Blau 1999; Duncan *et al.* 1998; Ruhm 2004; Chia 2008). These models difference out the family fixed-effect by comparing the outcomes of siblings - either at the same point in time or at the same age. Since this does not require multiple observations over time, this is one of the methods adopted in this thesis. Full details of the sibling-difference model are provided in Section 7.5.

[45] As Cameron and Trivedi (2009, p.251) explain, the fixed-effect estimator will be relatively imprecise for time-varying regressors that vary little over time. Moreover, the data used in this thesis contain at most two BMI observations for each child, which exhibit little variation due to the close proximity in which they were measured (Wave 6 and Wave 7 in HILDA). It is also unknown as to what extent this variation reflects true changes in adiposity or simply random fluctuations due to measurement error.

4 Review of Empirical Literature

Following the review of methodologies in Chapter 3, this chapter summarises the empirical evidence on the relationship between household income, mothers' employment and child/adolescent outcomes. Section 4.1 and Section 4.2 review the evidence on the "income gradient" in children's overall health and developmental outcomes, respectively, to place this study into context. In Section 4.3, the literature specific to child and adolescent overweight and obesity is reviewed, giving attention to the effects of both household income and also SES more generally. The chapter concludes by summarising the evidence on the relationship between maternal employment and child and adolescent overweight problems.

4.1 Household Income and Parent-Reported Child Health

There is a growing international literature on the relationship between household income and the subjective, parent-reported overall health status of children.[46] The majority of studies document a positive relationship between family income and children's health, with larger effects for permanent income than for current income (Case *et al.* 2002; Currie and Stabile 2003; Doyle *et al.* 2007; Proper *et al.* 2007), or at least a similar effect (Currie *et al.* 2007). This literature is summarised in Table 4.1.

[46] In these studies, the dependent variable (health) is measured on an ordinal scale (1=very good, 2=good, 3=fair, 4=bad, 5=very bad), with an ordered-probit model used to estimate the relationship.

Table 4.1. Income Effects on Parent-Reported Child Health

Dependent Variable = Parent-Reported Health Status of the Child (1=excellent, 5=poor)

Study	Country; Data	Ages of Children	Income Measure	Specification and controls	Findings
Case *et al.* (2002)	United States; National Health Interview Survey, PSID, NHANES	0-3; 4-8; 9-12; 13-17 years	Log of pre-tax family income + log of permanent family income	Ordered Probit. Controls for parents' education and employment, household size and composition, race, insurance, health at birth, genetics.	Significant and large effects of income with greater effects for older children. Stronger effects for permanent income.
Currie and Stabile (2003)	Canada; National Longitudinal Survey of Children and Youth	0-3; 4-8; 9-12; 13-15 years	Log of pre-tax family income + log of permanent family income	Ordered Probit. Similar controls to Case *et al.* (2002)	Significant and large effects of income with greater effects for older children. Stronger effects for permanent income.
Currie *et al.* (2007)	England; 1997-2002 Health Surveys of England, British Household Panel Survey	0-3; 4-8; 9-12; 13-15 years	Log of pre-tax family income + log of permanent family income [1]	Ordered Probit. Similar controls to Case *et al.* (2002) + child nutrition and family lifestyle factors.	Significant but modest effects of income. No change in the gradient with child age. Similar effects using permanent income. Inclusion of supplementary controls (nutrition & lifestyle) did not reduce size of income gradient.
Proper *et al.* (2007)	England; Avon Longitudinal Study of Parents and Children	0-7 years	Financial hardship indicators + disposable family income	Ordered Probit. Similar controls to Case *et al.* (2002) + maternal physical and mental health.	Small effects of income. No change in the gradient with child age. Stronger effects for permanent income but effect disappears after controlling for maternal health.
Doyle *et al.* (2007)	England; 1997-2002 Health Surveys of England	2-15 years	Log of total pre-tax family income + Income squared	Ordered Probit + IV-Probit. Similar controls to Case *et al.* (2002) + parental smoking indicators.	Significant but modest effects of income. No change in the gradient with child age. Stronger effects treating income as endogenous.
Khanam *et al.* (2009)	Australia; Longitudinal Survey of Australian Children (two waves)	0–3 and 4–8 years	Log of pre-tax family income + Log of permanent income (2 year average income)	Ordered Probit. Similar controls to Case *et al.* (2002) + child nutrition, parental physical and mental health and parental-related behaviours and lifestyle.	Significant effects of income with greater effects for older children. Income gradient disappears after including controls for parental health and child's nutrition.

Note 1: Results are robust to whether income was transformed to a *per adult equivalent* basis using the McClements family equivalence measure of income (see Lillard *et al.* 2009, p2-117).

The seminal paper by Case *et al.* (2002) used data from the US and found robust evidence of a positive effect of household income on child health, after controlling for parental education and other observable child and household characteristics. Importantly, the income 'gradient' in health was found to increase (steepen) as individuals move from infancy through late adolescence. That is, the adverse effects on health of low permanent income were found to accumulate over a child's life.[47]

The methodology of Case *et al.* (2002) was subsequently applied to pooled data on Canadian (Currie and Stabile 2003) and English children (Currie *et al.* 2007; Proper *et al.* 2007), with all studies finding evidence of a positive relationship between household income and the general health status of children. However, in contrast to Case *et al.* (2002) for the US and Currie and Stabile (2003) for Canada, the size of the income-health gradient was found to be relatively small in English children, with no evidence to suggest that it increases with the age of the child (Currie *et al.* 2007; Proper *et al.* 2007).[48] [49]

Recently, Khanam *et al.* (2009) investigated this relationship for Australia using two waves of data from the Longitudinal Survey of Australian Children (LSAC). Using similar controls to Case *et al.* (2002), the authors found evidence of an increasing income gradient by child age, albeit with smaller income coefficients than Case *et al.* (2002) and Currie and Stabile (2003). However, introducing controls for parental heath reduced the income coefficient to zero, suggesting that parental health is the potential mechanism by which low income translates into poor child health in Australia (Khanam *et al.* 2009, p.12).

---

[47] A doubling of family income was associated with a 4 per cent increase in the probability of a child aged 0-3 being rated as in excellent or very good health, increasing to 4.9 per cent for 4-8 year-olds, 5.8 per cent for 9-12 year-olds and 7.2 per cent for children aged 13-15 years

[48] More recently however Case *et al.* (2008) have re-examined the US and English data, comparing similar data (in terms of time periods). They find using similar years and methods reduces differences between the two countries.

[49] Slightly *larger* income effects on the health of English children were found by Doyle *et al.* (2007) using instrumental variables to account for endogeneity.

4.2 Household Income and Child Development

There is also a well-developed literature on the effect of household income on the developmental outcomes of children and adolescents. Table 4.2 summarises this literature, with a focus on studies that have explicitly accounted for endogeneity. As Table 4.2 shows, studies have generally found that the effects of annual income on child outcomes is small and statistically significant and decreases in size as more covariates are included (Blau 1999; Duncan *et al.* 1998; Taylor *et al.* 2004; Dahl and Lochner 2005; Mayer 1997).

Moreover, the magnitude of the income effect diminishes in size when accounting for the endogeneity of household income (Blau 1999; Mayer 1997; Dahl and Lochner 2005; Taylor *et al.* 2004; Shea 2000). Mayer (1997) was one of the first to contend with endogeneity, using a variety of approaches to eliminate biases caused by the omission of unobserved family and child characteristics.[50] In most cases, Mayer found small, insignificant effects of income after accounting for heterogeneity, concluding that once basic needs are met "parental income is not as important to children's outcomes as many social scientists have thought" (Mayer 1997, p.2). Similarly, other studies have found that the effect of income on child-development outcomes is largely or completely attenuated after accounting for unobservables using fixed-effects (Blau 1999; Taylor *et al.* 2004; Dahl and Lochner 2005).

Only two studies have found *larger* effects after accounting for endogeneity. However, the estimated income effects in these studies were either imprecise (Duncan *et al.* 1998) or based on questionable instrumental variables for household income (Maurin 2002). To date, no Australian study has examined the relationship

[50] Mayer used proxy-variables and looked at the *sources* of income which are arguably less correlated with unobserved parental characteristics (i.e. income from inheritance and asset income).

between household income and the developmental outcomes in children using methodologies that account for the endogeneity of income.

Another substantiative finding in this literature is that the income effect is non-linear, with changes in income having larger effects on developmental outcomes within poor households, particularly for changes in permanent income (Taylor *et al.* 2004, p.996; Duncan *et al.* 1998, p.414; Mayer 1997; Korenman *et al.* 1995). However, this finding is not universal, with Blau (1999) finding no evidence of non-linearity in the income effect on motor and social outcomes of infants (0-3 years), and cognitive and language outcomes for young children (3-7 years). Further, Doyle *et al.* (2007) found no evidence of a non-linear effect of income in the parent-reported health literature.

The most consistent finding in this literature is that of larger effects for permanent income than for annual income (Blau 1999; Taylor *et al.* 2004; Dahl and Lochner 2005; Mayer 1997; Kornenman *et al.* 1995), which is also consistent with the literature summarised in Section 4.1. This suggests that persistent poverty is a far stronger predictor of adverse outcomes than is transitory poverty or that measures of permanent income are more highly correlated with unobserved (stable) family characteristics, such as skill and motivation (Mayer 1997, p.87). Notwithstanding, the effects of permanent income are still modest, and Blau (1999, p.261) concludes that policies that affect family income will have "little direct impact on child development unless they result in *very large* changes in permanent income" [italics added].

Table 4.2. Income Effects on Child Development

Study	Country; Data; Age of Children	Income Measure	Outcomes Measure	Specification and controls	Findings
Blau (1999)	United States; NLSY; 0-3 and 3-7 years	Household annual income + mothers' nonwage income and wage rate + permanent income	Motor and social outcomes (0-3 years) cognitive and language outcomes (3-7 years)	OLS + individual, sibling and cousin fixed-effects + random effects. Controls for race, AFQT test score, household size and composition, mothers' age, education and marital status and fathers' education.	Significant and positive effects of annual income, with larger effects for permanent income. Insignificant effects of annual income after accounting for heterogeneity (permanent income still significant but modest).
Duncan *et al.* (1998)	United States; PSID; Up to 20 years	Permanent income (average pre-tax family income since birth)	Years of completed schooling + timing of first non-marital birth	OLS + sibling fixed-effects. Similar controls to Blau (1999)	Significant but modest protective effects of income, with larger income effects for low income families. Average income in first 5 years of life has greatest effect. Larger income effects in the sibling-difference model, but imprecise.
Taylor *et al.* (2004)	United States; NICHD Study of Early Child Care; 15-36 months	Permanent income (over 36 months) + income-to-needs ratio	Cognitive development, language performance and social behaviour	OLS + random-effects. Similar controls to Blau (1999) + maternal depression and home environment indicators	Significant and positive effects of income. Larger income effects for low income families and larger effects for permanent income. Smaller income effects accounting for endogeneity.
Dahl and Lochner (2005)	United States; NLSY; Children > 5 years	Annual disposable family income + permanent income (since birth)	Standardised math and reading test scores	OLS + fixed-effects + IV. Similar controls to Blau (1999).	Significant and positive effects of income assuming exogeneity, with larger effects for permanent income. Size and significance reduced substantially after accounting for endogeneity.
Mayer (1997)	United States; NLSY, PSID; Children and young adults	Annual income (logs, levels) + permanent income + sources of income (earnings, welfare, assets etc.)	Test scores, probability of teenage childbearing & dropping out of school, education and hourly wages	OLS Similar controls to Blau (1999), less marital status	Generally small, insignificant effects of income when accounting for unobserved heterogeneity. Larger effects for poorer families. Larger (but modest) effects of permanent income.
Korenman *et al.* (1995)	United States; NLSY; All ages	Current income-to-needs ratio[1] + 13 year average income-to-needs ratio	Socio-emotional development and behavioural problems.	OLS. Similar controls to Blau (1999) + index of home environment.	Significant and large effects of income with larger effects for permanent income and larger effects for low income families. Robust to all controls.
Shea (2000)	United States; PSID; Young adults	Log permanent income	Wages and years of schooling	OLS + 2SLS. Controls for age, gender, race, education, parents' education.	Significant and positive effects of income on children's human capital assuming exogeneity. IV (2SLS) estimates smaller than OLS estimates and insignificant.
Maurin (2002)	France; EPCV Survey; 1997; 11-12 years	Log of annual income	Held back in elementary school	OLS + 2SLS. Controls for gender, parents' education, household size and structure.	Significant and large negative effect of income on being held back in school 2SLS estimates three times larger than the corresponding OLS estimates

4.3 Household Income and Child Obesity

In comparison to the literature reviewed in Section 4.1 and 4.2, the literature examining the effect of income on child and adolescent obesity is far less developed (see Table 4.3).

Moreover, studies in this area have yielded conflicting results. The relationship between family income and child BMI/obesity has been found to be negative for US children (Anderson *et al.* 2007) and adolescents (Baum and Ruhm 2007); positive for US adolescent males (Monheit *et al.* 2007) and adolescent females (Gordon-Larsen *et al.* 2003); insignificant for US children and adolescents (Classen and Hoyakem 2005); and highly non-linear, with very high and very low incomes both associated with decreased probability of overweight (Hofferth and Curtin 2005).[51]

The two studies which looked at this relationship for Australia both found evidence of a negative effect of household income on child overweight (Wang *et al.* 2002a; Wake *et al.* 2007). However, these studies did not account for unobserved heterogeneity and other sources of endogeneity, and hence, the interpretation of the income effects, and their policy implications, are uncertain. To date, no published study has examined the household income-child obesity gradient in Australia *or* internationally using methods that adequately account for endogeneity.

[51] Hofferth and Curtin (2005) argue that this result provides support for the hypothesis that the quantity of food increases with income up to a certain point, after which it most likely leads to a higher quality rather than to a higher quantity.

Table 4.3. Income and SES-Effects on Child BMI, Overweight and Obesity

Study	Country; Data; Age of Children	Income (or SES) Measure	Obesity Measure[1]	Specification and controls	Findings
Anderson *et al.* (2007)	United States; NHANES; 2-11 years	Income-to-poverty line ratio, race, parents' education.	BMI [M]	OLS. Controls for child's age and gender.	Significant and negative relationship between income and BMI (economically small).
Baum and Ruhm (2007)	United States; NLSY; Over 14 years	Mothers' education + annual household income	BMI + Obesity (BMI $\geq$ 18 kg/m^2) [SR]	OLS and LPM. Controls for mothers' race, ethnicity, marital status, number of children, smoking, binge drinking and physical activity.	Significant and negative relationship between SES and BMI/obesity (economically small), which increases with age. Larger effects for females. Half the SES-obesity gradient remains after including all controls.
Classen and Hoyakem (2005)	United States; NLSY; 2 to 18 years	Dummies for income quartile	BMI + overweight (CDC) [M] or [PR]	Probit. Controls for mothers' BMI, education, labour supply and marital status, family size, health insurance, the youth's mental health, race, sex and health at birth.	Insignificant.
Monheit *et al.* (2007)	United States; See note 2; 12 to 19 years	Income-to-poverty line ratio	BMI + overweight (CDC) [PR]	OLS and Probit. Controls for race, mother's employment and weight, parents' education and smoking, family size.	Significant and positive relationship between poverty ratio and BMI/overweight for males. Insignificant for females.
Gordon-Larsen *et al.* (2003)	United States; See note 2; 12 to 20 years	Annual household income (and poverty dummies)	BMI + overweight (CDC) [SR]	OLS and Logistic regression. Controls for race, parental education, gender.	Significant and positive relationship between income and BMI/overweight for females. Insignificant for males.
Hofferth and Curtin (2005)	United States; PSID-CDS; 6 to 12 years	Income-to-poverty line ratio (and poverty dummies)	BMI + Overweight (CDC) [PR]	OLS and Logistic regression. Controls for gender, age, mothers' education.	Significant non-linear effects, with incomes 3 times below the poverty threshold associated with lower probability of obesity than children with incomes between the poverty line and 3 times the poverty line.
Booth *et al.* (2001)	Australia; See note 2; 2-17 years	SEIFA (IRSAD) + mothers' education	BMI + overweight & obese (IOTF) [M]	OLS. Controls for gender, age, ethnicity, location (urban/rural)	Insignificant.
Wang *et al.* (2002a)	Australia; NNS 1995; 7-15 years	Annual household income (indicators for income quintile)	BMI + overweight & obese (IOTF) [M]	Logistic regression. Controls for parental BMI, gender.	Significant and negative relationship between income and BMI/overweight/obesity for males. Insignificant for females.
Wake *et al.* (2007)	Australia; LSAC; 4-5 years	Mothers' education + equivalised income + indigenous + SEIFA	BMI + overweight & obese (IOTF) + [M]	Logistic regression. Controls for gender, age, number of siblings, birth-order, English-speaking ability, parents' BMI	Significant and negative relationship between income child obesity, controlling for other factors and dimensions of SES.

Note 1: 'CDC' denotes cut-offs based on gender and age-specific growth charts compiled by the Center for Disease Control and Prevention's National Center for Health Statistics; 'IOTF' denotes cut-offs based on age and gender specific thresholds developed by Cole *et al.* (2000) in accordance with the International Obesity Task Force (IOTF). [M], [PR] and [SR] denote measured, parent-reported and self-reported height and weight, respectively.
Note 2; Monheit *et al.* (2007) used data from the Medical Expenditure Panel Survey; Gordon-Larsen *et al.* (2003) use data from the National Longitudinal Study of Adolescent Health; Booth *et al.* (2001) used data from the National Nutrition Survey 1995, the NSW Schools Fitness and Physical Activity Survey 1997, and the Health of Young Victorians Study 1997.

The majority of Australian research has been directed at investigating the effects of SES (e.g. parents' education and occupation), rather than income, on child overweight/obesity. Consistent with the studies investigating the effects of household income, research in this area has generally uncovered a negative relationship between SES and child and adolescent obesity (O'Dea 2003; Salmon *et al.* 2005; O'Dea 2009; Booth *et al.* 2001).[52] However, as these studies typically report simple correlations between dimensions of SES and child/adolescent obesity and/or control for a (very) limited set of confounding factors, they do not establish whether low-SES causes child and youth obesity, or whether there is some other observed or unobserved factor(s) causing both low SES and obesity.

The most comprehensive study on the socioeconomic correlates of child obesity in Australia is Wake *et al.* (2007), which investigated the association between 12 indicators of SES and the likelihood of overweight in 4-5 year old Australian children.[53] This study made use of the rich set of demographic and household variables available in the first wave of the Longitudinal Survey of Australian Children (LSAC). Using a multivariable logistic regression model, the authors uncovered a substantial and significant socioeconomic gradient in weight outcomes. Furthermore, the authors find a significant, negative relationship between equivalised family income and child overweight even after controlling for other dimensions of SES and parents' BMI. Nevertheless, the authors did not account for endogeneity and hence the true causal effects are unclear.

---

[52] O'Dea (2003) found BMI to be significantly higher among low-SES children than middle- and high-SES children; Salmon *et al.* (2005) found higher increases in obesity rates among 9-13 year old children attending schools in low-SES areas compared with children attending schools in high-SES areas (a 60% vs. 52% increase) between 1985 and 2001; and O'Dea (2009) found that, between 2000 and 2006, significant increases in obesity occurred only among students from low SES schools ($p<0.05$), rather than middle ($p=0.32$) and high ($p=0.92$) SES schools.

[53] Wake *et al.* measure SES along several dimensions; the mother (lower maternal education), the family (lower income, less-skilled occupations, immigrant and indigenous status) and the community (SEIFA).

In summary, the effect of income on child overweight/obesity in Australia has received little attention. The larger international literature has generally uncovered modest effects of permanent income on child health and development, whilst the evidence is more mixed for child and youth obesity. Moreover, no study in Australia or internationally, has examined whether the observed relationship between income and overweight is causal. Given the theoretical reasons for the existence of an income-obesity gradient and the lack of prior research, this thesis contributes to the Australian and international literature by investigating the causal effect of household permanent income on youth obesity in Australia.

4.4 Maternal Employment: A Review of the Evidence

The relationship between the labour supply of mothers and the health of their children has been a subject of interest among empirical researchers for over three decades.[54] It is only recently, however, that economists have begun examining the link between maternal labour supply and child weight problems specifically, and whether this link is causal. This literature is summarised in Table 4.4.[55]

Research in this area has generally uncovered a positive and statistically significant relationship between maternal employment and obesity in children, despite differences in methodology, data, nationality and children’s ages. A positive effect was found by Classen and Hokayem (2005) using data on American children aged 2-18; Phipps *et al.* (2006) using data on 6 to 11-year-old Canadian children; Garcia *et al.* (2006) using data on Spanish children aged 2-15 years; and Takahashi *et al.*

[54] An early study by Edwards and Grossman (1978) found that participation by mothers in the labour market is detrimental to the health of their children, particularly via the effect on the child's nutritional status. The study by Blau *et al.* (1996) was the first to examine the relationship between maternal labour supply and child health accounting for the potential endogeneity of maternal labour supply. Using longitudinal data from the Philippines and fixed-effects, Blau found little evidence to suggest that maternal employment has a direct causal effect on child health.

[55] These studies have tended to focus on maternal, rather than parental employment given traditional gender roles assign meal preparation and child care responsibilities primarily to mothers. Previous literature has found no evidence of a relationship between paternal employment and childhood obesity (Phipps *et al.* 2006; Chia 2008).

(1999) using data on 3-year-old Japanese children. However, the labour supply effects estimated in these 'first-generation' studies are potentially biased, given that mothers working long hours may differ in unmeasured ways from those who do not, and that labour supply may be simultaneously determined with child obesity.

The seminal study of Anderson *et al.* (2003) was the first to contend with the issue of causality. The employment effects in this study were similar to the aforementioned studies and were robust to a range of econometric techniques controlling for observable and unobservable differences across individuals and families, such as instrumental variables and fixed-effects.[56] Chia (2008) also found no evidence of omitted variable bias using a sibling-difference specification and data on 0-11 year old Canadian children. Overall, however, evidence for the endogeneity of mothers' employment is mixed, with Ruhm (2004) finding some evidence of reverse causality bias[57] and recent studies for the US (Liu 2006) and Australia (Zhu 2007) finding that labour supply effects are biased downward in specifications assuming exogeneity.

[56] The authors show that these estimates are not economically large, when one considers that increases in maternal employment account for only around 6 percent of the overall increase in childhood obesity over the period 1975-1999.

[57] This is based on a statistically significant coefficient on a variable measuring mothers' employment in the year after the child's weight is observed (see Section 3.5.2).

Table 4.4. Maternal Employment Effects on Child BMI, Overweight and Obesity

Study	Country; Data; Age of Children	Maternal Employment Measure	Obesity measure [2]	Specification and controls	Findings
Anderson *et al.* (2003)	United States; NLSY; 3-11 years	Average weeks per year since birth *and* average hours per week (if working) since birth.	Overweight (CDC); [M] or [SR]	Probit + long-difference models + sibling-difference models + instrumental variables. Standard controls[1] + AFQT test score, child breastfed, mothers' weight status, permanent family income (since birth), % child's life mother married	A 10 hour increase in hours per week increases likelihood of child overweight by 2 to 4 percentage points. No effect of weeks worked per year. Robust to specification used and significant with exception of IV. No evidence of unobserved heterogeneity. Relationship confined to high income, white, educated families, with insignificant effects for children from less affluent families.[3]
Ruhm (2004)	United States; NLSY; 11 years	Average hours per week since birth (and squared) + hours worked per week in year *after* child assessment	Obese and "at risk of overweight" (CDC); [PR]	Probit + proxy variable approach + sibling fixed-effects + propensity score techniques. Standard controls[1] + mothers' AFQT score, child's health at birth and home environment, mothers' pre-pregnancy employment characteristics, home environment at age 14, and BMI, geographic indictors of crime, divorce and medical facilities.	A 20 hour increase in hours per week increases likelihood of child obesity by between 1.6 and 2.9 percentage points and 'at-risk of overweight' by between 2.3 and 4 percentage points. Robust to additional controls and proxies for unobservables. Relationship confined to 'advantaged' youths, with insignificant effects for 'disadvantaged' youths. Evidence of reverse-causality from child overweight to work hours. No evidence of non-linearity.
Zhu (2007)	Australia; LSAC (child cohort); 4 to 5 years	Average hours per week during previous month (includes 0 hours) *or* non-working part-time, full-time	Overweight (IOTF); [M]	2SLS linear probability model + multinomial treatment model with selection. Standard controls[1] + birth order, fathers' employment and mothers' nonwage income	2SLS model: 1 hour increase in hours per week increases likelihood of child overweight by 0.6 percentage points (6 times larger than OLS estimates). Selection model: Full-time employment increase probability of child overweight by 19 percentage points (relative to non-working), with no significant effect of part-time employment (relative to non-working).
Liu (2006)	United States; NLSY; 3 to 11 years	Full-time employment dummy	BMI + overweight (CDC); [PR]	Control function estimator + propensity score method + endogenous switching regression model Standard controls[1] + annual family disposable income	Average treatment effect of full-time employment on child's BMI is 2.214kg/m^2; or increase in likelihood of overweight of 14.2% (relative to a mother not working full-time). Exogeneity of mothers' full-time employment rejected using specification tests.
Chia (2008)	Canada; National Longitudinal Survey of Children and Youth; 0 to 11 years	Average hours per week over sample + lagged work hours (after child birth & when child started school)	BMI + overweight & obese (IOTF); [PR]	OLS + linear probability model + sibling fixed-effects. Standard controls[1] + log of permanent family income, mothers' marital status and health and child's participation in sports.	A 10 hour increase in hours per week increases likelihood of child overweight by 2.5 percentage points, with *larger* effects after accounting for unobserved heterogeneity. Larger effects for work hours in the early stages of child's life and for children with low family income.

Continued over page

Maternal Employment Effects on Child BMI, Overweight and Obesity (Continued)

Study	Country; Data; Age of Children	Maternal Employment Measure	Obesity measure [2]	Specification and controls	Findings
Scholder (2008)	England; National Child Development Study (Panel); 16 years	Full-time employment at pre-school age, age 7 and age 11	BMI + overweight (IOTF); [M]	Probit + random-effects + fixed-effects. Standard controls[1] + fathers' employment, income, child breastfed, mother smoked during pregnancy and fathers' socio-economic class at child's birth	Full-time employment at age 7 increases probability of overweight at age 16 by 5.5 percentage points, with no effect at earlier/later stages.
Classen and Hokayem (2005)	United States; NLSY; 2 to 18 years	Full-time, part-time and non-working	BMI + overweight (CDC); [M] or [PR]	Probit. Standard controls[1] + mothers' weight status and marital status, health insurance, child's mental health	Significant, positive effects of maternal employment on likelihood of child overweight (smaller estimates than Anderson *et al.* 2003)
Phipps *et al.* (2006)	Canada; National Longitudinal Survey of Children and Youth; 6 to 11 years	Hours per week (current) + 6 year average + combined maternal and paternal hours [11]	Overweight (CDC & IOTF); [PR]	Logistic regression. Standard controls[1] + current and permanent gross family equivalised income[2], fathers' schooling, child's physical activity	Significant, positive effect of *average* maternal work hours on likelihood of child overweight, with smaller effects for concurrent employment and no effect for fathers' employment.
Garcia *et al.* (2006)	Spain; National Health Survey 2003; 2 to 15 years	Full-time, part-time and non-working	BMI + overweight & obesity.[3] [SR]	OLS + Probit. Standard controls[1] + mothers' BMI.	Significant, positive effect of maternal employment on likelihood of child overweight.
Takahashi *et al.* (1999)	Japan; Toyama Study; 3 years	Full-time, part-time, non-working	Overweight ($BMI \geq 18$ kg/m^2); [M]	Logistic regression. Control for mothers' and fathers' weight status, birth weight, physical activity and nutrition,	Significant, positive effect of maternal employment on likelihood of child overweight.

Note 1: Standard controls include: child's age, child's gender, child's race, child's birth weight, mothers' education, mothers' age, location, family size and composition.
Note 2: 'CDC' denotes cut-offs based on gender and age-specific growth charts compiled by the Center for Disease Control and Prevention's National Center for Health Statistics; 'IOTF' denotes cut-offs based on age and gender specific thresholds developed by Cole *et al.* (2000) in accordance with the International Obesity Task Force (IOTF). [M], [PR] and [SR] denote measured, parent-reported and self-reported height and weight, respectively.
Note 3: Cut-offs based on 85^{th} and 97^{th} percentiles of the BMI distribution for each age and sex in Spain.

One important finding that has emerged from the literature is that the observed relationship is driven by children from high income, white and highly education families, despite the fact that children from such households are least likely to have weight problems (Anderson *et al.* 2003; Ruhm 2004).[58] These studies (for the US) have found that large amounts of work by mothers have no effect on the weight status of 'disadvantaged' youths, and that even very limited maternal employment has large adverse effects on 'advantaged' youths. The evidence for Canada points to the opposite effect, with Chia (2008) finding that the impact of maternal employment on obesity works mostly through those in the *lower* income quartiles. No study has examined this phenomenon for Australia.

The timing of mothers' employment is also important. Using separate dummy variables for employment at different ages, Chia (2008) find that labour supply in the earliest years of the child's life has the largest adverse effects on (future) obesity. Scholder (2008) examined a similar hypothesis using a similar methodology and longitudinal data on British children born in 1958, finding that children with a full-time employed mother during mid-childhood (aged 7-years) were 5.5 percentage points more likely to be overweight at age 16, with no significant effects of full-time employment at earlier/later ages.[59]

4.4.1 Studies for Australia

Despite the relevance of this issue to Australia, only one study has examined the effect of mothers' employment on the weight outcomes of Australian children (Zhu 2007). Using data on 4-5 year old children from the first wave of the Longitudinal Survey of Australian Children (LSAC), Zhu found that a 1 hour increase in maternal employment increases the likelihood of child being overweight by 0.6 percentage points. Interestingly, she also found evidence of a

---

[58] Anderson *et al.* (2003, p.492) and Ruhm (2004) estimate their models by income quartile, mothers' education and race/ethnicity, and a multivariate index of SES (Ruhm 2004).

[59] Scholder (2008, p.903) explains this result by noting that food preferences and habit formations in children develop around mid-childhood. However, the data for this study are dated, with children born in 1958 not part of the current obesity epidemic and working mothers in the 1960s and 1970s likely to be different from working mothers today.

non-linear relationship; with full-time employment increasing the probability of overweight by 19 percentage points (relative to not working), and part-time maternal employment having no significant effect (relative to not working). Notwithstanding, the instruments used in this analysis were questionable (conceptually and using statistical tests) and sensitivity tests on the results were not reported.

In summary, the international literature on the relationship between maternal employment and child obesity has uncovered consistent evidence of a positive (adverse) effect of employment, which is either unchanged, or increased, after accounting for endogeneity. This literature has also found different effects for children of different SES, and that maternal employment has larger effects at some ages (early and mid-childhood) than others. Furthermore, no study, in Australia or elsewhere, has analysed the effects of employment on *older* children – aged more than 12-years.[60] For this reason, and since the timing of mothers' employment appears to matter (Scholder 2008; Chia 2008), this thesis fills an important gap in the literature.

[60] The exceptions are Classen and Hokayem (2005) for US children aged 2-18 years and Garcia *et al.* (2006) for Spanish children aged 2-15 years. However, the effect sizes in these studies will correspond to an average of the effects on younger children and the effects for older children. Scholder (2008) examines weight outcomes among 16-year old English children, but using a lagged measure of maternal employment (prior to age 12).

5 Data, Sample and Variables

5.1 The Data

There are two nationally representative, longitudinal data sets for Australia that currently collect data on Australian children. These are the Household, Income and Labour Dynamics in Australia Survey (HILDA) and the Longitudinal Survey of Australian Children (LSAC). This thesis uses the former, primarily because HILDA has more years of data available, which allows for a more reliable (less noisy) calculation of permanent income.[61] Further, no published study has used HILDA to investigate the effect of household income and/or maternal employment on child or youth health outcomes in Australia.[62]

The HILDA Survey began in 2001 with a large national probability sample of Australian, privately occupied dwellings, covering 7,682 households and 19,914 individuals of various ages (Wilkins *et al.* 2009, p.v; Headey and Warren 2008, p.vi).[63] Information has been collected annually on members of the households that participated in Wave 1, on any children subsequently born to or adopted by them, and on persons who later joined one of the original households and had a child with one of the original sample members or their direct descendants. From Wave 2 onwards, information has also been collected on persons who joined one of the original households, but only for as long as they remained in the household (Watson 2009, p.2).

---

[61] HILDA is also a dedicated income survey that contains a measure of disposable household income which includes welfare benefits. The LSAC data set, on the other hand, does not provide a measure of disposable income (Department of Families, Community Services and Indigenous Affairs 2009). Further, LSAC uses a series of "abbreviated questions" to collect information on gross family income, which are likely to lead to an "underestimation of income" since respondents may forget to report income from minor income sources, such as family payments (Bradbury 2007, p.9).
[62] As reported in Chapter 4, the LSAC data has been used to explore the relationship between family income and parent-reported child health (Khanam *et al.* 2009) and the relationship between maternal employment and child overweight (Zhu 2007).
[63] Excluded from the scope of the HILDA survey, are the homeless and people living in most institutions (children living in boarding schools are included). People who move into institutions in subsequent years remain in the sample.

5.2 The Sample

In HILDA, information on height and weight was collected for the first time in Wave 6, with the same questions included again in Wave 7. This information is self-reported in the self-completion questionnaire (SCQ), which is completed by persons aged 15 years and older living within a sampled household. Accordingly, a BMI measure cannot be constructed for persons younger than 15. With this in mind, the main sample used in this analysis consists of children between the ages of 15 and 19 years at Wave 7.[64] Persons aged 20 years and older are necessarily excluded from the sample as there are insufficient years of data during adolescence available to construct valid measures of the key explanatory variables: permanent income and maternal employment.

For the purposes of this thesis, a child's 'adolescence' is defined as the four year period spanning the ages of 12 to 15 years.[65] The exception is 19 year olds, for whom adolescence is defined as the 3 year period between the ages 13 and 15, as necessitated by (limitations on) the data.[66] The key variables used in this analysis, namely permanent income and maternal employment, are measured exclusively during the years in which the child is defined as an adolescent. This is motivated by the fact that persons aged 15 and under are less likely to engage in income-generating activities and are more likely to be full-time dependent students, and hence, the potential for reverse causation running from the child obesity to household income is largely mitigated.[67] Moreover, this thesis seeks to

[64] The sample also includes a number of 15 to 18 year olds measured at Wave 6, for whom a BMI observation was missing in Wave 7 but available in Wave 6. This is explained below.

[65] For instance, Rodgers and Rodgers (2006, p.2) assert that "four years of data is the bare minimum required to measure permanent income". The choice of an income period longer than 4 years would have reduced the sample size, since data at earlier stages of adolescence are unavailable for the older members of the sample. Similarly, the sample would be reduced using a four-year data period at an earlier stage of the child's life, such as between the ages of 11 and 14 years.

[66] The decision to include 19-year olds in the sample was based on sample size considerations. The potential issue in retaining these observations is that income and maternal employment cannot be observed for these persons at age 12, and hence permanent income and measures of maternal employment are not constructed using an equivalent period of data as the 15, 16, 17 and 18-year olds in the sample. Thus, for the purpose of sensitivity analysis, a smaller sample that excludes 19 year-olds is considered in Section 8.1.4 and 8.2.5.

[67] As discussed in Chapter 3, reverse causality can still occur through parental employment

explicitly examine whether income and maternal employment in adolescence affects overweight.

The main sample of children is taken from the seven-year balanced panel of enumerated individuals who were aged 15- to 19-years in Wave 7. This ensured that annual income data are available for all sample members over the three or four year period comprising their adolescence. It could be argued that an unbalanced panel would be more appropriate, given that there are several 'redundant' years in which the youth is not classified as an 'adolescent' and not supplying information on height and weight. The issue with an unbalanced panel is that there is no appropriate way to weight the data. The cross-sectional weights are not appropriate, given that persons not present in all necessary waves would have to be excluded from the cross-section sample, nor are the longitudinal weights, because they require enumerated persons be present in *all* years between Wave 1 and Wave 7. Drawing the sample from a balanced panel circumvents this problem, since enumerated person longitudinal weights may be validly applied.[68]

Initially, this resulted in a sample of 1,078 children, consisting of 210 15-year olds, 261 16-year olds, 192 17-year olds, 225 18 year-olds and 190 19 year-olds. However, the sample was reduced following several sample restrictions, which are summarised in Table 5.1.

Table 5.1. Sample Selection

Remaining Observations	Observations Excluded	Total Youths
Youths	-	1,078
Mother living in household	18	1,060
Dependent child/student	55	1,005
Strictly +ve household income	10	995
Non-missing employment data	6	989
Non-missing BMI data	110	879
Not Pregnant	5	874
Final Sample Size		**874**

[68] Preliminary analysis revealed that 16 additional observations would be added by using an unbalanced panel. The cost of the inability to weight was deemed greater than the benefits of additional sample size. A sensitivity analysis using a sample taken from an unbalanced panel (with unweighted estimates) is conducted in Sections 8.1.4 and 8.2.5.

First, any child for whom a 'mother' was not present at any time over adolescence was excluded. For the purposes of this thesis, the child's 'mother' is defined as an own (by birth or adoption), step or foster mother who is co-resident in household (MIAESR 2009a, p.D4).

Next, 55 observations were excluded pertaining to children classified as anything other than a 'dependent child' or 'dependent student' during adolescence.[69] [70] This corresponds to the ABS (Cat. No. 6523.0, 2009a, p.76) definition of a 'dependent child' and ensures that household income measured when the child was 15 years old is strictly exogenous with respect to the child's overweight status.[71]

Following the approach of Headey and Warren (2008, p.52), a further 10 observations were excluded relating to those persons having zero or negative household income in at least one year during their adolescence, on the assumption that their income data are unreliable. In addition, six observations were dropped pertaining to those children missing employment data on the mother in every year during adolescence.

The sample was further reduced due to missing data on height and weight. In Wave 7, 223 youths had incomplete information on height and/or weight, representing approximately 22.5 per cent of the remaining sample. However, of those 223 persons, 113 had complete height and weight information in Wave 6, which was used in place of the Wave 7 missing BMI observation. These 113 observations are herein referred to as the 'supplemented observations'. These observations included only those persons aged 15-, 16-, 17- or 18-years in Wave 6, given that the 19-year olds in Wave 6 had only two years of data available during

[69] These included 46 non-dependent children, 4 lone persons, 2 unrelated to all household members, 1 other family member, 1 lone parent with a child less than 15 years of age and 1 couple family without children.

[70] Dependent children are aged 15 years and under, whilst dependent students are defined as persons aged 15-24, studying full-time, not working full-time and living in a household with their parent but without a partner or child of their own (MIAESR 2009c, p.R221).

[71] If non-dependent children and persons starting a new household are retained in the sample, household income may be endogenous, since individuals are likely to be contributing to household income in some way, which in turn may be influenced by their health.

adolescence. Following this strategy, the proportion of respondents with missing BMI information was reduced from 22.5 per cent to around 11.1 per cent of the remaining sample.[72] Note, however, that no more than one observation on BMI was used for each child and therefore this analysis is inherently cross sectional in nature.

Finally, since BMI is not a valid measure of excess body weight for pregnant women (WHO 2000, p.8), five females who reported to have become pregnant in the 12 months prior to the Wave 7 interview date (Wave 6 interview date for 'supplemented observations') were excluded from the sample.[73]

The final sample consisted of 874 person observations, made up of 183 15-year olds, 203 16-year olds, 175 17-year olds, 190 18-year olds and 123 19-year olds. This corresponds to 81.1 per cent of the initial sample. In 856 cases the mother was the child's own mother, representing 98 percent of the full sample; a step mother in 12 cases; and foster mother in 2 cases. The remaining 4 children had more than one mother during their adolescent years.[74] In 751 cases a father (own, step or foster) was present in the household some time during adolescence, which was an own father in 656 cases, a step father in 91 cases, a foster father in one case and both an own father *and* a step father in 3 cases.

[72] This corresponds closely with the percentage of persons for whom the BMI score is not available in the 2004-05 NHS; 12.1% for males aged 15-24 and 14.4% for females aged 15-24 (ABS Cat. No. 4363.0, 2006, p.88). The reasons for missing BMI scores were non-completion of the self-completion questionnaire (73 obs.), failing to disclose height and/or weight (34 obs.) and invalid or implausible responses (3 obs.).

[73] Three females pregnant in the year preceding the Wave 7 interview were *retained* in the sample as they were able to be observed prior to becoming pregnant (Wave 6). These females were added to the 'supplemented sample', which thus contained a total of 116 observations.

[74] In these four cases, the children spent part of their adolescent years with a biological (own) mother present in the household and part of their adolescent years with a step mother. In these cases, all predetermined maternal variables, including the mother's country of birth, education and age are measured for the biological (own) mother, whereas time-varying variables are measured for whichever mother is present in the relevant year. The same convention is used in cases where the child had both an own and step father during their adolescence.

5.3 Weights

The HILDA survey is designed to be representative of the Australian population. An important issue with HILDA and other longitudinal surveys is that over time the demographics of the general population change, and attrition occurs (Wilkins *et al.* 2009, p.v). Such factors may render the sample non-representative of the general population, which is problematic if the objective is to make inferences about the population using samples drawn from the data. Fortunately, this problem can be addressed, at least partly, by using the weights provided with the data set. That is, given the appropriate weights are applied, the HILDA sample should constitute a representative sample of all Australian (residing in non-remote areas), over time and/or at a particular point in time.

For the purposes of this thesis, the longitudinal enumerated person weights are appropriate for most purposes. These weights adjust for the non-response and attrition experienced at both the household and person level between Wave 1 and Wave 7, as well as for unequal probabilities of selection into the HILDA survey sample (Wooden and Watson 2007, p.217). Applying these weights ensures that the children in the seven-year balanced panel remain representative of all Australian children (of similar age) who were living in private dwellings in non-remote areas from 2001 to 2007. The (longitudinal) weights are appropriate given that the sample used in this analysis is drawn from a *balanced* panel of children, rather than a series of cross-sections.

5.4 Variables

5.4.1 The Dependent Variable

The dependent variable in this analysis is specified as a continuous variable for BMI and as binary variables for BMI category, with the latter defined according to the age- and gender- specific cut-offs developed in Cole *et al.* (2000) and presented in Chapter 3, Table 3.1. To determine the sensitivity of the results to the BMI category used, two binary outcomes are considered: the first equal to one if the youth is overweight *or* obese, and zero otherwise, and the second equal to one if the youth is obese, and zero otherwise.[75] The sample distribution of BMI is presented in Figure 5.1. The sample distribution of BMI category is presented in Table 5.2.

Figure 5.1. Body Mass Index

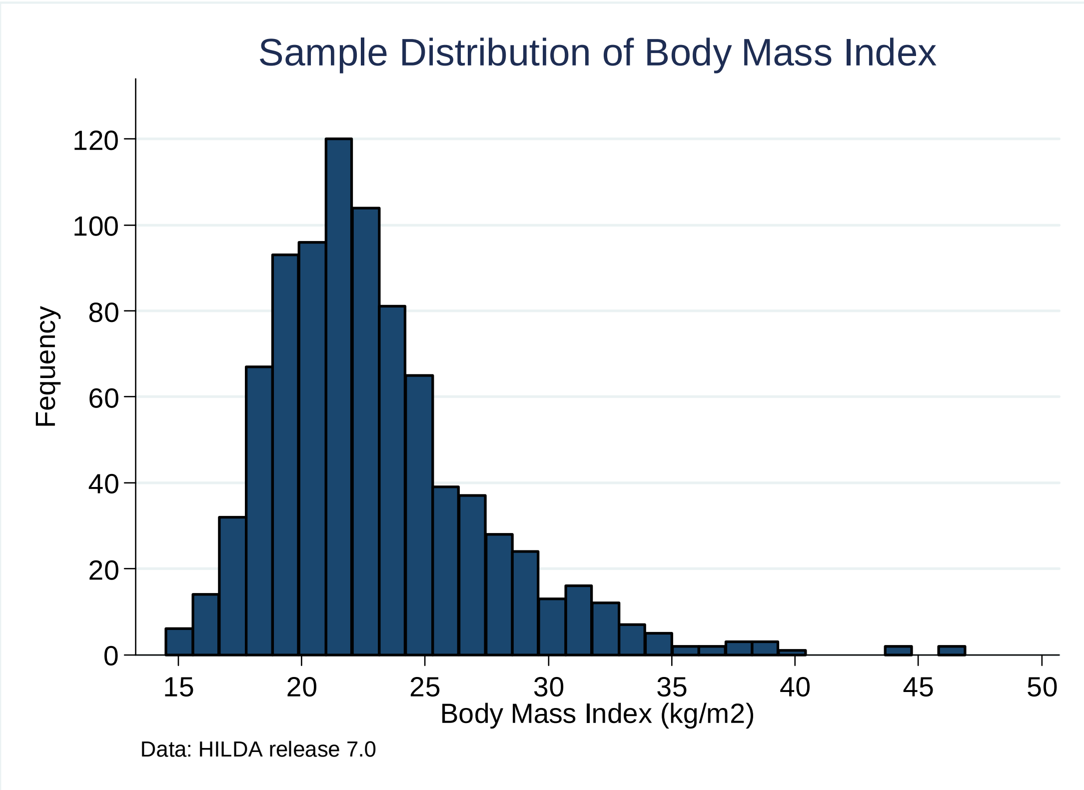


Table 5.2. Sample Distribution of BMI Category

	Number	Proportion
Normal weight	634	72.5%
Overweight or obese	240	27.5%
Obese	69	7.9%
Total	874	

There are a reasonable number of youths in the 'overweight or obese' and 'obese' categories.[76] A comparison of the prevalence rates with external aggregates is given in Section 6.1.

[75] This distinction recognises that obesity is worse than overweight in a medical sense, whereas the number of obese youths in the sample may be too few to obtain precise estimates.

[76] Two males in the sample lie outside the extreme BMI values reported by the ABS in their comparison of the self-reported and measured data collected as part of the 1995 NHS and NNS (=45.9kg/m^2), though only marginally (Wooden *et al.* 2008, p.6).

5.4.2 Permanent Income

The permanent income measure used in this thesis is constructed using annual data on the youth's real, household, equivalised disposable income during adolescence. Household financial year disposable income is the combined income of all household members after receipt of public transfers and deduction of taxes (Wilkins *et al.* 2009, p.25).[77]

Since previous work for Australia has shown that imputed rent on owner-occupied housing "contributes significantly to the average level of household income" (Saunders and Siminski 2004, p.364), and that home ownership varies across demographic groups (ABS 2008a, p.322),[78] household imputed rental value is added to disposable income. Household imputed rent is calculated as 4 per cent of the difference between the imputed house value and the remaining mortgage principal (Lillard *et al.* 2009, pp.2-5).[79] This assumes that the non-cash benefits received from investments in owner-occupied housing (i.e. shelter) are conceptually similar to cash payments (Rodgers and Rodgers 2009, p.6). Including imputed rent as a component of household income has not been done in the empirical literature on the income-child health relationship; hence the measure of income used in this thesis represents one of the most complete measures of household income used in the literature to date.[80]

---

[77] Household *gross* income consists of wages and salaries, business income, investment income, private pensions, private transfers, Australian government pensions and benefits, family tax benefits and maternity allowances (Watson 2009, p.51). Windfall income is excluded from gross income (to obtain a measure of 'regular' income) as are transfers in kind, including the Child Care benefit (Watson 2009, p.54).

[78] In particular, home ownership rates increase with the age of the reference person in the household and are particularly concentrated among couple-households (ABS, 2008a, p.322).

[79] The algorithm used to calculate household imputed rental values is taken from Lillard *et al.* (2009, pp.2-5). This algorithm is used to impute household rental values in the HILDA component of the Cross National Equivalent File (CNEF). Four per cent is assumed to reflect the opportunity cost of the funds invested in the property. Imputed rental values for the 38 youths living in public housing and the 10 youths living in rent-free housing was not imputed, since this imputation required access to an external program and data from the Census of Population and Housing.

[80] Section 8.1.4 and 8.2.5 examine whether the results are sensitive to the exclusion of imputed rental values from permanent RHED income.

Disposable income (including imputed rental value) is adjusted to take account of differences in household size and composition using the 'modified OECD' equivalence scale ABS (Cat. No. 6523.0, 2009a, p.57).[81] This assumes that resources are pooled and shared equally among household members (Wilkins *et al.* 2009, p.26) and that household members are able to improve their standard of living by taking advantage of economies of scale in consumption and household production, such that the costs of living increase less than proportionately with household size. Finally, annual disposable income *per adult equivalent* is converted into real terms using the Consumer Price Index, so that income in all waves is expressed in 2006-07 dollars (ABS Cat. No. 6401.0, 2009b).[82] For brevity, the youth's real, household, equivalised disposable income (including imputed rent on owner-occupied housing) is hereafter referred to as his or her "annual RHED income".

Following Rodgers and Rodgers (1993, p.31), this thesis defines *permanent income* as the maximum sustainable annual consumption level that the agent could achieve with his or her actual annual RHED income stream over adolescence, by saving and borrowing at prevailing market interest rates. This methodology represents an extension of the 'smoothed' or 'average' income measures used in the literature (Taylor *et al.* 2004; Phipps *et al.* 2006; Case *et al.* 2002; Blau 1999; Mayer 1997) in that it assumes that intra-year income transfers are performed at a cost equal to the market rate of interest; rather than at a zero cost. Since permanent income is based on individuals' annual RHED income during *adolescence*, this variable is hereafter referred to as "permanent RHED income".

[81] The modified OECD scale assigns a weight of one to the first adult in the household, a weight of 0.5 to each extra adult and a weight of 0.3 for every child under the age of 15 years. Hence, a household comprising two adults and two children less than 15 years old is said to contain 2.1 adult equivalents (= 1 + 0.5 + 0.6). This implies that this household would require an income 2.1 times that of a lone person household in order to achieve the same standard of living (ABS 2009a, p.57). Household *equivalised* income is calculated by dividing total household income by the number of *adult equivalents*. The equivalised income of the household is assigned to all persons living in that household.

[82] The consumer price index used is: CPI, All Groups, Weighted Average of the Eight Capital Cities. It is calculated as a weighted average of the quarterly CPI over the relevant financial year (corresponding to each Wave), with weights equal to the number of days in the four quarters.

Permanent RHED income is calculated using the numerical algorithm described in Rodgers and Rodgers (1993, p.37), which is reproduced in Appendix A. In accordance with the permanent income hypothesis (Friedman 1957), this algorithm assumes that agents smooth their consumption intertemporally by saving and borrowing, so as to consume the same income in every period within the planning horizon.[83] [84]

In these calculations, the interest rates on savings in each year is set equal to the average of the RBA indicator interest rates on cash management accounts on balances totalling $10,000 and $50,000 and the interest rate on term deposits of terms 6 and 12 months. The interest rate on borrowings in each year is also calculated from RBA (2009a, Table F05) data, equalling the average of the indicator lending rates for fixed and variable rates on unsecured loans in addition to the standard credit card interest rate. The annual rates on savings and borrowing are presented in Appendix A, and they range between 3.02 and 4.67 per cent for the savings rate and between 13.52 and 15.07 per cent for the borrowing rate.

Figure 5.2 presents the distribution of permanent RHED income in the sample. Most children had a permanent RHED income of less than $50,000, with mean and median permanent RHED income equal to $36,207 and $33,821, respectively (unweighted). However, for four children, permanent RHED income was in excess of $130,000, which is over $20,000 more than the child with the fifth largest permanent RHED income in the sample and more than five standard deviations from the sample mean. Including these children in the sample could potentially bias the regression estimates and hence, the main estimates exclude these outlying

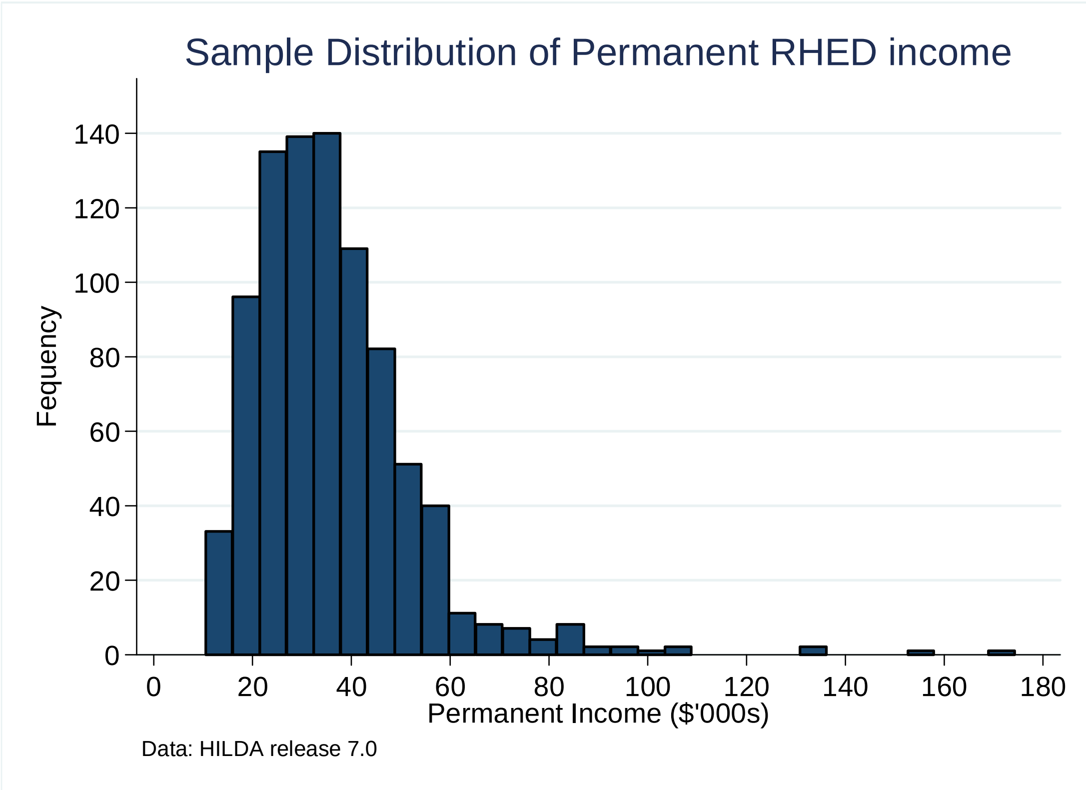
---

[83] This algorithm also assumes that each agent begins and ends the income period with wealth unchanged.

[84] Data on annual RHED income must be available in *all* years during adolescence in order to calculate permanent income, which is necessarily the case for all members of the sample, given that the sample is drawn from a balanced panel of individuals, and noting that missing and incomplete income data are imputed in HILDA (Wooden and Watson 2007, p.218).

observations. As a sensitivity analysis, the model is re-estimated including these observations in Section 8.1.4 and 8.2.5.

Figure 5.2. Permanent RHED Income



Given the sharply (right) skewed distribution of permanent RHED income (skewness coefficient = 2.311), the natural log of permanent RHED income is used (Khanam *et al.* 2009, p.5). This also accounts for the expected non-linear relationship between permanent income and the BMI of youths.

5.4.3 Maternal Employment

Mothers' employment is measured by the number of years in which the mother is observed to be working full-time, part-time, or not working at all, during adolescence. This is represented by two separate count variables with the number of years the mother was not working as the omitted category. Hence, the effects on BMI/obesity of working full-time and part-time are estimated *relative* to the effect of not working (in a given year). Two separate variables are also used to indicate the number of years the mother is non-responding (i.e. no person questionnaire) or living in a different household, respectively[85], and a dummy equalling one if the youth was observed in only three waves is used to ensure that the sum of these (five) variables was equal to four for all sample members.

In each year during adolescence, the mother's level of employment (part-time, full-time or otherwise) is observed only once (at the interview date). Thus, these variables will be a noisy measure of employment to the extent that the mother's interview-date employment level is an imperfect measure of her employment level over the preceding year. This is particularly true for mothers who work at a high intensity, but intermittently, over the course of an average year, and thus, under plausible assumptions, the coefficient estimates will be biased toward zero (Wooldridge 2002, p.75).[86] HILDA does not collect sufficiently detailed employment information for frequent intervals over the year (i.e. during every week) to construct a less erroneous measure of employment. Notably, the 'calendar' data does not distinguish part-time from full-time employment – a distinction shown to be important in prior work for Australia (Zhu 2007). Hence, information from the 'calendar' was not used in this analysis and magnitude of the bias stemming from "errors-in-variables" is explored using instrumental variables.

Matched data on the father is also used to construct four separate variables for paternal employment, namely the number of years the father is employed part-time; employed full-time; not present in the household; and non-responding. The omitted category is the number of years the father is non-working, with the 19-year old dummy defined previously ensuring that these (five) variables sum to four.[87]

---

[85] These are treated separately given that non-responding mothers still contribute positive time investments to the child.

[86] The assumption is that the errors of measurement are *classical* – uncorrelated with the 'true' proportion of the preceding year the mother spent in that employment category and uncorrelated with the disturbance in the primary relationship. The errors of measurement could induce a bias if mothers working in certain industries and occupations (such as school teaching) are more or less likely to be employed from September to November, which is the period most interviews are conducted (Watson 2009, p.94).

[87] The ability to include separate controls for the employment status of the father is a significant advantage to using HILDA. As Ruhm (2004, p.10) asserts, the fact that paternal employment has

An alternative specification of maternal employment, as a continuous measure of the average hours worked per week (during adolescence), was also considered.[88] The advantage of this specification is that it takes account of hours worked within the categories of part-time and full-time employment. However, a measure of average hours (across the four observations) was not used as it cannot distinguish a mother who works at a very high intensity for only part of adolescence, from a mother who works consistently but at an intermediate intensity. For example, a mother working 40 hours a week in one year during adolescence (and not working at all in the other three years) would be treated identically to a mother working 10 hours a week in all four years.

5.4.4 Basic "Background" Variables

The first set of control variables introduced into regression models in this thesis are the "core" demographic variables. These predetermined variables are unlikely to be chosen jointly with, or determined by, income and maternal employment.[89]

First, all estimated models include controls for age and gender (=1 if female), on the grounds that boys are expected to be heavier than girls (ABS Cat. No. 4364.0, 2009c, p.8) and that BMI is anticipated to be increasing in age (Cole *et al.* 2000, p.1).[90] Further, age may not be orthogonal to permanent RHED income. That is, since real incomes have consistently increased over the seven year-panel, higher

---

been ignored in "most prior research" is a "significant limitation dictated by severe constraints on the data available for fathers [for the US]. The omission of paternal employment (and other paternal factors) will lead to omitted variables bias, if, for instance, the father works fewer hours per week to compensate for an increase in the intensity of maternal employment.

[88] This specification of maternal employment has been used in several studies (Anderson *et al.* 2003; Ruhm 2004; Chia 2008)

[89] As argued by Blau (1999, p.262) and Mayer (1997, pp.63-66), specifications which are useful from a policy perspective do not hold fixed any variables that will actually change in response to changes in income and employment, or variables that are jointly chosen with income and employment.

[90] A lower correlation is expected between the age of the youth and the probability that he or she is overweight or obese given that the cut-offs formulated by Cole *et al.* (2000) account for age-specific differences in body size and composition. There still may be a correlation, however, given that the cut-off points were developed using a different reference population to that of Australian youths.

permanent RHED incomes will be calculated for younger members of the sample (on average), noting that adolescence constitutes a more recent four-year period of data for these individuals.[91]

The models also include a dummy variable equalling one if the child was Aboriginal and/or Torres Strait Islander, given that indigenous youths are generally heavier than non-indigenous youths (ABS Cat. No. 4719.0, 2008b, p.11; Wilkins *et al.* 2009, p.186), and that indigenous families may have fewer employment opportunities and experience more hardship than non-indigenous families.[92] Also, to account for the youth's ethnicity, the models include five binary indicator variables for the mother's country of birth, with Australian-born mothers representing the base category.[93]

5.4.5 Parents' Education

Parents' education is commonly seen as a productivity shifter in the child health production model (see Section 2.3.2), with higher education levels augmenting the efficiency of health production. Parents' education is also a correlate of permanent income and employment and so its effects should be held constant. Yet, since education is potentially jointly chosen with income, controls for education are omitted from the initial specification(s). In this analysis, mothers' and fathers' level of education are represented by three separate dummy variables, indicating whether the respective parent's highest education level is (i) a University degree (ii) a trade Certificate or Diploma (iii) Year 12 or (iv) Year 11 or less (the omitted category). These variables are constructed using the most recent data available during adolescence, or using the child's own reports of their parent's highest

[91] Median permanent RHED income ($2006-07; weighted) equals $36,147 for 15-year olds, $35,901 for 16-year olds, $33,986 for 17-year olds, $32,854 for 18-year olds and $33,432 for 19-year olds.

[92] There are 33 indigenous children in the sample.

[93] The country groupings include Australia (688 obs.); the Oceania region (excluding Australia) (22 obs.); the UK, Ireland, North and West Europe (53 obs.); Southern and Eastern Europe (24 obs.); Asia (50 obs.); and all 'other' countries (37 obs.). Country of birth information is reported by the mother (or the child if mother's data is missing). The 4-digit country codes are used to assign country groupings using the classifications provided in the Standard Australian Classification of Countries (ABS Cat. No. 1269.0, 1998).

education level in cases where either parent had not reported his or her education. Information on fathers' highest education was missing for 72 youths in the sample, and hence a dummy denoting the presence of missing data is used.[94]

5.4.6 Neighbourhood Characteristics

Evidence for Australia suggests that the concentration of fast-food outlets is greatest among lower SES neighbourhoods and that these areas have, on average, fewer kilometres of cycling and walking paths than middle- and high-SES areas (Kavanagh *et al.* 2007, pp.29, 50). To the extent that neighbourhood factors have effects on youth BMI/obesity independent of permanent RHED income (and employment), regression analysis should control for its effect. Conversely, neighbourhood SES may also mediate the effects of income on obesity, given that higher incomes allow a family to 'purchase' a higher SES area. For this reason, neighbourhood factors are not controlled initially, given that the "full" effect of income (including that due to changes in neighbourhood) would not be measured.

However, in some models, neighbourhood factors are included using the ABS Socio-Economic Index for Areas (SEIFA) 2001, Index of Relative Socio-Economic Advantage and Disadvantage (IRSAD) (ABS Cat. No. 2039.0, 2003).[95] The SEIFA-IRSAD provides a summary of the socioeconomic advantage and disadvantage of people living within a geographic area, *relative* to other areas in Australia. It incorporates information on the education, occupation, unemployment and incomes of persons in the area, along with other measures of economic and social resources, such as the proportion of occupied dwellings with

[94] 219 mothers and 199 fathers has a University qualification; 242 mothers and 378 fathers had a trade Certificate or Diploma; 127 mothers and 59 fathers highest education was Year 12; and 286 mothers and 166 fathers had completed less than Year 11.

[95] The ABS currently produce four neighbourhood indices; socio-economic disadvantage; socio-economic advantage and disadvantage; neighbourhood economic resources; and neighbourhood education and occupation (ABS 2003, p.2). The Index of Relative Socio-Economic Advantage and Disadvantage (IRSAD) was deemed most appropriate, since it has the broadest coverage of all available indices. This index is constructed by the ABS using data from the 2001 Census of Population and Housing.

internet connections and the incidence of single-parent households (ABS Cat. No. 2039.0, 2003, p.22).

In HILDA, the IRSAD variable classifies households into deciles, with the lowest and highest deciles representing the least and most disadvantaged neighbourhoods, respectively. These deciles are recoded into three categories (represented by three dummy variables) to indicate whether the child lived in a ‘disadvantaged’, ‘middle’ or ‘advantaged’ geographic area. The highest and lowest three deciles of the IRSAD are used to define an ‘advantaged’ and ‘disadvantaged’ suburb, respectively, whilst the remaining four deciles define a ‘middle’ ranked neighbourhood (the omitted category). Each youth is assigned to their modal neighbourhood category for the adolescent period.[96]

5.4.7 Mothers’ Age and Marital Status

The estimated models also control for the mother’s age (and its square) and marital status, measured by the number of years the mother is legally married during adolescence.[97] However, following Blau (1999, p.265), these are not controlled initially, given that the mother’s age is determined by the age of the mother when the child was born, which is potentially a choice variable, and that marital status may *depend* on income.[98]

---

[96] If a child spent an equal number of years living in two neighbourhood categories, they are classified as having lived in the higher (i.e. more advantaged) of the two. The number of youths living in ‘disadvantaged’ and ‘advantaged’ were 231 and 277, respectively, with the remaining 366 youths living in a ‘middle’ suburb.

[97] Single parent families are likely to have less time to invest in their children, and less capacity to engage in paid employment. In addition, emotional and behavioural problems associated with divorce or separation may lead to eating disorders. Mothers’ age could be correlated with the efficiency of home and market production.

[98] Mayer (1997, p.65) cites evidence that low income individuals are more likely to separate and divorce, and are also less likely to marry when they have a child, relative to high income couples. Thus, marital status may not remain constant after an exogenous change in income, and hence should not be held fixed if the objective is to measure the full, “policy-relevant”, effect of income. In the child health literature, some studies control for marital status (Bradbury 2007; Case *et al.* 2002) and some do not (Zhu 2007; Blau 1999; Mayer 1997).

5.5 Proxy Variables

This analysis uses an extensive set of proxy variables to account for as much heterogeneity as possible.

First, following Anderson *et al.* (2003), the models include maternal and paternal BMI as proxies for the genes and common home environment that parents and their children share.[99] [100] However, unlike genetics (which are predetermined), parental BMI can also be *influenced* by household income and maternal employment, and hence, the coefficients on parental BMI may capture (i.e. “pick up”) a portion of the income and labour supply effect.[101] For this reason, and since mothers’ country of birth should already capture part of the variance in youth obesity attributable to genetics, parental BMI is omitted from the ‘preferred’ specification.

Second, as a proxy for mothers’ intelligence, some models include the mother’s self-reported ability in reading and mathematics (relative to an “average” adult) (2 variables) and the extent that she agreed/disagreed with the statements: “I do not feel comfortable when working out amounts like discounts, the GST or percentages” (1 variable) and “I am good with numbers and calculations” (1 variable) (MAIESR 2009c, p.R197).[102]

---

[99] Parental BMI is measured concurrent with the youth’s BMI (Wave 6 or Wave 7), since a lagged measure of parental BMI could not be constructed. Data was available for the mother in 707 cases and for the father in 569 cases. Missing observations were set to zero and dummy variables denoting missing values were created.

[100] Omitting controls for genetics will lead to biased estimates if shared genetic factors are correlated with both household income/employment and youth BMI.

[101] Changes in maternal employment that affect a child’s BMI (i.e. from changes in eating patterns and time use) could also affect parents’ BMI; and a change in household income will improve the quality of goods consumed by both the parent and the child. For this reason, controlling for parental BMI may be over-controlling for the effects of the common home environment.

[102] The data are from Wave 7, with all responses recoded to be measured on a 1-10 scale with higher scores reflecting higher (perceived) ability in the relevant domain. In the 106 cases in which data were unavailable, the relevant scores were set equal to zero and a (single) dummy variable equalling one for missing data was created. These responses will be influenced by the respondents’ education, self-confidence and the reference group that he or she deems to represent the “typical” or “average adult”.

Third, parental smoking (2 dummies) is used as a proxy for the parent's rate of time preference (Heckman 2007, p.5),[103] their attitudes and behaviours toward health, and also the child's initial health stock, given these variables may be correlated with smoking during pregnancy.[104]

Fourth, following the logic of Mayer (1997, p.82) the ratio of household income from Australian public transfers and foreign pensions to gross household income is used as a proxy for unmeasured parental characteristics, such as diligence, motivation and ability.[105]

Finally, four variables are used as proxies for the unobserved *quality* of home investments and/or the mother's tastes for employment. These include the mother's score on the mental health sub-scale of the Medical Outcomes Study Short Form (SF-36) (see Ware and Gandek 1998, p.903) (1 variable),[106] and binary indicator variables for whether the mother's own mother was in paid employment when the mother was aged 14 years (1 variable), whether the mother's own mother *and* father were present in her household when she was aged 14 (1 variable) and whether the mother worked at any time in the 12 months prior to the birth of the child (1 variable).[107]

---

[103] Higher rates of time preference are expected to be negatively correlated with income (due to lower human capital investment) and negatively correlated with investments in children's health, since less weight is given to future outcomes (Heckman 2007, p.5).

[104] In the full sample, 180 mothers and 159 fathers smoked cigarettes (during adolescence). In 147 cases this information was missing for the father (a dummy variable for missing data is used).

[105] This ratio is averaged over each financial year during adolescence. It indicates the degree of welfare reliance of a household, with ratios of 0 and 1 indicating no welfare reliance and complete welfare reliance, respectively. Half the youths in the sample had a transfer payments-to-gross income ratio of less than 0.082.

[106] Scores range from 0 to 100 with a score of 100 representing optimal functioning in terms of mental health and low scores indicating symptoms associated with anxiety and depression (Ware and Gandek 1998, p.903). This score is averaged over adolescence in each year in which a score was available.

[107] For the 58 respondents with missing information on pre-birth employment, a dummy denoting the presence of missing data was specified.

5.6 Instrumental Variables

This section details the instrumental variables for permanent RHED income and maternal employment used in this thesis.

5.6.1 Permanent RHED Income: Grandfathers' Occupation SES

The instruments for permanent RHED income relate to the SES of the youth's maternal and paternal grandfather. The key assumption is that differences in grandfathers' SES are associated with differences in permanent income but do not lead to changes in youth BMI/obesity, except indirectly through permanent income (Cameron and Trivedi 2009, p.173).[108]

Grandfather's social class is measured using the Australian Socioeconomic Index 2006 (AUSEI06), which provides a measure of the SES of the grandfather's *occupation*. This scale was developed using data from the 2006 Australian Census of Population and Housing based on the methodology developed for the International Socioeconomic Index (ISEI). It conceptualises occupation as the mediating factor that converts education inputs into monetary outputs (McMillan *et al.* 2009, p.126), and hence, should be highly correlated with parental income.

The grandfather's AUSEI06 score is calculated using the parent's responses to questions about the occupation of their own father when they were aged 14. The AUSEI06 scores, which are imputed in HILDA, range between 0 and 100, with higher scores denoting higher occupational status.[109] Separate variables are used to represent the occupation SES of the youth's maternal and paternal grandfather.[110] Full details on the construction of these variables, including the author's

---

[108] This assumption would be violated, if, for instance, the youth has frequent interaction with his or her grandfather (in which case there may also be a direct relationship).

[109] Medical practitioners are at the top of the scale (100), followed by other health professionals (94) and legal professionals (91), whereas labourers are at the bottom of the scale (0).

[110] The occupational status score of the maternal and paternal *grandmother* was not used, as preliminary analysis revealed it was only weakly related to permanent RHED income; reflecting that most grandmothers were probably homemakers. The overidentifying restrictions tests also indicated that these variables were not suitably exogenous; reflecting that grandmothers possibly have frequent interaction with the child (a direct effect on the outcome).

imputation of status scores where occupation data are unavailable, are available in Appendix B (Section 12.1.2).

One limitation with the AUSEI06 is that the current stratification processes in Australia in which the scale is based (data from the 2006 Census), may not suitably reflect the structure of the labour markets that prevailed when the mother and father were aged 14, due to gradual structural changes in the economy, in particular the emergence of more specialised occupations and more skill-intensive production.[111]

5.6.2 Maternal Employment: Local Economic Conditions

This thesis uses local economic conditions and variables external to the mother and family as instruments for part-time and full-time maternal employment. First, following several studies (Anderson *et al.* 2003; James-Bardumy 2005; Cawley and Liu 2007) the local unemployment rate is used; given that higher unemployment rates make it harder for women to find employment and that the unemployment rate should be uncorrelated with the unobserved determinants of youth obesity. However, as previous studies have found, the unemployment rate oftentimes exhibits little variation between states/areas and over time, and hence is inefficient. This is particularly true using HILDA for which the most disaggregated geographic-region identifier in the general release files is the 'major statistical region'.[112] For this reason, the unemployment rate is measured at the level of the mother's industry division (current or previous) and Australian State/Territory. The benefit of such disaggregation is that the instrument exhibits more variation, and hence, is more likely to produce efficient results.[113] The downside is that

---

[111] The advantage to using this scale, as opposed to controlling for occupational categories directly, is that it preserved degrees of freedom, explicitly models the 'occupational status' of those unemployed and not in the labour force, and can provide more variation that a set of dummies.

[112] The major statistical region identifies Sydney, Melbourne, Brisbane, Adelaide and Perth statistical divisions and the balance of each state, in addition to Tasmania, the Northern Territory and the ACT.

[113] To explore this proposition, measures of the overall (and female) unemployment rate by state and balance of state were also constructed. The coefficient of variation of these variables (0.15 to

mothers may exhibit a high degree of mobility between industries, and hence the unemployment rate in a given industry may not adequately reflect the labour market conditions they actually face. Full detail of the construction of this variable, which involves matching external data from the ABS (Cat. No. 6291.0.55.003, 2009d) to the mother's information, is given in Appendix B (Section 12.1.3).

The second instrument for mothers' employment is the average number of hours worked by employed females in the same industry, occupation and State/Territory as the mother (or previous industry-occupation if not employed). For instance, it may be the norm for a woman employed as a manager in the Victorian finance industry to work more hours per week than a woman working in sales in the retail industry in South Australia.[114] However, it could be argued that mothers self-select into industries and occupations based on unobserved factors also correlated with youth obesity. In this case, this instrument would not be suitably independent of the error process (the same could be argued of the first instrument). This cannot be resolved *a priori*. This variable is constructed by matching external data from the ABS (Cat. No. 6291.0.55.003, 2009d) to the mother, with full details in Appendix B (Section 12.1.4).

The first instrument is conceived as an instrument for the labour force participation decision; distinguishing working from non-working mothers. In comparison, the second instrument is conceived more as an instrument for hours worked, conditional on working. That is, it is used to identify the separate effects of part-time and full-time employment on the dependent variable.

0.16) was around three times smaller than the coefficient of variation for the unemployment measure used in this thesis (0.46).

[114] Often the mother will be faced with a decision to work the standard number of hours or not at all.

6 Descriptive Statistics

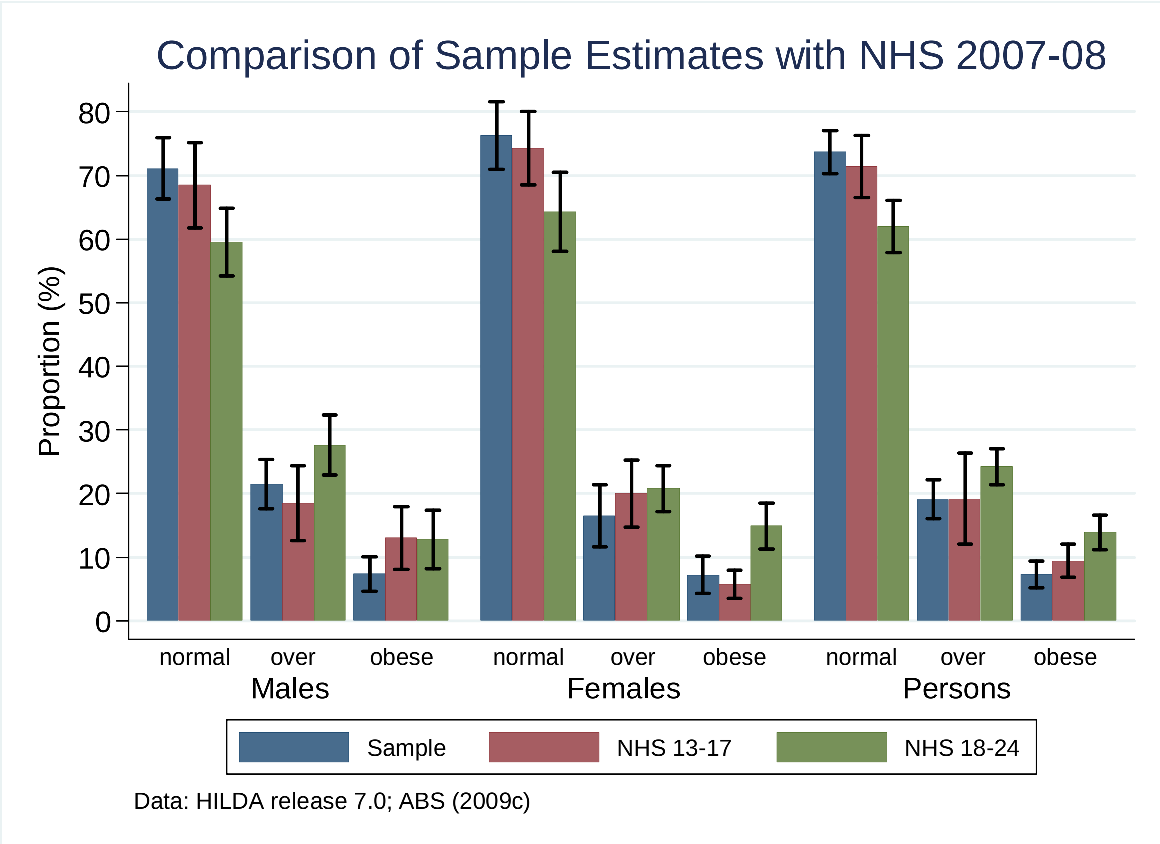
This chapter presents summary statistics for the main estimation sample. It is divided into four sections. Section 6.1 examines whether the sample proportions of overweight and obesity are representative of Australian youths, using benchmark data from a nationally representative health survey. Section 6.2 presents the descriptive statistics for the primary estimation sample. Section 6.3 provides a description of overweight and obese youths in Australia, including their health, wellbeing and life-satisfaction and behaviours and attitudes toward health. Lastly, Section 6.4 examines the relationships between maternal employment and parental time allocation within the household and the sample correlations between parental time allocation and BMI outcomes.

The HILDA survey used a complex sample design involving clustering and stratification to select the initial sample of households (Wilkins *et al.* 2009, p.vi). The implication is that standard errors cannot be validly constructed using formulae that apply to simple random sampling, even after the appropriate population weights have been applied. The standard errors of the estimates reported in this chapter are calculated using the Jackknife methodology. Full details of this procedure, including its formulae are presented in detail in Appendix C.

6.1 Comparison of HILDA Estimates with NHS (2007-08)

To determine whether the proportions of obese and overweight youths in the sample are representative of the broader population of Australian youths, Figure 6.1 compares the (population-weighted) sample proportions of overweight (excluding obesity) and obesity with the most recent National Health Survey (NHS 2007-08), conducted by the ABS (2009c, Table 17). Since the age categories reported in the NHS 2007-08 do not exactly correspond to the ages of sample members, two sets of benchmark estimates are used: first for persons aged 13-17 years and second for persons aged 18-24 years.[115] The 95 per cent confidence intervals are also indicated at the top of each bar in the figure.

Figure 6.1. Overweight and Obesity Prevalence Rates in Sample and NHS 2007-08



This analysis indicates that there is a significant under-representation of obese males in the main sample with 7.4 per cent of sample males classified as obese, compared to 13.0 per cent and 12.8 per cent of 13-17 and 18-24 year-old males in the NHS, respectively. These differences are significant at the 5 per cent levels of significance (p difference = 0.050 and 0.046, respectively), using independent z tests (Appendix C Equation 13.5). This is not surprising given that the NHS 2007-08 estimates are based on *measured* height and weight data (ABS 2009c, p.55), whilst the BMI data in the sample are self-reported.[116]

[115] The sample used in this thesis and the NHS 2007-08 estimates both use the overweight and obese cut-offs developed by Cole *et al.* (2000) (ABS 2009c, p.55).

[116] Respondents have a tendency to overstate height and understate weight (Wang *et al.* 2002b).

6.2 Characteristics of the Estimation Sample

Table 6.1 presents sample statistics for the variables used in the empirical analysis of BMI in Chapter 8. The three BMI categories correspond to those used in the estimated models, namely normal (and under-) weight, overweight or obese and obese. For brevity, the combined overweight and obese category is herein referred to as 'overweight'. Estimates are weighted, with standard errors calculated using the jackknife method.

Higher socioeconomic status is generally associated with lower BMI outcomes. Table 6.1 shows that youths who were in the normal weight range had mean permanent RHED income $1,440 higher, than overweight youths. Furthermore, normal weight youths were approximately 37 per cent less likely to be living in a 'disadvantaged' SEIFA suburb (p difference = 0.015) and 34 per cent more likely to have a University educated father (p difference = 0.003). The socioeconomic status gradient is even more pronounced when considering obese youths. However, these disparities need not reflect causal effects.

Table 6.1 also demonstrates that youths whose mothers supply large amounts of part-time labour tend to be less overweight, whilst non-working mothers tend to have heavier children. Yet, these differences are not statistically significant at the 10 per cent level.

There is no significant relationship between age or gender and BMI category, which reflects the use of age- and gender- adjusted cut-off points.[117] Indigenous status and mothers' country of birth are also not significantly different between BMI categories, possibly because of the small number of observations in each category (and hence a lack of statistical power) rather than a lack of relationship.

[117] The weighted correlation between BMI and age is 0.12, and is statistically different to zero.

Table 6.1. Descriptive Statistics of Explanatory Variables by BMI Category

	Normal Weight (N=634)	*Overweight (N=240)*	*Obese (N=69)*
Socioeconomic Status			
perminc ($000)	36.733 (0.867)	35.293 (1.476)	32.448 (1.979)**
ln(perminc)	10.432 (0.023)	10.372 (0.036)*	10.297 (0.043)***
SEIFA_disad	0.205 (0.024)	0.281 (0.034)**	0.388 (0.080)**
SEIFA_mid	0.469 (0.028)	0.438 (0.041)	0.470 (0.077)
SEIFA_adv	0.327 (0.029)	0.281 (0.038)	0.141 (0.039)***
Parental Education			
med_Yr 11	0.319 (0.026)	0.374 (0.036)	0.287 (0.078)
med_Yr 12	0.160 (0.019)	0.130 (0.026)	0.154 (0.045)
med_Cert	0.281 (0.021)	0.264 (0.032)	0.360 (0.077)
med_Uni	0.240 (0.023)	0.231 (0.034)	0.199 (0.048)
fed_Yr 11	0.187 (0.021)	0.268 (0.033)**	0.279 (0.065)
fed_Yr 12	0.079 (0.015)	0.068 (0.016)	0.073 (0.034)
fed_Cert	0.473 (0.031)	0.491 (0.037)	0.492 (0.085)
fed_Uni	0.261 (0.021)	0.172 (0.026)***	0.157 (0.048)**
Maternal Employment			
years NE (mum)	0.981 (0.068)	1.069 (0.121)	1.381 (0.263)
years PT (mum)	1.572 (0.082)	1.427 (0.121)	1.240 (0.206)
years FT (mum)	1.208 (0.087)	1.239 (0.102)	1.171 (0.221)
Demographic Variables			
female	0.513 (0.019)	0.447 (0.040)	0.491 (0.070)
age (years)	16.745 (0.062)	16.942 (0.116)	16.869 (0.288)
indigenous	0.036 (0.012)	0.068 (0.029)	0.057 (0.035)
mcob_Aust	0.749 (0.028)	0.757 (0.040)	0.740 (0.059)
mcob_Oceanic	0.026 (0.008)	0.045 (0.018)	0.083 (0.050)
mcob_NW Euro	0.064 (0.014)	0.058 (0.024)	0.016 (0.016)
mcob_SE Euro	0.039 (0.012)	0.027 (0.016)	0.014 (0.014)
mcob_Asia	0.076 (0.014)	0.065 (0.022)	0.075 (0.044)
mcob_Other	0.076 (0.014)	0.065 (0.022)	0.075 (0.044)
Supplementary Variables			
age mother (years)	43.731 (0.277)	43.568 (0.404)	42.917 (0.772)
years married	2.761 (0.087)	2.823 (0.111)	2.741 (0.259)
years NE (dad)	0.194 (0.037)	0.299 (0.074)	0.387 (0.153)
years PT (dad)	0.137 (0.022)	0.145 (0.037)	0.108 (0.048)
years FT (dad)	2.474 (0.082)	2.467 (0.129)	2.309 (0.267)
Proxy Variables			
BMI mum	26.238 (0.333)	27.989 (0.553)***	30.916 (1.194)***
BMI dad	27.689 (0.287)	29.207 (0.555)***	30.156 (0.726)***
transfers (units)	0.200 (0.013)	0.221 (0.027)	0.256 (0.045)
mental health	72.508 (0.723)	73.838 (1.284)	72.149 (2.024)
mum smoke	0.206 (0.020)	0.219 (0.041)	0.256 (0.069)
dad smoke	0.196 (0.023)	0.291 (0.037)**	0.214 (0.057)
prebirth	0.661 (0.023)	0.713 (0.029)	0.738 (0.063)
mmemploy	0.498 (0.024)	0.455 (0.040)	0.447 (0.070)
bothparents	0.832 (0.022)	0.809 (0.029)	0.865 (0.048)
reading	8.358 (0.095)	8.381 (0.131)	8.369 (0.191)
math	7.184 (0.091)	7.283 (0.171)	7.144 (0.286)
GST	7.262 (0.126)	6.962 (0.206)	6.930 (0.352)
numbers	7.706 (0.112)	7.556 (0.160)	7.254 (0.342)

Notes: *, ** and *** means estimates for normal-weight and overweight or obese (or for normal-weight and obese) are statistically different at 10%, 5% and 1% levels of significance, respectively. Estimates weighted using enumerated person longitudinal weights; standard errors calculated using the jackknife procedure (in parentheses). Variable definitions are provided in Table 14.1 (Appendix D). 'years NE' denotes the number of years the mother/father is not in paid employment. The 'overweight' sample includes obese.

As expected, there is a strong positive correlation between parental BMI and youth overweight/obesity, highlighting the important role for the genetics and/or the common household environment that parents and their children share. The elasticity of the youth's BMI with respect to their mother's and father's BMI is 0.19 and 0.24, respectively.[118]

The remainder of Table 6.1 demonstrates few significant differences in observed characteristics between overweight/obese and normal weight youths, with the exception of paternal smoking, which is associated with increased overweight (p difference = 0.024).

6.3 Who are the Obese? A Profile of Overweight Youths in Australia

An overweight-health gradient is important to validating the theoretical approach taken in this analysis, which assumes overweight to enter into health inversely. To examine the extent to which overweight and obesity is associated with lower health and wellbeing in Australian youths, Section 6.3.1 presents a brief descriptive analysis. The health-related behaviours and attitudes of overweight and obese Australian youths are analysed in Section 6.3.2.

6.3.1 Health, Wellbeing and Life-Satisfaction

Table 6.2 presents a selection of health, wellbeing and life-satisfaction indicators by BMI category, for the main sample of youths (N=874). In most cases these are measured concurrently with BMI, and derive from the youth's own responses to HILDA questionnaires.[119]

---

[118] Elasticities are obtained by probability-weighted OLS regression of log youth BMI on log parents' BMI. Both elaticities are significant at the 0.1 per cent level. Excluding step fathers from the regression sample increases the youth-father elasticity to 0.30, suggesting a role for both genetics and common environment in this observed correlation. Excluding step-mothers produces no significant change in the youth-mother elasticity.

[119] Respondents with missing information on any of these characteristics are omitted when computing sample statistics.

Table 6.2. Health, Wellbeing and Life Satisfaction

	Normal Weight (N=634)	*Overweight (N=240)*	*Obese (N=69)*
SF-36 Health Scale			
General Health	74.85 (0.70)	71.68 (1.32)**	66.43 (2.76)***
Physical Functioning	93.19 (0.92)	90.59 (1.53)	83.98 (4.54)**
Role Physical	91.84 (1.05)	92.08 (1.60)	90.20 (3.65)
Bodily Pain	82.54 (0.78)	85.32 (1.45)*	83.43 (2.63)
Vitality	65.05 (0.89)	64.91 (1.26)	63.91 (2.37)
Mental Health	74.89 (0.64)	74.17 (1.32)	70.04 (2.12)**
Role Emotional	86.85 (1.14)	88.16 (1.62)	85.64 (3.82)
Social Functioning	87.32 (0.67)	88.55 (1.50)	84.54 (2.70)
Kessler Psychological Distress Scale (K10)			
Low	0.58 (0.02)	0.57 (0.04)	0.44 (0.06)**
Moderate	0.24 (0.02)	0.24 (0.03)	0.29 (0.06)
High	0.14 (0.02)	0.13 (0.03)	0.16 (0.05)
Very High	0.04 (0.01)	0.07 (0.02)	0.11 (0.04)
Medically Diagnosed Health Conditions			
Asthma	0.17 (0.02)	0.21 (0.03)	0.19 (0.06)
Depression, Anxiety	0.05 (0.01)	0.05 (0.02)	0.05 (0.03)
Limiting health condition	0.19 (0.02)	0.24 (0.03)	0.28 (0.07)
Social Indicators			
Member of a Club	0.48 (0.02)	0.42 (0.04)	0.31 (0.07)**
Many Friends	5.44 (0.06)	5.15 (0.09)***	5.10 (0.18)*
Life Satisfaction			
Satisfaction: Overall	8.30 (0.05)	8.13 (0.09)*	8.14 (0.20)
Satisfaction: Health	8.37 (0.07)	7.90 (0.13)***	7.63 (0.22)***
Satisfaction: Community	7.00 (0.12)	6.81 (0.14)	6.65 (0.31)

Notes: see note on Table 6.1.

A mean comparison of the eight sub-scales derived from the SF-36 Health Survey (Ware and Gandek 1998, p.903) indicates that overweight and obese youths are in poorer general health, relative to their normal weight counterparts.[120] For obese youths, the mean difference exceeds *five* for 'general health' - the threshold indicating "clinical or social meaningfulness" in the SF-36 (Butterworth and Crosier 2004, p.7). The mean difference also exceeded five for 'physical functioning' between obese and normal weight youths (p difference = 0.039). Significant health penalties were also found for overweight youths in terms of bodily pain (p difference = 0.099), and for the obese in terms of mental health (p difference = 0.026), though these differentials were not large enough to be clinically meaningful.

[120] Scores on each of the eight SF-36 sub-scales range between 0 and 100, with 100 representing optimal functioning. The SF-36 was developed to meet the psychometric standards required for group comparisons, to quantify relative disease burden and to enable profiling of functional health and well-being (Ware and Gandek 1998, p.903). The first four scales represent dimensions of physical health, whereas the latter four represent dimensions of mental health.

Obese youths were around 24.1 per cent less likely to exhibit low levels of psychological distress, relative to youths of normal weight (p difference = 0.021), according to the Kessler Psychological Distress Scale (K10).[121] Conversely, overweight and obese youths were not significantly more likely to be diagnosed with asthma, depression/anxiety or a limiting health condition, than a normal weight youth.[122] [123]

From a social perspective, overweight and obese youths were less likely to report being a current member of a sporting, hobby or community-based organisation, or to have ‘a lot of friends’, relative to youths of normal weight. Overweight and obese youths were also relatively less satisfied with their life, and, in particular, their health.

6.3.2 Health Behaviours and Attitudes

Table 6.3 shows that overweight and obese youths are significantly more likely to be dissatisfied with their current weight, relative to normal weight youths (p difference = 0.000). Surprisingly, half of overweight youths and a quarter of obese youths do not consider themselves to be overweight, which may reflect either a lack of information regarding what constitutes an ‘acceptable weight’ or a degree of measurement error in the self-reported height and weight data.

Overweight and obese youths are more likely than normal weight youths to engage in dieting behaviour (p difference = 0.000), they are more likely to skip breakfast and they consume smaller quantities of vegetables. Conversely, there is an interesting lack of relationship between exercise frequency and overweight

---

[121] The Kessler Psychological Distress Scale (K10) is a 10-item measure of non-specific psychological distress. It seeks to “measure the level of current anxiety and depressive symptoms a person may have experienced in the four weeks prior to interview” (Wooden 2009, p.2). It is effective at screening for serious mental disorders, as it is designed to have optimal sensitivity at the upper end of the population distribution in regards to psychological distress (Wooden 2009, p.2).

[122] The relevant questions referred only to medically diagnosed conditions that are expected to last six months or more, and hence they are likely to result in underestimates of the conditions.

[123] It is important to note that many of the adverse health consequences of obesity, such as cardiovascular disease, do not develop until adulthood (Daniels 2006, p.47).

prevalence; and between overweight/obesity status and the propensity to purchase meals from an outlet. There were also no significant differences in risky behaviour, such as smoking and binge drinking, between BMI categories.

Table 6.3. Health Behaviours and Attitudes

	Normal Weight (N=634)	*Overweight (N=240)*	*Obese (N=69)*
Body Weight			
Satisfaction with current weight;			
satisfied / very satisfied	0.52 (0.02)	0.32 (0.04)***	0.12 (0.04)***
neither satisfied nor dissatisfied	0.31 (0.02)	0.27 (0.03)	0.34 (0.06)
dissatisfied / very dissatisfied	0.17 (0.02)	0.41 (0.04)***	0.54 (0.07)***
Self-rated weight status;			
acceptable or underweight	0.93 (0.01)	0.50 (0.04)***	0.25 (0.05)***
overweight	0.07 (0.01)	0.50 (0.04)***	0.75 (0.05)***
Diet and Exercise			
Number of diets in past 12 months;			
none	0.83 (0.02)	0.58 (0.04)***	0.43 (0.08)***
one	0.09 (0.01)	0.21 (0.03)***	0.25 (0.06)***
> one	0.08 (0.01)	0.21 (0.03)***	0.31 (0.06)***
Purchase meals from outlet (days per week);			
breakfast	0.42 (0.05)	0.41 (0.07)	0.40 (0.11)
lunch	2.07 (0.09)	1.94 (0.13)	2.11 (0.26)
dinner	1.09 (0.07)	1.04 (0.07)	1.05 (0.18)
Eats breakfast (days per week)	5.28 (0.12)	5.02 (0.18)	4.58 (0.24)***
Serves of fruit per day	1.35 (0.06)	1.41 (0.08)	1.30 (0.12)
Serves of vegetables per day	1.97 (0.06)	1.88 (0.08)	1.58 (0.16)**
Exercise frequency;			
< once week	0.14 (0.01)	0.14 (0.02)	0.16 (0.05)
1 - 3 times week	0.41 (0.03)	0.42 (0.04)	0.44 (0.09)
> 3 times week	0.45 (0.02)	0.45 (0.03)	0.41 (0.06)
Risky Behaviours			
Smokes Cigarettes	0.14 (0.02)	0.15 (0.03)	0.14 (0.04)
Binge drinks alcohol	0.69 (0.02)	0.68 (0.03)	0.62 (0.08)

Notes: see note on Table 6.1.

6.4 Maternal Employment and Parental Time Investments: Is there a Relationship?

This section explores how full-time, part-time and non-employed mothers (and their partners) allocate their non-market time, using the time use data in HILDA. The objective is to determine whether increased maternal work intensity (L_{Wt}) is associated with a corresponding decrease in parental time investments in child health (L_{Qt}), which is the assumption underlying the empirical models.

The time use data derive from the time-use component of the Self-Completion Questionnaire (SCQ), where respondents are asked how much time they would spend in a certain activity in a "typical week", for each of nine separate categories of activities.[124] These questions were included in all seven waves of HILDA, and hence an *average* of time the mother allocated to each activity during adolescence (in the years she was present in the household and responding), is measured. Similar averages are computed for the child's father.

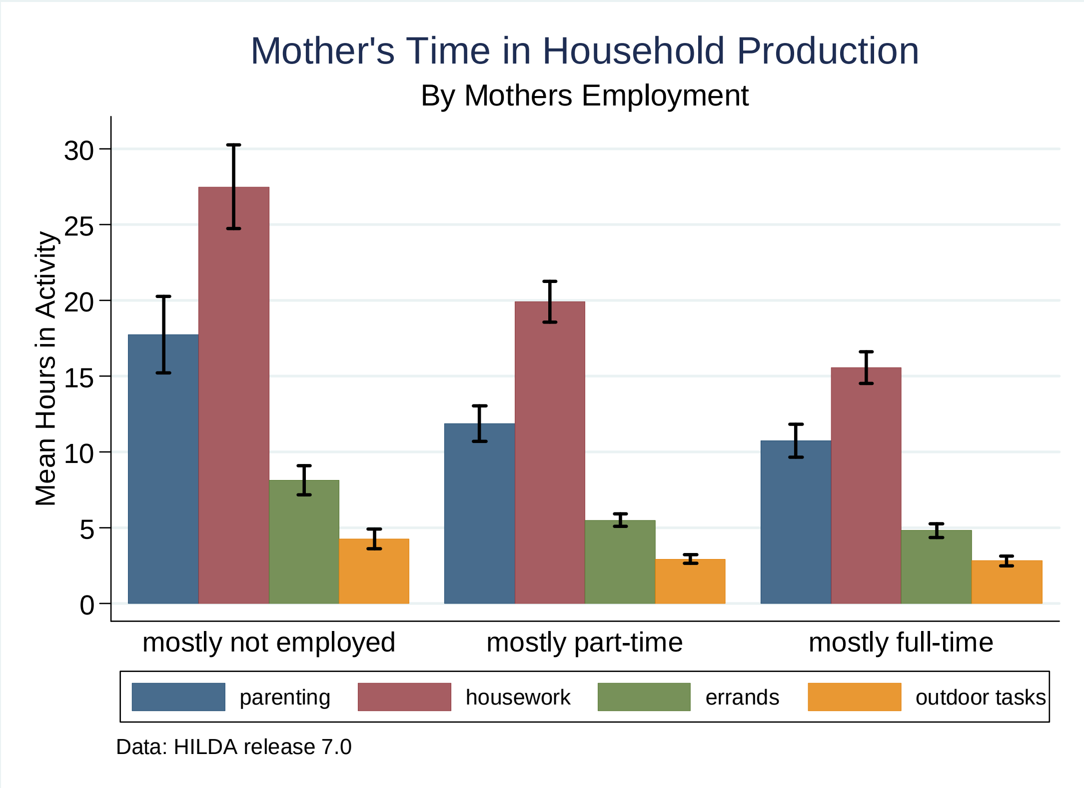
Time spent in the following activities is of interest; parenting and playing with children; housework; household errands; and outdoor tasks.[125] Each of these activities is assumed to contribute to child health in some way (L_{Qt}), compared with other forms of non-market time, such as time in volunteer and charity work; commuting to paid employment; playing with other people's children; and caring for elderly adult relatives (L_{Lt}). Figure 6.2 presents the mother's average time spent in each activity, by her employment status, along with the associated 95 per cent confidence bands.[126]

[124] These time use data are likely to be of lower quality than data collected using time-diaries (such as the ABS Time Use Surveys and the LSAC Time Use Diary), in which respondents record the activities they partook over specified time intervals over a given period (Siminski 2006, p.2),

[125] 'Parenting and playing with children' includes playing with own children, helping them with personal care, teaching, coaching or actively supervising them, or getting them to child care, school and other activities. 'Housework' includes preparing meals, washing dishes, cleaning the house, washing clothes and ironing. 'Outdoor tasks' includes home maintenance (repairs, improvements, painting etc.), car maintenance or repairs and gardening. 'Household errands' includes shopping, banking, paying bills, and keeping financial records but does not include driving children to school and to other activities (MIAESR 2009b, p.9).

[126] A mother is 'mostly full-time employed' if she is employed full-time for at least three of the four adolescent years, or two of the three adolescent years in the case of 19-year olds. Similarly, a mother is 'mostly part-time employed' or 'mostly not employed' if she works part-time, or not at all, for at least three of the four years, respectively (or two of the three years in the case of 19-year olds).

Figure 6.2.



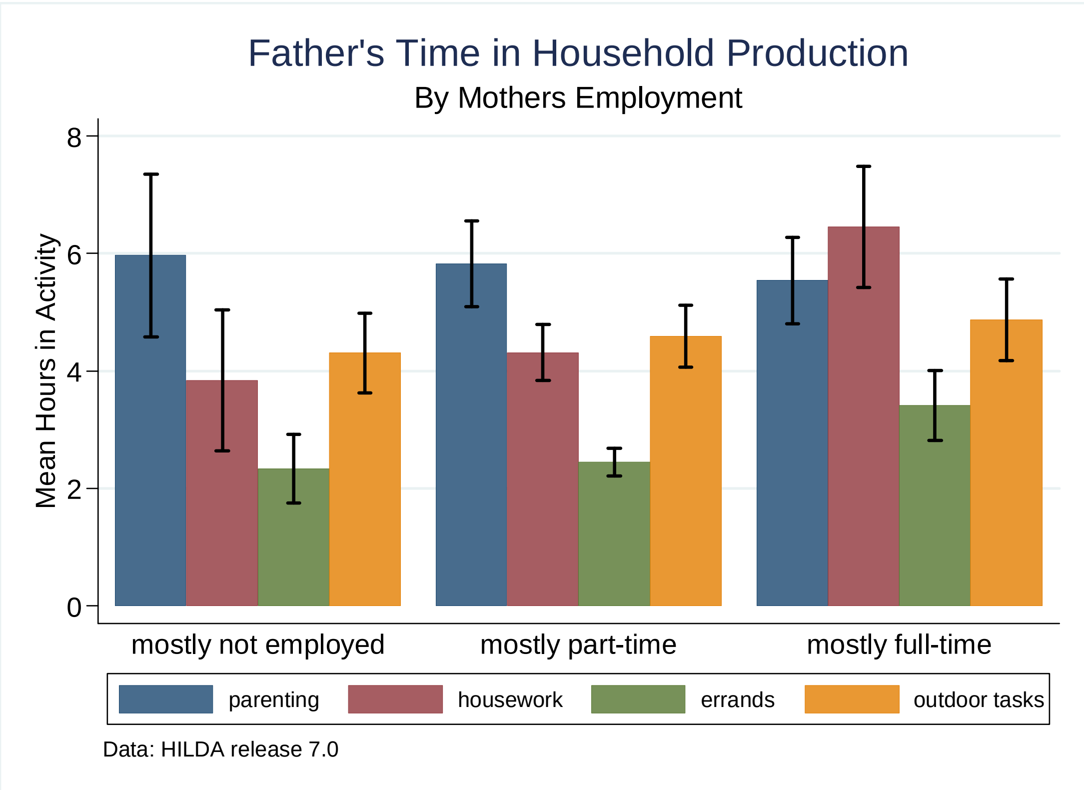
Mothers 'mostly not employed' during adolescence spent an average of 17.8 hours a week in parenting activities. This was around 50 per cent more time than mostly part-time employed mothers (11.9 hours per week) and 66 per cent more time than mostly full-time employed mothers (10.7 hours per week), suggesting that maternal time investments are inversely related to work intensity. Interestingly however, the relationship is highly nonlinear, with full-time working mothers spending on average only 1.2 fewer hours in parenting activities than part-time employed mothers, despite spending an additional 21.5 hours in employment-related activities, on average, in a typical week.[127] Similar stories can be told for time spent in housework, household errands and outdoor tasks. This suggests that, compared with mothers in part-time employment, mothers in full-time employment may act to "protect" the most productive time investments in their children, by cutting back least on activities directly involved in producing health. This raises important questions regarding the validity of maternal employment as a

[127] Employment related activities include both time spent in employment and time spent travelling to and from the place of work.

proxy for time investments in the child, which will be discussed further in Chapter 8.

Furthermore, there is evidence of substitution between maternal and paternal time as maternal work intensity increases, as shown in Figure 6.3, which depicts the average time the father spends in each activity, by *mothers'* employment.

Figure 6.3.



Fathers with full-time employed partners spend significantly more time doing housework and household errands in a typical week, compared to fathers with non-working or part-time employed partners. Hence, there appears to be at least some substitutability between parents. Figure 6.3 also shows that fathers spend little time playing with their children, regardless of how much the mother works.

6.4.1 Health Production and Youth BMI Category

Finally, Table 6.4 presents the mean time the mother and father spent in each household activity, by youth BMI category. This indicates no clear relationship between parental time investments and the overweight and obesity outcomes of children. However, this does not imply the absence of a causal relationship.[128]

Table 6.4. Time in Household Activities by BMI Category

	Normal Weight (N=634)	*Overweight (N=240)*	*Obese (N=69)*
Mother's Time Investment (hours)			
Parenting	13.09 (0.54)	12.57 (0.87)	12.60 (1.45)
Housework	20.45 (0.51)	20.31 (0.93)	20.37 (2.14)
Household errands	6.05 (0.20)	5.70 (0.29)	5.40 (0.64)
Outdoor tasks	3.46 (0.14)	3.16 (0.17)	2.67 (0.33)
Total time	43.08 (1.03)	41.73 (1.76)	41.04 (3.88)
Father's Time Investment (hours)			
Parenting	7.07 (0.31)	6.87 (0.60)	7.67 (1.20)
Housework	6.10 (0.37)	6.22 (0.60)	5.87 (0.75)
Household errands	3.43 (0.28)	3.53 (0.31)	3.48 (0.46)
Outdoor tasks	5.72 (0.22)	5.16 (0.29)	6.02 (0.75)**
Total time	22.30 (0.91)	21.77 (1.42)	23.04 (2.33)

Notes: see note on Table 6.1.

[128] Other factors, such as mothers' innate ability, may be correlated with both time spent in certain activities and her child's BMI.

7 Methodology

This section outlines the methods to be used in the empirical analysis in Chapter 8. Section 7.1 and Section 7.2 outline the baseline least squares and probit specifications, respectively. The linear and non-linear IV specifications are outlined in Section 7.3 and Section 7.4.

7.1 Ordinary Least Squares (OLS)

Firstly, linear models of the form;

$$\begin{aligned} y_i &= \alpha + \delta \mathrm{I}_i + \mathbf{E}_i \boldsymbol{\gamma} + \mathbf{x}_i \boldsymbol{\beta} + u_i \\ &= \mathbf{x}_i' \boldsymbol{\lambda} + u_i \end{aligned} \tag{7.1}$$

are specified, where $i = 1,...,n$ and y_i is a continuous dependent variable measuring BMI, $\mathbf{x}_i$ is a vector of exogenous individual, household and neighbourhood variables associated with the i^{th} youth, I_i is a continuous variable denoting the natural log of permanent RHED income, $\mathbf{E}_i$ is a vector of maternal employment variables, u_i is a disturbance and α, δ, $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are the parameters to be estimated. The effect of a one per cent increase in permanent income on BMI is given by $\delta/100$. $\boldsymbol{\gamma} = \begin{pmatrix} \gamma^{PT} \\ \gamma^{FT} \\ \gamma^{MS} \\ \gamma^{NR} \end{pmatrix}$ denotes the 4×1 vector of estimated parameters associated with $\mathbf{E}_i$, where γ^{PT} and γ^{FT} give the effect on BMI *ceteris paribus* of one additional year in part-time or full-time employment, relative to not working, respectively.[129]

[129] γ^{MS} and γ^{NR} give the effect on BMI of an additional year not present in the household and not-responding to the Continuing Person Questionnaire, respectively.

Estimation of (7.1) is by the method of least squares, with cluster-robust standard errors[130] and probability weights. The estimation procedure is carried out by sequentially adding sets of regressors, as mentioned in Section 5.4.

7.2 Single Equation Probit Model

As discussed in Section 3.1.3, the dependent variable y_i can alternatively be modelled as a binary outcome that equals one if the i^{th} youth is overweight ($y_i = 1$) and zero otherwise ($y_i = 0$).[131] Regression analysis of a binary dependent variable requires a probability model, which is specified as a probit model in this thesis.

Using the notation in (7.1) and assuming the error term u_i is $N(0,1)$, the probability of the i^{th} youth being overweight is given by evaluating the area under the standard normal distribution between $-\infty$ and the index value, $\mathbf{x}_i'\boldsymbol{\lambda}$. This is equivalent to evaluating the value of the standard normal *cumulative* distribution function at $\mathbf{x}_i'\boldsymbol{\lambda}$. That is, the probability of the youth being overweight is given by;

$$\begin{aligned} \text{Prob}\left(y = 1 \middle| \mathbf{x}_i'\right) &= \text{Prob}\left(u_i < \mathbf{x}_i'\boldsymbol{\lambda} \middle| \mathbf{x}_i'\right) \\ &= \int_{-\infty}^{\mathbf{x}_i'\boldsymbol{\lambda}_i} \varphi(\text{z})\text{dz} \\ &= \Phi\left(\mathbf{x}_i'\boldsymbol{\lambda}\right) \end{aligned} \tag{7.2}$$

where $\varphi(\text{z}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ is the standard normal density function, and $\Phi(\text{z}) = \int \varphi(\text{z})$ is the cumulative distribution function (CDF).

---

[130] Since the data set contains siblings, all standard errors (in all estimated models in Chapter 8) are clustered on the mother's identification code. Failing to do this will lead to underestimated standard errors, since the observations of siblings may be correlated (Baum and Ford 2004, p.888). Standard errors reported in this analysis are also robust to heteroskedasticity (Cameron and Trivedi 2009, p.83). Standard errors are *not* calculated using the Jackknife procedure in Chapter 8 since Jackknife standard errors (which are 'clustered' on a higher unit of analysis i.e. the Census Collection District) are downwardly biased, relative to standard errors clustered on the mother (the lowest unit of analysis).

[131] Alternatively, the outcomes could be obese or not obese, respectively.

7.2.1 Maximum Likelihood Estimation

The probit model is a non-linear regression model and is estimated using maximum likelihood estimation (MLE) rather than OLS (Wooldridge 2006, p.586; Stock and Watson 2007, pp.396-7). The general theory of MLE for random samples implies that, under very general conditions, the MLE is consistent, asymptotically normal and asymptotically efficient (Wooldridge 2002, p.385).

7.2.2 Interpretation of Probit Estimates; Marginal Effects

In non-linear regression models, the marginal (or partial) effect does not equal the relevant slope coefficient. Given $\text{Prob}\left(y_i = 1 \middle| \mathbf{x}_i'\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\lambda}\right)$ (Equation 7.2) and using the chain rule of differentiation, the marginal effect on the probability of a success for a unit change in a continuous independent variable x_j is given by;

$$\frac{\partial p_i}{\partial x_j} = \frac{\partial \Phi\left(\mathbf{x}_i'\boldsymbol{\lambda}_i\right)}{\partial\left(\mathbf{x}_i'\boldsymbol{\lambda}_i\right)} \cdot \frac{\partial\left(\mathbf{x}_i'\boldsymbol{\lambda}_i\right)}{\partial x_j} = \varphi\left(\mathbf{x}_i'\boldsymbol{\lambda}_i\right) \cdot \lambda_j \tag{7.3}$$

Since the evaluation of the standard normal density function at $\mathbf{x}_i'\boldsymbol{\lambda}_i$ depends on the particular values taken by each explanatory variable, the marginal effect of a unit change in x_j will depends on the assumed values of all other explanatory variables (Gujarati 2003, p.611).

However, if x_j is a dummy variable, its marginal effect is calculated as the difference between the predicted probability of success when $x_j = 1$ and the predicted probability of success when $x_j = 0$, holding all other variables constant at their assumed values.

7.3 Two-Stage Least Squares (2SLS)

The OLS and probit model above both assume the RHS variables to be uncorrelated with the disturbance term, a condition necessary for consistent estimation of the parameters. This assumption is relaxed using the two-stage least squares (2SLS) and IV-probit models described in this section and in Section 7.4, respectively, which explicitly allow the unobservable error to be correlated with permanent income, maternal employment or both.

Consider again the structural model (7.1), which is reproduced below with slightly different notation;

$$y_{1i} = \alpha + \mathbf{y}_{2i}\boldsymbol{\beta}_1 + \mathbf{x}_i\boldsymbol{\beta}_2 + u_i \qquad i = 1, \ldots, n \qquad (7.4)$$

where y_{1i} is the BMI of the i^{th} youth, $\mathbf{y}_{2i}$ is a vector of m endogenous variables, that could include permanent income I_i and/or maternal employment $\mathbf{E}_i$, $\mathrm{Cov}(\mathbf{y}_{2i}, u_i) \neq 0$; and $\mathbf{x}_i$ is a k_1 vector of exogenous individual, household and neighbourhood variables associated with the i^{th} youth, $\mathrm{Cov}(\mathbf{x}_i, u_i) = 0$. Endogeneity of $\mathbf{y}_{2i}$ can come from any of the sources discussed in Section 3.4.

Estimating the structural model using OLS will yield inconsistent estimators of $\boldsymbol{\beta}_1$. In this situation, the method of instrumental variables (IV) can provide consistent estimation. This approach requires at least m observable instrumental variables $\mathbf{z}_i$ (one for each endogenous variable) satisfying the conditions of relevance and exogeneity, as discussed in Section 3.5.3.

For all linear regression models used in this analysis, the IV estimator is Two-Stage Least Squares (2SLS).[132] This estimator is obtained in two stages. First, a first-stage regression (i.e. reduced-form model) is estimated for each endogenous variable in

[132] 2SLS is most efficient in the class of all linear IV estimators under the assumptions of independence and homoscedasticity of u_i (Wooldridge 2002, p.96).

Equation (7.4), relating that regressor to the k_1 exogenous variables in the structural model and the k_2 instrumental variables, where $k_2 \geq m$;[133]

$$y_{2ji} = \mathbf{x}_i \boldsymbol{\pi}_{1j} + \mathbf{z}_i \boldsymbol{\pi}_{2j} + v_{ji}, \qquad j = 1,...,m \tag{7.5}$$

where $\boldsymbol{\pi}_{1i}$ and $\boldsymbol{\pi}_{2i}$ are (vectors of) unknown regression coefficients and v_{ji} is a zero mean error term(s). The fitted values, $\widehat{y_{2ji}}$, are obtained from OLS regression of (7.5).

In the second stage of 2SLS, the structural equation is estimated using OLS after replacing $\mathbf{y}_{2i}$ with the respective fitted values, $\widehat{\mathbf{y}_{2i}}$;

$$y_{1i} = \omega_0 + \widehat{\mathbf{y}_{2i}} \boldsymbol{\omega}_1 + \mathbf{x}_i \boldsymbol{\omega}_2 + \varepsilon_i \tag{7.6}$$

where $\omega_0, \boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are the 2SLS estimators of α, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, respectively. If the assumptions of instrument relevance and exogeneity are satisfied, the parameters will be consistently estimated by OLS of (7.6), and the 2SLS estimator will be normally distributed in large samples (Stock and Watson 2007, p.464). [134]

It is also important to note that 2SLS yields consistent point and interval estimates of a discrete endogenous regressor (i.e. binary or count data), which for this analysis includes maternal employment, $\mathbf{E}_i$ (Baum 2007, p.95).

---

[133] The regression coefficients must be exactly identified, $k_2 = m$, or overidentified, $k_2 > m$, for IV regression to be possible (Verbeek 2008, p.146).

[134] The standard errors reported by OLS estimation of the second-stage regression are incorrect because they do not recognise that it is the second stage of a two-stage process (Stock and Watson 2007, p.437). The correct formula standard errors is applied automatically in STATA.

7.4 IV-Probit

Consider the latent variable version of the model developed in Section (7.2). The departure from the standard probit is that permanent income, y_2, is modelled as an endogenous variable, which is assumed to be a linear function of the exogenous variables that determine overweight $\mathbf{x}_1$, a set of other exogenous variables $\mathbf{z}$ (i.e. instruments), and an error term u_2;

$$y_1{*} = \mathbf{x}_1\boldsymbol{\beta}_1 + y_2\beta_2 + u_1 \tag{7.7}$$
$$y_2 = \mathbf{x}_1\boldsymbol{\delta}_1 + \mathbf{z}\boldsymbol{\delta}_2 + u_2 \tag{7.8}$$

$$y_1 = \begin{cases} 1 & \text{if } \; y_1{*} > 0 \\ 0 & \text{if } \; y_1{*} \leq 0 \end{cases} \tag{7.9}$$

Again, the identifying assumption is that the instruments $\mathbf{z}$ are correlated with y_2 but are uncorrelated with the residual probability the youth is overweight, u_1.[135]

The IV-probit assumes u_1 and u_2 have a bivariate normal distribution, each with mean zero and covariance $\text{cov}(u_1, u_2)$, where

$$\text{cov}(u_1, u_2) = \begin{bmatrix} 1 & \rho\sigma_2 \\ \rho\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Hence, this structural-model approach completely specifies the distributions of $y_1{*}$ and y_2 in Equations (7.7) and (7.8). As a result, the approach relies heavily on the distributional assumptions, with consistent estimation requiring both multivariate normality and homoskedasticity of the errors u_1 and u_2 (Cameron and Trivedi 2009, p.467). Since u_2 is normally distributed, it is assumed that y_2 given $\mathbf{x}_1$ and $\mathbf{z}$ is normal, and hence y_2 should exhibit features of a normal random variable. For this reason, maternal employment – a *discrete* variable – should not be endogenously modelled using IV-probit. Instead, to account for

[135] Equations (7.7) and (7.9) constitute the structural equation, whilst (7.8) is the reduced form equation for permanent income. If u_1 and u_2 are correlated y_2 is endogenous; otherwise y_2 is exogenous (Wooldridge 2002, p.472).

endogeneity in maternal employment (with a binary outcome), a simple IV linear probability model is used. Conversely, IV-probit is appropriate for modelling the endogeneity in permanent RHED income, which is specified as a continuous variable – hence joint normality of (u_1, u_2) is plausibly satisfied. The advantage of the IV-probit over the IV-LPM model is that the latter approach ignores the binary nature of the dependent variable.

7.4.1 Maximum Likelihood Estimation of IV-Probit

The IV-probit model in this thesis is estimated using maximum likelihood (MLE). MLE is more efficient than any two-stage procedure and provides direct estimates of the parameters of interest for computing average partial effects (Wooldridge 2002, p.476). The MLE estimators are consistent and efficient as long as the underlying distributional assumptions hold (Wooldridge 2002, p.477).

7.4.2 Limitations of IV Analysis

Aside from the difficulty in locating suitable instruments, IV can suffer a number of limitations. First, the IV estimator is *inconsistent* if the instruments are correlated with the error term in Equation (7.6) (Wooldridge 2002, pp.101-2). In this case, IV achieves nothing since the instrumented variable(s) are still endogenous (Bound *et al.* 1993).

Second, the IV estimator will not be normally distributed in large samples if the instruments are *weak*, that is, only marginally relevant (Stock and Watson 2007, p.436). This implies there will be no theoretical justification for the usual methods of statistical inference, even asymptotically (Stock and Watson 2007, p.440). In addition, a weak correlation between the instrument(s) and the endogenous regressor(s) will exacerbate any problems associated with a correlation between the instrument and the residual in the second-stage regression (Bound *et al.* 1993; Wooldridge 2002, p.102).

Third, most of the justification for use of IV is asymptotic and its performance in small samples can be poor (Wooldridge 2002, p.101; Baum 2007, p.35). Notably, even when IV estimators are consistent, they can be biased in finite samples. This arises because the IV estimator is not centred on $\boldsymbol{\beta}_1$ in finite samples even though it is consistent for $\boldsymbol{\beta}_1$ in infinite samples (Cameron and Trivedi 2009, p.176).

Appendix E outlines a set of statistical tests that can be used to assess the validity of the assumptions underlying the IV approach. These tests, which are used in the empirical analysis, include the test for weak instruments (Section 15.1.1), the test for instrument exogeneity (Section 15.1.2) and the test for regressor endogeneity (Section 15.1.3-15.1.4).

7.5 Sibling-Difference Models

The final approach to accounting for omitted variables in this thesis assumes that unobserved heterogeneity is family-specific. That is, Equation (7.1) contains a component α_f which represents family-specific unobserved heterogeneity;

$$y_i = \delta \mathrm{I}_i + \mathbf{E}_i \boldsymbol{\gamma} + \mathbf{x}_i \boldsymbol{\beta} + \alpha_f + u_i \qquad (7.10)$$

If a correlation between $\mathbf{E}_i$ and α_f exists, and if the family heterogeneity component is unobserved, then estimates of the labour supply effect will be biased. To control for this source of bias the fixed-effects technique takes the difference between the BMI/overweight/obese outcomes of *siblings* in the sample to difference out the family fixed-effect, α_f;

$$y_i - y_j = \delta\left(\mathrm{I}_i - \mathrm{I}_j\right) + \left(\mathbf{E}_i - \mathbf{E}_j\right)\boldsymbol{\gamma} + \left(\mathbf{x}_i - \mathbf{x}_j\right)\boldsymbol{\beta} + \left(\alpha_{if} - \alpha_{jf}\right) + \left(u_i - u_j\right) \qquad (7.11)$$

for siblings i and j at the same point in time (Anderson *et al.* 2003; Blau 1999). In this fixed-effects technique the parameters are identified by variation within sibling-pairs. Hence, to identify γ the siblings must be of different ages at the time weight is observed.[136]

The differencing procedure can account for fixed unobserved genetic and/or environmental influences specific to the mother and household that is *constant within pairs* ($\alpha_{if} = \alpha_{jf}$). Irrespective, γ will still be biased if unobserved genetic and environmental factors correlated with mother's employment are not constant across siblings;[137] if the variables are measured with error; or if there is a simultaneous determination of weight and maternal employment.

For this analysis, the sibling fixed-effects is used exclusively to estimate γ. Conceptually, it is harder to justify variation in *permanent* income between siblings, and on a practical level, there is too little variation in differenced income to obtain precise estimates.

136 For example, the effects of maternal employment will be identified for a youth whose mother worked part-time during his or her own adolescence and worked full-time (or not at all) during the adolescence of a co-resident sibling.

137 For example, if one of the children suffers a chronic health condition the mother may make time or resource compensations for that child and hence the sibling fixed-effects model will be biased (Anderson *et al.* 2003, p.484).

8 Econometric Results

This chapter presents the empirical results, initially the effects of permanent RHED income and then the effects of maternal employment. For brevity, the combined overweight and obese category is herein referred to as 'overweight'.

8.1 Estimates of Permanent Income

Section 8.1.1 discusses the baseline OLS and probit results for the main sample. The estimation results are presented in Table 8.1 for BMI and Tables 8.2 and 8.3 for overweight and obesity, respectively. The estimates using proxy variables and IV methods are discussed in Section 8.1.2 and estimates for gender subgroups are presented in Section 8.1.3. Sensitivity tests on the main estimation results are presented in Section 8.1.4.

8.1.1 OLS and Probit Models

In Specification (1), a negative and significant relationship is found between permanent RHED income and youth BMI ($p=0.029$), overweight ($p=0.048$) and obesity ($p=0.004$), controlling only for basic child, parental and household characteristics. These income effects however are small in magnitude, even when considering relatively large changes in income. For instance, a $10,000 increase in permanent RHED income is associated with a 0.266kg/m^2 decrease in BMI, a 2.44 percentage point decrease in the probability of overweight and a 1.75 percentage point decrease in the probability of obese, for a youth whose permanent RHED income is equal to the (weighted) mean permanent RHED income in the sample (=$33,197).[138]

[138] The marginal effect of $10,000 permanent income was calculated by multiplying the marginal effect of log permanent income by the inverse of the (weighted) sample mean permanent RHED income (the level at which the marginal effect is evaluated), multiplied by $10,000. The marginal effects reported in Table 8.2-8.3 can be directly interpreted as the effects on the probability of overweight and obese respectively of a doubling of permanent RHED income, for a youth whose permanent RHED income equals the sample mean of log permanent RHED income.

Table 8.1. OLS and 2SLS Estimates of Permanent RHED Income
DV = BMI

	(1)		(2)		(3)		(4)	
	OLS	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*
ln(perminc)	-0.882**	-1.430	-0.665	-1.546	-0.302	-1.051	-0.089	0.005
std. error	(0.403)	(0.999)	(0.457)	(1.616)	(0.478)	(1.887)	(0.672)	(4.927)
female	-0.184	-0.183	-0.131	-0.123	-0.121	-0.116	-0.014	-0.014
age	0.376***	0.363***	0.404***	0.388***	0.404***	0.392***	0.316*	0.319
med_Yr 12			-0.297	-0.191	-0.243	-0.174	-0.232	-0.239
med_Cert			0.176	0.219	0.221	0.241	0.367	0.368
med_Uni			0.304	0.528	0.371	0.513	0.190	0.181
fed_Yr 12			-1.017	-0.909	-0.984	-0.931	-0.847	-0.846
fed_Cert			-0.494	-0.371	-0.467	-0.386	-0.321	-0.320
fed_Uni			-1.186**	-0.882	-1.123**	-0.928	-0.799	-0.808
SEIFA_disad					0.778*	0.645	0.818*	0.825
SEIFA_adv					-0.237	-0.100	-0.108	-0.117
BMI mum							0.130***	0.130***
BMI dad							0.165***	0.165***
Parent Employment							X	X
Proxy Variables							X	X
Constant	26.341***	32.237***	24.086***	33.309*	20.152***	28.009	10.578	9.462
Hansen J	—	2.339	—	1.301	—	1.347	—	0.375
Prob > χ^2 (Hansen J)	—	0.505	—	0.729	—	0.718	—	0.945
First-stage F	—	47.362	—	14.439	—	11.171	—	2.820
F test of exogeneity	—	0.427	—	0.338	—	0.173	—	0.000
Prob > F	—	0.514	—	0.561	—	0.678	—	0.985
Adjusted R^2	0.019	0.016	0.020	0.015	0.023	0.020	0.077	0.077
F	2.310	2.140	2.080	1.970	2.030	1.990	2.340	2.320
Prob > F	0.009	0.016	0.006	0.009	0.005	0.006	0.000	0.000
N	866	866	866	866	866	866	866	866

Notes: *, ** and *** means estimates are statistically significant at 10% and 5% and 1% levels, respectively. Standard errors are robust and clustered on the mother's identification code (in parentheses for permanent RHED income). Estimates are weighted. Income outliers are excluded. All specifications include dummies for mothers' country of birth, supplemented observations and for 'own' mothers; Specification (4) includes the parental employment dummy variables and proxy variables listed in Appendix D (Table 14.1) (with missing value dummies). Instruments include maternal and paternal grandfathers' occupational status score and the average unemployment rate by mothers' industry division and state. Four observations have a zero population weight.

Table 8.2. Marginal Effects of Permanent RHED Income
DV = Overweight

	(1)		(2)		(3)		(4)	
	probit	*IV-probit*	*probit*	*IV-probit*	*probit*	*IV-probit*	*probit*	*IV-probit*
ln(perminc)	-0.081**	-0.060	-0.073	-0.032	-0.056	0.032	-0.092	0.211
std. error	(0.041)	(0.095)	(0.046)	(0.156)	(0.048)	(0.190)	(0.073)	(0.533)
female	-0.049	-0.049	-0.043	-0.043	-0.042	-0.042	-0.031	-0.032
age	0.012	0.013	0.016	0.017	0.016	0.018	0.007	0.016
med_Yr 12			-0.036	-0.041	-0.031	-0.039	-0.034	-0.053
med_Cert			-0.026	-0.028	-0.025	-0.027	-0.030	-0.026
med_Uni			0.031	0.020	0.033	0.015	-0.002	-0.030
fed_Yr 12			-0.074	-0.079	-0.079	-0.084	-0.089	-0.085
fed_Cert			-0.048	-0.053	-0.046	-0.056	-0.044	-0.041
fed_Uni			-0.111**	-0.123*	-0.112**	-0.131**	-0.100*	-0.123*
SEIFA_disad					0.064	0.081	0.076*	0.096*
SEIFA_adv					0.014	-0.002	0.014	-0.015
BMI mum							0.008**	0.009**
BMI dad							0.014***	0.014***
Parent Employment							X	X
Proxy Variables							X	X
Hansen J	—	3.204	—	0.911	—	0.875	—	0.495
Prob > χ^2 (Hansen J)	—	0.361	—	0.823	—	0.832	—	0.920
First-stage F	—	47.362	—	14.439	—	11.171	—	2.820
Wald χ^2 test of exogeneity	—	0.065	—	0.073	—	0.233	—	0.329
Prob > χ^2	—	0.799	—	0.787	—	0.629	—	0.566
Pseudo R^2	0.016	—	0.027	—	0.029	—	0.076	—
Wald χ^2	14.978	11.260	28.253	24.419	31.005	28.614	80.546	80.120
Prob > χ^2	0.184	0.422	0.058	0.142	0.055	0.096	0.004	0.004
N	866	866	866	866	866	866	866	866

Notes: See note on Table 8.1. Marginal effects (at the mean) are calculated at the mean values of the continuous variables. For dummy variables the marginal effect of a change from 0 to 1 is calculated. Hansen J and first-stage F statistic calculated using an equivalent 2SLS-linear probability model. The dependent variable equals one if the youth is overweight *or* obese; and zero otherwise.

Table 8.3. Marginal Effects of Permanent RHED Income
DV = Obese

	(1)		(2)		(3)		(4)	
	probit	*IV-probit*	*probit*	*IV-probit*	*probit*	*IV-probit*	*probit*	*IV-probit*
ln(perminc)	-0.058**	-0.048	-0.052**	-0.016	-0.024	0.027	-0.047*	0.029
std. error	(0.020)	(0.052)	(0.022)	(0.095)	(0.023)	(0.113)	(0.030)	(0.239)
female	-0.002	-0.002	0.000	0.000	0.004	0.003	0.019*	0.019
age	0.001	0.001	0.002	0.003	0.002	0.003	-0.001	0.001
med_Yr 12			0.024	0.019	0.028	0.022	0.024	0.018
med_Cert			0.033	0.031	0.034	0.033	0.039*	0.041*
med_Uni			0.030	0.018	0.038	0.025	0.032	0.023
fed_Yr 12			-0.019	-0.022	-0.012	-0.015	-0.004	-0.003
fed_Cert			-0.015	-0.020	-0.010	-0.016	-0.002	-0.002
fed_Uni			-0.033	-0.042	-0.023	-0.034	-0.006	-0.012
SEIFA_disad					0.037	0.049*	0.037**	0.046*
SEIFA_adv					-0.039**	-0.047*	-0.027**	-0.034*
BMI mum							0.005***	0.005***
BMI dad							0.006***	0.006***
Parent Employment							X	X
Proxy Variables							X	X
Hansen J	—	0.812	—	3.669	—	3.878	—	0.689
Prob > χ^2 (Hansen J)	—	0.847	—	0.300	—	0.275	—	0.876
First-stage F	—	47.362	—	14.439	—	11.171	—	2.820
Wald χ^2 test of exogeneity	—	0.049	—	0.171	—	0.250	—	0.112
Prob > χ^2	—	0.825	—	0.679	—	0.617	—	0.738
Pseudo R^2	0.037	—	0.045	—	0.064	—	0.189	—
Wald χ^2	15.800	8.401	21.960	15.082	28.530	27.179	110.990	109.049
Prob > χ^2	0.149	0.677	0.234	0.656	0.098	0.130	0.000	0.000
N	866	866	866	866	866	866	866	866

Notes: See note on Table 8.1. Marginal effects (at the mean) are calculated at the mean values of the continuous variables. For dummy variables the marginal effect of a change from 0 to 1 is calculated. Hansen J and first-stage F statistic calculated using an equivalent 2SLS-linear probability model. The dependent variable equals one if the youth is obese; and zero otherwise.

Moreover, the effects of permanent income are reduced in size and significance as controls for parents' education and neighbourhood factors are introduced. Controlling for parents' education (Specification 2) reduces the effects of permanent RHED income by around 25 per cent for BMI and 10 per cent for overweight and obese, which remains statistically significant only for obesity (p=0.019).[139] This suggests that parents' education accounts for a large part of the observed relationship between family income and youth BMI/overweight in Australian youths.

Interestingly though, parents' education as a whole is not a significant determinant of youth weight outcomes. The dummy variables representing parents' education are jointly insignificant in the OLS (p=0.455) and probit models (p=0.329; p=0.895), based on the F-test and Wald χ^2 tests, respectively. This is in line with recent evidence for Australia which found only "weak" effects of parents' education on the health of children (Khanam *et al.* 2009, p.10). Individually, significant effects are predicted only for youths whose fathers are University educated, who are 1.186 kg/m^2 lighter (p=0.032) and 11.1 percentage points less likely to be overweight (p=0.038), than a youth whose father's highest qualification is Year 11 or less. However, with only imperfect measures of permanent income the unobservable component of family income could also be loading onto the paternal education coefficients.

Including controls for neighbourhood SES (Specification 3) reduces the coefficient on permanent RHED income in all models. However, as discussed in Section 5.4.6, the neighbourhood variables may control for "too much". That is, neighbourhood

[139] The specification of permanent RHED income in logarithmic form implies that the marginal effects of a given increase in income are larger at lower level of income. To examine whether the logarithmic functional form is appropriate, specifications were estimated that interacted permanent RHED income (in level form) with three dummy variables denoting whether the youth's permanent RHED income was in the first (highest), second and third quartile of the permanent RHED income distribution, respectively; with the fourth (lowest) income quartile as the omitted category. The results (not reported) indicated that a given increase in income does have larger effects for lower income households; however, the results were sensitive to the exclusion of outlying permanent RHED income observations.

may mediate the relationship between family income and youth weight if increased permanent income enables a family to "purchase" a higher SES neighbourhood with better amenities and fewer fast food outlets. If part of a change in permanent income operates through neighbourhood selection, the "full" or "policy-relevant" effect of income will not be measured in Specification 3.

The relative advantage and disadvantage in a youth's suburb does have an independent, albeit *marginally* significant, effect on weight, after controlling for household income.[140] As argued for fathers' education, this may also reflect the possibility that the neighbourhood variables are stronger correlates of a family's "true" permanent income than the 3-4 year measure of permanent RHED income used in this analysis. This hypothesis will be investigated in the following section, using IV.

8.1.2 Proxy and Instrumental Variables

Specification (4) introduces several covariates as proxies for the unobservable factors suspected to be correlated with permanent RHED income and youth weight (see Appendix D, Table 14.1).[141] In this specification, the permanent income parameter remains small and insignificant, suggesting little (if any) bias from unobserved heterogeneity.[142]

---

[140] Compared to a youth who lived in a 'middle' ranked suburb, a youth who lived in a 'disadvantaged' suburb is around 3.7-4.9 percentage points *more* likely to be obese; and is 2.7-4.7 percentage points *less* likely to be obese if they lived in an 'advantaged' suburb.

[141] Specification (4) also includes controls for parental employment, mothers' age and marital status.

[142] The explanatory power of the model is greatly increased in Specification (4), with the R^2 and Pseudo R^2 statistics increasing by around 3-4 fold. This is largely due to the introduction of controls for mothers' and fathers' BMI, which are jointly and independently significant at the 1 per cent level. Models were also estimated which interacted the mother's and father's BMI with separate dummy variables that equalled one if the respective parent was a step or foster parent, and zero otherwise (not reported). The interaction terms were insignificant, suggesting that parental BMI mostly captures the effects of the home environment, rather than the genetic transmission of excess weight. The income parameter was virtually unchanged if parental BMI was omitted from the model (remaining small and insignificant), suggesting it is not highly correlated with permanent RHED income.

Using IV (2SLS and IV-probit), a similar conclusion is reached.[143] The IV point estimates of permanent RHED income are broadly similar to the baseline estimates, albeit slightly larger for BMI and slightly smaller for binary overweight and obesity. However, the standard errors are around 2-4 times larger using IV and the instruments do not allow enough power to reject that the true coefficient is zero. It is not surprising then, that the null hypothesis of exogeneity in permanent RHED income cannot be rejected in any specification ($p>0.1$), suggesting that the baseline OLS and probit estimates (Table 8.1-8.3) are not driven by endogeneity.

Providing the instruments are valid, this also implies that the IV estimates could be discarded in favour of the more efficient OLS and probit estimates. As shown in Table 8.1-8.3, *all* instruments are suitably exogenous according to the Hansen J test of overidentifying restrictions[144]; and the instruments are not weak, at least in Specification (1)-(3).[145] However, a low first-stage F-statistic in Specification (4) flags a weak instrument problem, and hence, the IV estimates for permanent RHED income should be interpreted with caution in that specification.

Finally, the baseline and IV coefficients on parents' education and the SEIFA dummies are compared to investigate whether these coefficient are picking up the unobserved component of "true" permanent income not captured by permanent RHED income. Since IV (in theory) accounts for measurement error, any downward bias in the coefficients on fathers' education and 'SEIFA_disad' should

---

143 In IV (2SLS and IV-probit) specifications in Table 8.1-8.3, the only endogenous variable is permanent RHED income, which is instrumented using the occupational status score of the youth's maternal and paternal grandfather and the average unemployment rate by the mother's industry division and Australian state.

144 The Hansen J statistic reported in Table 8.1-8.3 has a χ_3^2 distribution, so the 5 per cent critical value is 7.81. For the IV-probit, the Hansen J statistic is calculated using an equivalent linear probability model, with a critical value equal to 7.81.

145 In Specification (1)-(3), the F-statistic for joint significance of the instruments in the first-stage regression exceed the rule of thumb ($F>10$) and the critical value for the 'maximal IV relative bias' test (=10.27). However, in Specification (3), the null hypothesis of weak instruments cannot be rejected using the critical values for the (conservative) 'maximal IV size' test (=24.58). See Appendix E (Section 15.1.1) for details of these tests.

be reduced after using IV. Similarly, any upwards in 'SEIFA_adv' would also be reduced.[146]

Comparison of the baseline and IV coefficients indicates that this is true for BMI (Table 8.1), though there is no clear pattern in the binary outcome models.[147] For this reason, and in view of the finding of the exogeneity test, it could be tentatively concluded that the measure of permanent RHED income used in this analysis provides a reasonable proxy for "true" permanent income.

8.1.3 Estimates by Gender

Table 8.4 presents the estimates of the OLS and probit models (Specification 4) by gender.[148] The results show that the effects of permanent RHED income differ by gender for overweight and obese; however, no differences exist for BMI.[149] For a girl with sample mean income (=$33,197), a $10,000 increase in permanent RHED income decreases the probability of overweight by 7.1 percentage points (p=0.016) and the probability of obese by 2.3 percentage points (p=0.068). In contrast, neutral effects are predicted for boys. Independent samples z tests show these differences are marginally significant for overweight (p=0.097) and obesity (p=0.073). This is in line with previous research which has uncovered gender differentials in the effects of household income on weight (Monheit *et al.* 2007; Gordon-Larsen *et al.* 2003; Baum and Ruhm 2007).

[146] This assumes that the unobserved component of income is *negatively* correlated with youth weight.

[147] It could be argued that parents' BMI and smoking and youth indigenous status are also capturing part of the variance in "true" income not accounted for by permanent RHED income (if any). Instrumenting for permanent RHED income, however, does not lead to a systematic decrease in these coefficients, relative to estimates assuming exogeneity of income.

[148] The instrumental variables estimates are excluded because the estimates become even more imprecise with smaller subgroups.

[149] The effects of permanent RHED income are insignificant for BMI for girls (p=0.528) and boys (p=0.670) and the difference in the estimates are also insignificant (p=0.883).

Table 8.4. Marginal Effects of Permanent RHED Income, By Gender

	(4)					
		Girls (*N*=433)			**Boys (*N*=433)**	
	OLS (BMI)	*probit (overweight)*	*probit (obese)*	*OLS (BMI)*	*probit (overweight)*	*probit (obese)*
ln(perminc)	-0.642	-0.235***	-0.075*	-0.430	0.012	-0.001
	(1.017)	(0.098)	(0.041)	(1.023)	(0.112)	(0.002)

Notes: See note on Table 8.1. The specification estimated is the same as specification (4) of Table 8.1, with the sample limited to the specified gender. In the probit specifications, marginal effects (at the mean) are calculated at the mean values of the continuous variables. For dummy variables the marginal effect of a change from 0 to 1 is calculated.

8.1.4 Robustness Checks

The sensitivity of the main results (Tables 8.1-8.3, Specification 4) to sample restrictions and the specification of permanent income is examined in Table 8.5. The main results are reproduced in row (A). Overall, the results are not sensitive to the sample restrictions, with the effects of permanent RHED income remaining small and statistically insignificant after including the 4 outlying permanent RHED income observations (B); dropping the 153 youths with less than four years of data during adolescence (C); or dropping the 382 observations with missing data on at least one variable, rather than setting missing values equal to zero and including dummy variables for missing data (D). Similarly, the effects of permanent RHED income remain small and insignificant if the estimation sample is drawn from an unbalanced panel, rather than a balanced panel, and estimates are unweighted (E). The main conclusions are also unchanged if imputed rent on owner-occupied housing is excluded from the calculation of permanent RHED income (F); or if household income is not equivalised but a separate control for the log of household size is included (G).[150] [151]

[150] The purpose of test (G) is to examine whether equivalised income imposes restrictions on the size of the separate effects of income and family size. The coefficients on log household size were insignificant in all specifications.

[151] Models were also estimated that included controls for the average annual percentage increase or decrease in annual RHED income over adolescence (i.e. the "slope" of family income). The coefficient on the "slope" of income was insignificant in all models (controlling for permanent RHED income) which could suggest that the timing of annual income is unimportant in determining weight outcomes. However, in light of the finding of nil effects of income, this result is not taken as evidence in support of the permanent income model of consumption and saving.

Table 8.5. Robustness Checks on the Main Results

Sample	N	OLS (BMI)	(4) probit (overweight)	probit (obese)
(A)	866	-0.089	-0.092	-0.047*
		(0.627)	(0.073)	(0.030)
(B)	870	0.205	-0.034	-0.020
		(0.646)	(0.071)	(0.033)
(C)	713	0.522	-0.021	-0.021
		(0.792)	(0.082)	(0.027)
(D)	484	0.049	-0.036	-0.043
		(0.926)	(0.097)	(0.029)
(E)	890	0.337	-0.008	-0.011
		(0.638)	(0.069)	(0.028)
(F)	866	-0.151	-0.045	-0.043
		(0.685)	(0.074)	(0.030)
(G)	866	-0.020	-0.039	-0.056**
		(0.694)	(0.076)	(0.030)

Notes: See note on Table 8.1. The specification estimated is the same as Specification (4) of Table 8.1. Marginal effects are calculated at mean values of explanatory variables.

8.1.5 Discussion

In sum, this analysis shows that the effects of permanent income on youth weight outcomes are small/zero and statistically insignificant after introducing controls for confounding covariates; albeit with modest negative effects for female youths. This could point to weakness in the ability of the underlying theory to explain the effects of income on youth BMI. As Section 2.3 outlines, the health of children is only one argument of the household's utility function (and obesity is only one dimension of health), and hence it may be restrictive to assume that positive shocks to family income unambiguously lead to purchases of goods and services used in the 'production' of lower BMI adolescents. For example, the utility the household derives from other "commodities", such as "entertainment", could see the additional income spent on goods that have no effect, or even a detrimental effect, on the youth's weight, such as video games, cinema tickets and motor vehicles.

8.2 Estimates of Maternal Employment

The estimates for maternal employment are presented in this section. First, the results for the main sample are discussed; initially for the baseline OLS/probit estimates and then for the specifications accounting for endogeneity. The main results are then interpreted in light of economic theory; followed by estimates for SES subgroups and sensitivity tests on the main results.

8.2.1 OLS and Probit Models

With controls only for basic child, parental and household characteristics (Specification 1), a youth whose mother worked part-time for one additional year is around 0.293kg/m^2 ($p=0.034$) lighter and 1.4 percentage points ($p=0.049$) less likely to be obese relative to a youth whose mother had not worked in that year.[152] Conversely, the coefficient on full-time employment is statistically insignificant, as is the coefficient on part-time employment with a combined overweight/obesity outcome. This suggests that there may be some protective effects of limited amounts of maternal labour supply for older children, albeit concentrated at the upper end of the BMI distribution.

[152] Multiplying these estimates by four shows that a youth whose mother worked part-time in all four years during adolescence is more than 5.6 percentage points less likely to be obese and 1.172kg/m^2 lighter.

Table 8.6. OLS, 2SLS and Sibling-Difference Estimates of Maternal Employment
DV = BMI

	(1)		(2)		(3)		(4)		
	OLS	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*	*OLS*	*2SLS*	*Sibling-Difference*
years PT	-0.293**	-0.152	-0.272**	-0.102	-0.286*	-0.327	-0.269	-0.399	-0.352
	(0.134)	(0.317)	(0.138)	(0.322)	(0.160)	(0.442)	(0.163)	(0.423)	(0.628)
years FT	-0.072	-0.471*	-0.071	-0.467	-0.132	-0.822	-0.176	-0.702	-0.824
	(0.141)	(0.263)	(0.147)	(0.289)	(0.180)	(0.587)	(0.181)	(0.564)	(0.866)
Parents' Education			X	X	X	X	X	X	—
Income + Other					X	X	X	X	—
Proxy Variables							X	X	—
First-stage F: PT	—	41.172	—	42.598	—	38.639	—	42.413	—
First-stage F: FT	—	82.781	—	60.167	—	30.048	—	28.767	—
F test of exogeneity	—	1.999	—	1.863	—	1.447	—	0.618	—
Prob > F	—	0.136	—	0.156	—	0.236	—	0.540	—
χ^2 test of exogeneity: PT	—	1.443	—	1.725	—	1.777	—	0.545	—
Prob > χ^2	—	0.230	—	0.189	—	0.183	—	0.460	—
χ^2 test of exogeneity: FT	—	3.409*	—	3.430*	—	2.876*	—	1.268	—
Prob > χ^2	—	0.065	—	0.064	—	0.090	—	0.260	—
Adjusted R^2	0.020	-0.011	0.023	-0.009	0.028	-0.014	0.079	0.062	0.018
F	2.040**	1.590*	1.961***	1.750**	1.668**	1.620**	2.370***	2.350***	1.200
Prob > F	0.011	0.072	0.006	0.019	0.011	0.015	0.000	0.000	0.289
N	866	866	866	866	866	866	866	866	231

Notes: *, ** and *** means estimates are statistically significant at 10% and 5% and 1% levels, respectively. Standard errors are robust and clustered on the mother's identification code (in parentheses). Estimates are weighted. Income outliers are excluded. Specifications (1)-(4) include dummies for supplemented observations, for 19-year olds and for 'own' mothers. Specifications (1)-(4) also include controls for number of years the mother is non-responding or living in a different household and for the basic child, parental and household characteristics listed in Table 14.1 (Appendix D); Specifications (2)-(4) include controls for parental education; Specifications (3)-(4) include controls for fathers' employment, mothers' age and marital status, permanent RHED income (excluding mothers' wage and salary income) and SEIFA dummies. Specification (4) includes the proxy variables listed in Table 14.1 (Appendix D). Instruments include the average unemployment rate by mothers' industry division and state and the average hours worked by females in the mother's occupation, industry and state. Hansen J statistic not reported as there are no overidentifying restrictions. Sibling-difference specification includes controls for age, gender, fathers' education and permanent RHED income (excluding mothers' wage and salary income).

Table 8.7. Marginal Effects of Maternal Employment
DV = Overweight

	(1)		(2)		(3)		(4)		
	probit	*IV-LPM*	*probit*	*IV-LPM*	*probit*	*IV-LPM*	*probit*	*IV-LPM*	*Sibling-Difference*
years PT	-0.014	-0.004	-0.013	0.006	-0.011	0.002	-0.015	-0.007	-0.002
	(0.012)	(0.031)	(0.013)	(0.031)	(0.016)	(0.044)	(0.016)	(0.043)	(0.061)
years FT	-0.004	-0.035	-0.004	-0.029	-0.005	-0.047	-0.015	-0.045	-0.106
	(0.013)	(0.024)	(0.013)	(0.027)	(0.017)	(0.055)	(0.018)	(0.057)	(0.081)
Parents' Education			X	X	X	X	X	X	—
Income + Other					X	X	X	X	—
Proxy Variables							X	X	—
First-stage F: PT	—	41.172	—	42.598	—	38.639	—	42.418	—
First-stage F: FT	—	82.781	—	60.167	—	30.048	—	28.767	—
F (test of exogeneity)	—	1.457	—	1.097	—	1.291	—	0.567	—
Prob > F	—	0.234	—	0.334	—	0.275	—	0.568	—
χ^2 (test of exogeneity: PT)	—	0.960	—	1.432	—	2.147	—	0.978	—
Prob > χ^2	—	0.327	—	0.231	—	0.143	—	0.323	—
χ^2 (test of exogeneity: FT)	—	2.512	—	2.143	—	2.225	—	1.035	—
Prob > χ^2	—	0.113	—	0.143	—	0.136	—	0.309	—
Pseudo R^2	0.013	—	0.025	—	0.044	—	0.084	—	0.054
Adjusted R^2	—	0.000	—	-0.015	—	-0.016	—	0.023	—
Wald χ^2	12.593	—	26.271	—	44.551	—	90.990***	—	13.340
Prob > χ^2	0.634	—	0.240	—	0.1064	—	0.000	—	0.272
F	—	0.830	—	1.190	—	1.380*	—	2.010***	—
Prob > F	—	0.647	—	0.253	—	0.075	—	0.000	—
Log likelihood	-491.309	—	-485.242	—	-475.961	—	-456.238	—	-144.422
N	866	866	866	866	866	866	866	866	231

Notes: see note on Table 8.6. First-stage F statistics calculated using an equivalent linear probability model. The dependent variable equals one if the youth is overweight *or* obese; and zero otherwise.

Table 8.8. Marginal Effects of Maternal Employment
DV = Obese

	(1)		(2)		(3)		(4)		
	probit	*IV-LPM*	*probit*	*IV-LPM*	*probit*	*IV-LPM*	*probit*	*IV-LPM*	*Sibling-Difference*
years PT	-0.014**	-0.005	-0.014*	-0.006	-0.016**	-0.019	-0.012**	-0.024	-0.037
	0.007	0.021	0.007	0.022	0.007	0.029	0.005	0.028	(0.027)
years FT	-0.007	-0.024	-0.007	-0.027	-0.015**	-0.045	-0.013**	-0.035	-0.084**
	0.007	0.016	0.007	0.018	0.008	0.037	0.006	0.033	(0.037)
Parents' Education			X	X	X	X	X	X	—
Income + Other					X	X	X	X	—
Proxy Variables							X	X	—
First-stage F: PT	—	41.172	—	42.598	—	38.639	—	42.418	—
First-stage F: FT	—	82.781	—	60.167	—	30.048	—	28.767	—
F (test of exogeneity)	—	1.061	—	1.099	—	0.775	—	0.215	—
Prob > F	—	0.347	—	0.334	—	0.461	—	0.807	—
χ^2 (test of exogeneity: PT)	—	1.132	—	0.980	—	0.946	—	0.031	—
Prob > χ^2	—	0.287	—	0.322	—	0.331	—	0.860	—
χ^2 (test of exogeneity: FT)	—	2.091	—	2.141	—	1.564	—	0.343	—
Prob > χ^2	—	0.148	—	0.143	—	0.211	—	0.558	—
Pseudo R^2	0.034	—	0.047	—	0.108	—	0.201	—	0.142
Adjusted R^2	—	-0.017	—	-0.019	—	-0.003	—	0.04	—
Wald χ^2	14.639	—	25.458	—	73.474***	—	114.463***	—	22.120**
Prob > χ^2	0.403	—	0.228	—	0.000	—	0.000	—	0.024
F	—	1.930**	—	1.510*	—	1.180	—	1.150	—
Prob > F	—	0.018	—	0.065	—	0.223	—	0.220	—
Log likelihood	-217.411	—	-214.389	—	-200.685	—	-179.747	—	-70.035
N	866	866	866	866	866	866	866	866	231

Notes: see note on Table 8.6. First-stage F statistics calculated using an equivalent linear probability model. The dependent variable equals one if the youth is obese; and zero otherwise.

It is possible, however, that the coefficient on part-time employment is picking up some of the beneficial effect of SES (particularly income) on weight. However, introducing controls for parents' education (Specification 2), permanent RHED income (*excluding* mothers' wage and salary income)[153], neighbourhood SES, fathers' employment and mothers' marital status (Specification 3) does not attenuate the part-time employment parameter, though its statistical significance drops slightly for youth BMI (p=0.075).[154] There is still the possibility, however, that this result is driven by remaining heterogeneity or reverse causality bias. This is explored in the following section.

8.2.2 Proxy and Instrumental Variables and Sibling Fixed-Effects

The first strategy to account for unobserved heterogeneity is to introduce the full set of proxy variables into the model (Specification 4).[155] Here, mothers' part-time and full-time employment remain individually and jointly (Prob>χ^2=0.032) significant for obesity; but are neither individually nor jointly significant (Prob>F=0.229) for BMI or for overweight (Prob>χ^2=0.693). In all cases however the point estimates change very little, suggesting little evidence of omitted variable bias.

Similarly, accounting for endogeneity using IV has no material effect on the coefficient estimates of part-time employment.[156] Nevertheless, the IV estimates for mothers' full-time employment are around 6-8 times *larger* relative to specifications assuming exogeneity. Though, in most cases the null hypothesis of

[153] The calculation of permanent RHED income excluding the mother's wage and salary income is described in Appendix B (Section 12.1.1).

[154] Controlling for SES and fathers' employment causes the coefficient on full-time employment to become significant for youth obesity at the 5 per cent level (p=0.045), with a one year increase in full-time employment reducing the probability of obese by around 0.015.

[155] These variables are listed in Table 14.1 (Appendix D).

[156] Part-time and full-time employment are instrumented using the unemployment rate by the mother's industry division and state and the average hours worked by females in the mother's occupation, industry and state. The IV specifications in Tables 8.6-8.8 do not instrument for permanent RHED income, as it was found to be exogenous in Section 8.1.2. Estimation is by 2SLS, which corresponds to an IV linear probability model in the case of Table 8.7 and Table 8.8. IV-probit was not used for the reasons discussed in Section 7.4.

exogeneity of part-time and full-time employment cannot be rejected at conventional levels of significance.[157]

The instruments themselves are highly correlated with mothers' part-time and full-time employment, evidenced by high F-statistics for joint significance of the instruments in the first-stage regressions, which are well in excess of the critical values proposed by Stock and Yogo (2005). The Hansen J test of overidentifying restrictions could not be employed to formally test whether the instruments are independent of the error process, given the equations are exactly identified.

The estimates of mothers' part-time and full-time employment using sibling-differences are similar to those using IV; part-time employment has similar effects to those estimated using OLS/probit and the effects of full-time employment are substantially larger when accounting for endogeneity.[158] This finding suggests that mothers select into full-time employment based on unobserved factors which are positively correlated with youth weight. Notwithstanding, both the IV and sibling fixed-effects estimates are imprecise, and hence, a caveat is placed on these results.

8.2.3 Income and Substitution Effects

This analysis highlights that the effects of mothers' employment on weight are either negative or statistically insignificant, after accounting for endogeneity. Intuitively, this would suggest that the "income effect" of employment dominates, or at least offsets, the "substitution effect" of maternal employment.[159]

---

[157] The exception is in Specifications (1)-(3) (in Table 8.6) for mothers' full-time employment, in which the exogeneity null is rejected at the 10 per cent level. However, the proxy variables appear to soak up much of the heterogeneity, as there is no evidence against the exogeneity of full-time employment in Specification (4).

[158] Standard errors are robust and clustered on the mother (multiple sibling-pairs per mother). Estimates are weighted using the average of the enumerated person longitudinal weights of siblings.

[159] That is, since Specifications (3) and (4) (Tables 8.6-8.8) measure permanent RHED income excluding mothers' income from wages and salary, the coefficient on mothers' part-time and full-time employment should pick up the net effect of decreased maternal time in the home (*a priori* a positive effect on weight) and increased earned income (*a priori* a negative effect on weight), as discussed in Section 3.3.

This explanation, however, is not supported by the data. As Tables 16.1-16.3 (Appendix F) demonstrates, re-estimating the models with the mother's wage and salary income included in permanent RHED income does not cause the coefficients on part-time and full-time employment to become positive. Full-time employment continues to have a negative (insignificant) effect, and part-time employment continues to have a negative and significant effect on obesity, at the 5 per cent level. This implies that employment leads to decreases in weight through channels other than household income.

The economic theory of health production does allow for employment to have *no* direct effect on health, holding income constant. That is, mothers may reduce time in leisure (L_{Lt}) as hours of employment (L_{Wt}) rise to "protect" the time spent producing children's health (L_{Qt}). The descriptive evidence (time use data) presented in Section 6.4 provides some support for this hypothesis. However, the theoretical explanation for a *negative* relationship between employment and weight is less clear.

One possibility is that there are weaknesses in the underlying theory, at least in terms of its predictions for weight outcomes in adolescents. The theory predicts that increased employment will lead to decreased health; but does not imply that *all* dimensions of health will decrease. Time spent producing health may also involve keeping children 'safe' by driving them to and from school or supervising sedentary study routines to improve their 'cognition'. This time may increase the child's overall stock of health, but may have *adverse* effects on BMI. Further research is required to explore these mechanisms – using detailed time use data.[160]

[160] Mothers may also allocate non-market time to producing other "commodities" which ultimately have adverse effects on children's BMI. Another possibility is that mothers acquire nutrition- and health-related information from their place of employment.

8.2.4 Estimates for SES Subgroups

Table 8.9 displays results for subsamples stratified by permanent RHED income, maternal education and neighbourhood advantage-disadvantage (SEIFA). All specifications control for the full-set of observable characteristics and proxy variables (equivalent to Specification 4 of Table 8.6-8.8).[161] Substantial favourable effects of part-time and full-time employment are predicted for advantaged youths, compared to neutral effects for the less advantaged. Particularly noteworthy are the highly significant reductions in BMI and overweight associated with employment of highly educated mothers (more than year 12 education) and employment of women from households with permanent RHED income greater than the (weighted) sample median income (=$34,462).[162]

Table 8.9. Marginal Effects of Maternal Employment for SES Subgroups

		years PT (4)		**years FT** (4)	
Group	*N*	*OLS (BMI)*	*probit (overweight)*	*OLS (BMI)*	*probit (overweight)*
		High SES			
> median income	432	-0.828***	-0.084***	-0.739***	-0.076***
		(0.257)	(0.024)	(0.265)	(0.023)
medcat > Yr 12	457	-0.730**	-0.044*	-0.756**	-0.061**
		(0.324)	(0.023)	(0.326)	(0.024)
High SES area	274	-0.389	-0.017	-0.590**	-0.009
		(0.258)	(0.021)	(0.282)	(0.023)
		Low SES			
<= median income	435	-0.023	0.022	0.117	0.000
		(0.212)	(0.021)	(0.307)	(0.030)
medcat <= Yr 12	409	-0.058	-0.011	0.239	0.011
		(0.172)	(0.020)	(0.244)	(0.023)
Low SES area	227	-0.274	-0.008	0.372	0.012
		(0.378)	(0.034)	(0.482)	(0.040)

Note: See note on Tables 8.6-8.8. The specification estimated is the same as model (4) of those tables, with the sample limited to the specified group. '> median income' refers to permanent RHED income greater than weighted sample median (=$34,462). 'medcat > Yr 12' refers to mothers with a trade Certificate or University Degree and 'medcat <= Yr 12' refers to youths whose mothers' highest education is Yr 12 or less. 'high SES area' and 'low SES area' means the youth lived in an 'advantaged' or 'disadvantaged' neighbourhood, respectively.

[161] Instrumental variables and sibling-difference estimates are not presented, since the estimates become even less precise with smaller subgroups. Subgroup analysis is not undertaken for binary obesity as there is too little variation in the dependent variable within the subgroups.
[162] The size and significance of the estimates were not materially affected by controlling for all sources of permanent RHED income (including mothers' wage and salary income). This suggests that the beneficial employment effects for advantaged youths do not work through income.

This finding could suggest that higher SES mothers are more focused on the education and safety dimensions of the “health” commodity (rather than excess weight) relative to low SES mothers, and hence, if these women are not employed, they spend time in activities that ultimately affect the child’s weight adversely. It may also indicate that high SES mothers are in the best position to provide highly quality market goods in their absence.

8.2.5 Robustness Checks

Table 8.10 presents the results of several sensitivity tests on the main results for maternal employment (Tables 8.6-8.8, Specification 4), which are reproduced in row (A). The results are not sensitive to the inclusion of the 4 outlying income and wealth observations (B); or to dropping the 152 observations with less than 4 years of income and employment data during adolescence, despite there being a slight increase in the size and significance of the effects of part-time employment (C). Furthermore, the results are not sensitive to taking the sample from an unbalanced panel, rather than a balanced panel, and using unweighted estimates (D). Finally, row (E) examines the sensitivity of the results to exclusion of the 382 individuals with missing data on one or more variables (rather than including dummy variables for missing data and setting missing observations to zero), finding that part-time and full-time employment are negatively and significantly related to the probability of overweight. In all cases (A)-(E), the direction of the point estimates are the same; providing some confidence in the main conclusions.

Table 8.10. Robustness Checks on the Main Results

		Years PT (4)			Years FT (4)		
Sample	*N*	*OLS (BMI)*	*probit (overweight)*	*probit (obese)*	*OLS (BMI)*	*probit (overweight)*	*probit (obese)*
(A)	866	-0.269	-0.015	-0.012**	-0.176	-0.015	-0.013**
		(0.163)	(0.016)	(0.005)	(0.181)	(0.018)	(0.006)
(B)	870	-0.257	-0.013	-0.012**	-0.167	-0.014	-0.013**
		(0.164)	(0.016)	(0.006)	(0.181)	(0.018)	(0.006)
(C)	714	-0.379**	-0.023	-0.013***	-0.260	-0.023	-0.010**
		(0.173)	(0.016)	(0.005)	(0.198)	(0.019)	(0.005)
(D)	890	-0.235	-0.020	-0.015***	-0.201	-0.031	-0.012*
		(0.158)	(0.015)	(0.006)	(0.182)	(0.020)	(0.007)
(E)	484	-0.260	-0.049**	-0.012**	-0.264	-0.048**	-0.009
		(0.191)	(0.020)	(0.006)	(0.212)	(0.022)	(0.006)

Notes: See note on Table 8.6. The specification estimated is the same as Specification (4) of Table 8.6-8.8. Marginal effects are calculated at mean values of explanatory variables.

The following chapter presents the major findings and conclusions for this thesis and suggested areas of further research.

9 Summary and Conclusions

9.1.1 Motivation and Objectives of Thesis

The rising economic and social costs of paediatric obesity warrant an investigation into its contributing factors. Emerging trends toward increased prevalence of obesity among children from low SES households (O'Dea 2009) and increased labour force participation of mothers with dependent children (Figure 1.1), suggest that household investments – both money and time – may be important. However, no study to date has investigated the *causal* effects of household income and maternal employment on youth overweight and obesity in Australia.

As such, and drawing on insights from household production theory, this thesis had two objectives: (i) to provide the first estimates, for Australia and internationally, on the causal effects of permanent income on youth weight outcomes and (ii) to provide the first evidence on the causal relationship between maternal employment and obesity for older children (aged over 12 years), in the Australian and international literatures.

To achieve this aim, data from the first seven waves of HILDA were exploited. Proxy variables, instrumental variables and sibling-differences were used to account for potential endogeneity in the two variables of interest.

9.1.2 Main Findings

The empirical results show that there are only modest differences between the weight outcomes experienced by youths in different income classes. Moreover, the analysis presented here finds no evidence that permanent income *causes* these differences: the income gradient in weight is reduced to zero as additional covariates are included in the model and after accounting for endogeneity. This suggests that income transfers are not a feasible approach to reducing the BMI of low income youths.

This result is broadly in line with the existing literature, which finds that, after accounting for confounding factors, family income does not have a large effect on child and adolescent weight outcomes (Anderson *et al.* 2007; Classen and Hoyakem 2005; Baum and Ruhm 2007) or child health and development more generally (Mayer 1997; Blau 1999; Khanam *et al.* 2009).

In line with several recent investigations (Baum and Ruhm 2007; Monheit *et al.* 2007), there is also some evidence of differential income effects by gender: with larger effects for females. However, this result was only found for binary overweight and obesity, and was not evident for the continuous measure of BMI.

The empirical analysis also implies that Australian youths have lower BMI and are less likely to be obese if their mother worked part-time or full-time during adolescence (relative to not-working). This negative effect remains after accounting for the potential endogeneity of mothers' employment; and was robust to sample restrictions. However, the labour supply effects were also imprecise in some models (particularly IV and sibling fixed-effects), and estimates based on larger sample sizes are required to provide confidence in these estimates.

Further analysis revealed this relationship to be driven by youths from high SES households. It is not fully understood why maternal job holding is particularly beneficial for high SES youths, though the analysis reveals that the effects do *not* operate through household income. A tentative conclusion is that time allocation decisions within the household are influenced by the mother's decision to work. For high-SES youths, these decisions inadvertently promote increased caloric expenditure and/or reduced caloric intake. However, the precise mechanisms remain unclear.

Overall, these results contrast with the Australian and international evidence for children and *young* adolescents, which generally uncovers a positive effect of maternal employment on children's weight outcomes (Anderson *et al.* 2003; Ruhm 2004; Zhu 2007). Moreover, the literature focussing on younger children generally finds larger adverse effects for high SES children (Anderson *et al.* 2003; Ruhm 2004), though not always (Chia 2008).

These results suggest some scope for policies to encourage high-SES mothers into work. However, in view of the large adverse effects of employment on the weight of young children (Zhu 2007), such policies may prove detrimental in aggregate.

9.1.3 Limitations and Areas for Further Research

One limitation of the empirical analysis is the relatively small sample size, which contributes to large estimated standard errors, particularly when using IV and sibling-differences. This imprecision is magnified by the inherent measurement error in the BMI data, meaning that, in many cases, there is insufficient statistical power to reject that the true coefficients are zero, particularly when controlling for multiple (correlated) dimensions of SES.

Future research could exploit the larger sample sizes and measured height and weight data available in the LSAC data set (particularly as more waves become available), to obtain more precise estimates of the income and labour supply effects. In combination with HILDA, the LSAC data, which collects information on young children, could also be used to explore whether the income gradient in obesity changes as children age; and whether the timing of maternal employment matters in Australia.

These types of analyses will also be feasible as more waves of HILDA become available. A longer panel will also allow permanent income to be calculated over a longer period, and enable a better approximation of children's lifetime exposure to maternal employment. Other econometric approaches may also be more feasible, including individual fixed-effects and 'long-difference' models.

Another avenue for future research is to provide a more comprehensive analysis of the effects of household income and maternal employment for subgroups in society. Notably, the newly available Longitudinal Survey of Indigenous Children

(LSIC) could be used to identify the effects of maternal employment and household income on obesity for Indigenous Australian children and the NHS 2007-08 could be used to analyse other ethnic subgroups. Research should also be directed at investigating the causal effects of permanent income and maternal employment on *older* adolescents in other countries.

As discussed previously, a priority for future research is to clarify the relationship between maternal employment and time allocation decisions in the household; and the implications for youth obesity. Several studies have found that an increase in employment is associated with a much smaller reduction in time spent with children (Bittman *et al.* 2004; Craig 2005; Bianchi 2000). However, more research could be done into the consequences of time allocation decisions on the adiposity of youths. This would require high quality parental time use data matched with youth BMI. In Australia, such data does not exist for older children; though LSAC comes close to meeting these requirements for younger children.

More fundamentally, there may be scope for further development of economic theory relating household production to youth obesity, given that the predictions of child health production models have not been substantiated here.

Lastly, the research and empirical methods used in this thesis could be extended and adapted to the analysis of the effects of household permanent income on the self-reported overall health of youths (using HILDA), to complement the research undertaken in this domain for younger children (Khanam *et al.* 2009).

10 References

Access Economics (2006). *The Economic Costs of Obesity*. Canberra.

Anderson, P. M., Butcher, K. F., & Levine, P. B. (2003). Maternal Employment and Overweight Children. *Journal of Health Economics.* 22(3), 477-504.

Anderson, P. M., Butcher, K. F., & Schanzenbach, D. W. (2007). Childhood Disadvantage and Obesity: Is Nurture Trumping Nature? Working Paper No. 13479, National Bureau of Economic Research.

Australasian Society for the Study of Obesity (ASSO) (2004). Available: www.asso.org.au.

Australian Bureau of Statistics (1998). *Standard Australian Classification of Countries (SACC)*. Canberra, ABS. Cat. No. 1269.0.

Australian Bureau of Statistics (2003). *Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA) 2001*. Canberra, ABS. Cat. No. 2039.0.

Australian Bureau of Statistics (2006). *2004-05 National Health Survey Users' Guide – Electronic Publication*. Canberra, ABS. Cat. No. 4364.0.

Australian Bureau of Statistics (2008a). *2008 Year Book Australia*. Canberra, ABS. Cat. No. 1301.0.

Australian Bureau of Statistics (2008b). *Overweight and Obesity in Adults, Australia 2004-05*. Canberra, ABS. Cat. No. 4719.0.

Australian Bureau of Statistics (2009a). *Household Income and Income Distribution Australia*. Canberra, ABS. Cat. No. 6523.0.

Australian Bureau of Statistics (2009b). *Consumer Price Index Australia, June 2008.* Canberra, ABS. Cat. No. 6401.0.

Australian Bureau of Statistics (2009c). *National Health Survey, Summary of Results, 2007-08*. Canberra, ABS. Cat. No. 4364.0.

Australian Bureau of Statistics (2009d). *Labour Force, Australia, Detailed, Quarterly, May 2009*. Canberra, ABS. Cat. No. 6291.0.55.003.

Australian Institute of Health and Welfare (AIHW) (2004). *A Rising Epidemic: Obesity in Australian Children and Adolescents*. Risk Factors Data Briefing No. 2.

Australian Taxation Office (ATO) (2008). *Taxation Statistics, Personal Tax, Table 5c*. Canberra.

Averett, S., & Korenman, S. (1996). The Economic Reality of the Beauty Myth. *Journal of Human Resources.* 31(2), 304-330.

Baum II, C. L., & Ford, W. F. (2004). The Wage Effects of Obesity: A Longitudinal Study. *Health Economics.* 13(9), 885-899.

Baum II, C. L., & Ruhm, C. J. (2007). Age, Socioeconomic Status and Obesity Growth. Working Paper No. 13289, National Bureau of Economic Research.

Baum, C. F. (2007). Instrumental Variables: Overview and Advances. Slide Show presentation to UKSUG 13, September, London, England.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal.* 75(299), 493-517.

Becker, G. S. (1981). *A Treatise on the Family*. Harvard University Press, Cambridge, Massachusetts.

Becker, G. S., & Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *The Journal of Political Economy.* 81(2), S279-S288.

Bianchi, S. M. (2000). Maternal Employment and Time with Children: Dramatic Change or Surprising Continuity? *Demography*, 37(4), 401-414.

Bittman, M., Craig, L., & Folbre, N. (2004). Packaging care: What Happens When Children Receive Nonparental Care? In N. Folbre & M. Bittman (Eds.), *Family Time: The Social Organization of Care*, pp.133–151, Routledge, London.

Blau, D. M. (1999). The Effect of Income on Child Development. *Review of Economics & Statistics.* 81(2), 261-276.

Blau, D. M., Guilkey, D. K., & Popkin, B. M. (1996). Infant Health and the Labor Supply of Mothers. *Journal of Human Resources.* 31(1), 90-139.

Blau, F. D., & Grossberg, A. J. (1992). Maternal Labor Supply and Children's Cognitive Development. *Review of Economics & Statistics.* 74(3), 474-481.

Booth, M. L., Wake, M., Armstrong, T., Tien, C., Hesketh, K., & Mathur, S. (2001). The Epidemiology of Overweight and Obesity among Australian Children and Adolescents, 1995-97. *Australian and New Zealand Journal of Public Health.* 25(2), 162-169.

Bound, J., Jaeger, D. A., & Raker, R. (1993). The Cure Can Be Worse than the Disease: A Cautionary Tale Regarding Instrumental Variables. Technical Paper No. 137, National Bureau of Economic Research.

Bradbury, B. (2003). *Child Poverty: A Review*. Report No. 3/03, Social Policy Research Centre, University of New South Wales.

Bradbury, B. (2007). *Child Outcomes and Family Socio-Economic Characteristics; Final Report of the Project: LSAC Outcomes and the Family Environment*. Report No. 10/07, Social Policy Research Centre, University of New South Wales.

Butterworth, P., & Crosier, T. (2004). The Validity of the SF-36 in an Australian National Household Survey: Demonstrating the Applicability of the Household Income and Labour Dynamics in Australia (HILDA) Survey to Examination of Health Inequalities. *BMC Public Health.* 4(44), 1-11.

Cameron, A. C. & Trivedi, P. K. (2009). *Microeconometrics Using STATA*. College Station, STATA Press.

Canning, H., & Mayer, J. (1966). Obesity: Its Possible Effect on College Acceptance. *New England Journal of Medicine.* 275, 1172-1174.

Case, A., Lee, D., & Paxson, C. (2008). The Income Gradient in Children's Health: A Comment on Currie, Shields and Wheatley Price. *Journal of Health Economics.* 27(3), 801-807.

Case, A., Lubotsky, D., & Paxson, C. (2002). Economic Status and Health in Childhood: The Origins of the Gradient. *American Economic Review.* 92(5), 1308-1334.

Cawley, J. (2004). The Impact of Obesity on Wages. *Journal of Human Resources.* 39(2), 451-474.

Cawley, J., & Liu, F. (2007). Maternal Employment and Childhood Obesity: A Search for Mechanisms in Time Use Data. Working Paper No. 07-29, National Poverty Center.

Cawley, J., Joyner, K., & Sobal, J. (2006). Size Matters: The Influence of Adolescents' Weight and Height on Dating and Sex. *Rationality and Society.* 18(1), 67-94.

Chase-Lansdale, P. L., Moffitt, R. A., Lohman, B. J., Cherlin, A. J., Coley, R. L., Pittman, L. D., *et al.* (2003). Mothers' Transitions from Welfare to Work and the Well-Being of Preschoolers and Adolescents. *Science.* 299(5612), 1548-1552.

Chia, Y. F. (2008). Maternal Labour Supply and Childhood Obesity in Canada: Evidence from the NlSCY. *Canadian Journal of Economics.* 41(1), 217-242.

Classen, T., & Hokayem, C. (2005). Childhood Influences on Youth Obesity. *Economics & Human Biology.* 3(2), 165-187.

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a Standard Definition for Child Overweight and Obesity Worldwide: International Survey. *British Medical Journal.* 320, 1240-1243.

Coleman, J. S. (1988). Social Capital and the Creation of Human Capital. *American Journal of Sociology.* 94(3), s95-s120.

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between Poverty and Psychopathology: A Natural Experiment. *Journal of the American Medical Association.* 290(15), 2023-2029.

Craig, L. (2005). How do they Find the Time? A Time-Diary Analysis of how Working Parents Preserve their Time with Children. Discussion Paper No. 136, Social Policy Research Centre, University of New South Wales.

Currie, A., Shields, M. A., & Price, S. W. (2007). The Child Health/Family Income Gradient: Evidence from England. *Journal of Health Economics.* 26(2), 213-232.

Currie, J. (2008). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. Working Paper No. 13987, National Bureau of Economic Research.

Currie, J., & Stabile, M. (2003). Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children? *The American Economic Review,* 93(5), 1813-1823.

Cutler, D. & Lleras-Muney, A. (2006). Education and Health: Evaluating Theories and Evidence. Working Paper, No.12352, National Bureau of Economic Research.

Dahl, G. B., & Lochner, L. (2005). The Impact of Family Income on Child Achievement. Working Paper No. 11279, National Bureau of Economic Research.

Daniels, S. R. (2006). The Consequences of Childhood Overweight and Obesity. *Future of Children.* 16(1), 47-67.

Datar, A., & Sturm, R. (2006). Childhood Overweight and Elementary School Outcomes. *International Journal of Obesity.* 30(9), 1449-1460.

Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA). (2009). *Data Dictionary.* Available online : http://www.aifs.gov.au/growingup/data/datadict/index.html

Doyle, O., Harmon, C., & Walker, I. (2007). The Impact of Parental Income and Education on Child Health: Further Evidence for England. Discussion Paper No. 06/07, UC Geary Institute.

Drewnowski, A., & Specter, S. E. (2004). Poverty and Obesity: The Role of Energy Density and Energy Costs. *American Journal of Clinical Nutrition.* 79(1), 6-16.

Duncan, G. J., Yeung, J. W., Brooks-Gunn, J., & Smith, J. R. (1998). How Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review.* 63(3), 406-423.

Edwards, L. N., & Grossman, M. (1978). Children's Health and the Family. Working Paper No. 256, National Bureau of Economic Research.

Fertig, A., Glomm, G., & Tchernis, R. (2006). The Connection between Maternal Employment and Childhood Obesity: Inspecting the Mechanisms. Working Paper No. 2006-020, Center for Applied Economics and Policy Research, Indiana University.

Finkelstein, E. A., Ruhm, C. J., & Kosa, K. M. (2005). Economic Causes and Consequences of Obesity. *Annual Review of Public Health.* 26, 239-257.

Friedman, M. (1957). *A Theory of the Consumption Function.* Princeton University Press, Princeton.

García, E., Labeaga, J. M., & Masagué Ana, C. O. (2006). Maternal Employment and Childhood Obesity in Spain. Working Paper No. 2006-17, Fundación de Estudios de Economía Aplicada.

Gordon-Larsen, P., Adair, L. S., & Popkin, B. M. (2003). The Relationship of Ethnicity, Socioeconomic Factors, and Overweight in U.S. Adolescents. *Obesity Research.* 11(1), 121-129.

Gortmaker, S. L., Must, A., Perrin, J. M., Sobol, A. M., & Dietz, W. H. (1993). Social and Economic Consequences of Overweight in Adolescence and Young Adulthood. *The New England Journal of Medicine.* 329(14), 1008-1012.

Gregg, P., Washbrook, E., Propper, C., & Burgess, S. (2005). The Effects of a Mother's Return to Work Decision on Child Development in the UK. *The Economic Journal.* 115, F48–F80.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *The Journal of Political Economy.* 80(2), 223-255.

Grossman, M. (2000). The Human Capital Model. In: Culyer, A. J. & J.P. Newhouse (Eds.), *Handbook of Health Economics*. North Holland, 347-408.

Gujarati, D. (2003). *Basic Econometrics*, 4th Edition. McGraw-Hill Higher Education, New York.

Guo, S. S., Wu, W., Chumlea, W. C., & Roche, A. F. (2002). Predicting Overweight and Obesity in Adulthood from Body Mass Index Values in Childhood and Adolescence. *The American Journal of Clinical Nutrition.* 76(3), 653-658.

Guryan, J., Hurst, E., & Kearney, M. S. (2008). Parental Education and Parental Time with Children. Working Paper No. 13993, National Bureau of Economic Research.

Haveman, R., & Wolfe, B. (1995). The Determinants of Children's Attainment: A Review of Methods and Findings. *Journal of Economic Literature.* 33(4), 1829-1878.

Hayes, C. (2008). HILDA Standard Errors: A Users Guide. HILDA Project Technical Paper Series No. 2/08. Institute of Applied Economic and Social Research, University of Melbourne.

Headey, B., & Warren, D. (2008). *Families, Incomes and Jobs, Volume 3: A Statistical Report on Waves 1 to 5 of the HILDA Survey*. Institute of Applied Economic and Social Research, University of Melbourne.

Heckman, J. (2007). The Technology and Neuroscience of Capacity Formation. *Proceedings of the National Academy of Sciences*. 104(33), 13250-13255.

Hofferth, S. L., & Curtin, S. (2005). Poverty, Food Programs, and Childhood Obesity. *Journal of Policy Analysis and Management.* 24(4), 703-726.

Huston, A., C., & Aronson, S. R. (2005). Mothers' Time with Infant and Time in Employment as Predictors of Mother-Child Relationships and Children's Early Development. *Child Development.* 76(2), 467-482.

Jacobson, L. (2000). The Family as Producer of Health - an Extended Grossman Model. *Journal of Health Economics.* 19(5), 611-637.

James-Burdumy, S. (2005). The Effect of Maternal Labor Force Participation on Child Development. *Journal of Labor Economics.* 23(1), 177-211.

Kaestner, R., & Grossman, M. (2008). Effects of Weight on Children's Educational Achievement. Working Paper No. 13764, National Bureau of Economic Research.

Kavanagh, A. M., Thornton, L. E., Tattam, A., Thomas, L., Jolley, D., & Turrell, G. (2007). *Place Does Matter for Your Health: A Report of the Victorian Lifestyle and Neighbourhood Environment Study.* University of Melbourne.

Khanam, R., Nghiem, H. S., & Connelly, L. B. (2009). Child Health and the Income Gradient: Evidence from Australia. *Journal of Health Economics.* 28(4), 805-817

Korenman, S., Miller, J. E., & Sjaastad, J. E. (1995). Long-Term Poverty and Child Development in the United States: Results from the NLSY. *Children and Youth Services Review.* 17(1-2), 127-155.

Li, M., Byth, K., & Eastman, C. J. (2007). Childhood Overweight and Obesity by Socio-Economic Indexes for Areas. *The Medical Journal of Australia.* 187(3), 195-200.

Lillard, D. R., Grabka, M., Freidin, S., Lipps, O., & Snider, K. (2009). Codebook for the Cross-National Equivalent File 1980-2007 BHPS - GSOEP - HILDA - PSID - SHP - SLID.

Liu, E. (2006). Maternal Full-Time Employment and Childhood Obesity. PhD Dissertation, University of Southern California.

Magarey, A., Daniels, L., & Boulton, T. (2001). Prevalence of overweight and obesity in Australian children and adolescents: reassessment of 1985 and 1995 data against new standard international definitions. *The Medical Journal of Australia.* 174:561-564.

Maurin, E. (2002). The Impact of Parental Income on Early Schooling Transitions: A Re-Examination Using Data over Three Generations. *Journal of Public Economics.* 85(3), 301-332.

Mayer, S.E. (1997). *What Money Can't Buy*. Harvard University Press, Cambridge, Massachusetts.

McMillan, J., Beavis, A., & Jones, F. L. (2009). The AUSEI06: A New Socioeconomic Index for Australia. *Journal of Sociology.* 45(2), 123-149.

Melbourne Institute of Applied Economic and Social Research (2009a). *Derived Variable Coding Framework*. University of Melbourne.

Melbourne Institute of Applied Economic and Social Research (2009b). *Continuing Person Questionnaire*. University of Melbourne.

Melbourne Institute of Applied Economic and Social Research (2009c). *Responding Person Coding Framework*. University of Melbourne.

Monheit, A. C., Vistnes, J. P., & Rogowski, J. A. (2007). Overweight in Adolescents: Implications for Health Expenditures. Working Paper No. 13488, National Bureau of Economic Research.

Morris, P. A., & Gennetian, L. A. (2003). Identifying the Effects of Income on Children's Development Using Experimental Data. *Journal of Marriage and Family.* 65(3), 716-729.

Mukhopadhyay, S. (2008). Do Women Value Marriage More? The Effect of Obesity on Cohabitation and Marriage in the USA. *Review of Economics of the Household.* 6(2), 111-126.

O'Dea, J. A. (2003). Differences in Overweight and Obesity among Australian Schoolchildren of Low and Middle/High Socioeconomic Status. *Medical Journal of Australia.* 179(1), 63-65.

O'Dea, J., & Dibley, M. (2009). Obesity Increase among Low SES Australian Schoolchildren between 2000 and 2006: Time for Preventive Interventions to Target Children from Low Income Schools? *International Journal of Public Health* (In Press).

Olshansky, S. J., Passaro, D. J., Hershow, R. C., Layden, J., Carnes, B. A., Brody, J., *et al.* (2005). A Potential Decline in Life Expectancy in the United States in the 21st Century. *The New England Journal of Medicine.* 352(11), 1138-1145.

Phipps, S. A., Lethbridge, L., & Burton, P. (2006). Long-Run Consequences of Parental Paid Work Hours for Child Overweight Status in Canada. *Social Science & Medicine.* 62(4), 977-986.

Propper, C., Rigg, J., & Burgess, S. (2007). Child Health: Evidence on the Roles of Family Income and Maternal Mental Health from a UK Birth Cohort. *Health Economics*. 16: 1245-1269.

Reserve Bank of Australia (RBA) (2009). *RBA F04 Retail Deposit and Investment Rates.* Available online: www.rba.gov.au/statistics/bulletin/index.html.

Reserve Bank of Australia (RBA) (2009). *RBA F05 Indicator Lending Rates.* Available online: www.rba.gov.au/statistics/bulletin/index.html.

Rodgers, J. R., & Rodgers, J. L. (1993). Chronic Poverty in the United States. *The Journal of Human Resources.* 28(1), 25-54.

Rodgers, J. R., & Rodgers, J. L. (2006).Chronic and Transitory Poverty in Australia 2001-2004. Economics Working Paper Series No. 06-28, University of Wollongong

Rodgers, J. R., & Rodgers, J. L. (2008). Contributions of Longitudinal Data to Poverty Measurement in Australia. Economics Working Paper Series, University of Wollongong.

Rodgers, J. R., & Rodgers, J. L. (2009). Chronic and Transitory Poverty over the Life Cycle. Paper presented to the 2009 HILDA Survey Research Conference (refereed stream), 16-17 July, University of Melbourne, Melbourne, Australia.

Rodgers, J., R , Siminski, P., & Bishop, J. (2009). Changes in Poverty Rates During the Howard Era. *The Australian Economic Review.* 42(3), 300-320.

Romer, D. (2006). *Advanced Macroeconomics,* 3rd Edition. McGraw-Hill Irwin, New York.

Ruhm, C. J. (2004). Maternal Employment and Adolescent Development. Working Paper No. 10691, National Bureau of Economic Research.

Sabia, J. J. (2007). The Effect of Body Weight on Adolescent Academic Performance. *Southern Economic Journal.* 73(4), 871(830).

Salmon, J., Timperlo, A., Cleland, V., & Venn, A. (2005). Trends in Children's Physical Activity and Weight Status in High and Low Socio-Economic Status Areas of Melbourne, Victoria, 1985-2001. *Australian & New Zealand Journal of Public Health.* 29(4), 337-342.

Saunders, P., & Siminski, P. (2005). Home Ownership and Inequality: Imputed Rent and Income Distribution in Australia. *Economic Papers.* 24(4), 346-367.

Scholder, S.V.H.K (2008). Maternal Employment and Overweight Children: Does Timing Matter? *Health Economics.* 17(8), 889-906.

Shea, J. (2000). Does Parents' Money Matter? *Journal of Public Economics.* 77(2), 155-184.

Siminski, P. (2006). The Effect of Earnings on Housework: Pros and Cons of HILDA's Time Use Data Items, Paper presented to the ACSPRI Social Science Methodology Conference, 10-13 December, University of Sydney, Sydney, Australia.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica.* 65(3), 557-586.

Stifter, C., A. , Coulehan, C., M. , & Fish, M. (1993). Linking Employment to Attachment: The Mediating Effects of Maternal Separation Anxiety and Interactive Behavior. *Child Development.* 64(5), 1451-1460.

Stock, J. H., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In: D. W. K. Andrews & J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80-108). Cambridge University Press, Cambridge.

Stock, J.H., & Watson, M.W. (2007). *Introduction to Econometrics,* 2nd Edition. Pearson Education, New York.

Takahashi, E., Yoshida, K., Sugimori, H., Miyakawa, M., Izuno, T., Yamagami, T., Kagamimori, S. (1999). Influence Factors on the Development of Obesity in

3-Year-Old Children Based on the Toyama Study. *Preventive Medicine.* 28(3), 293-296.

Taras, H., & Potts-Datema, W. (2005). Obesity and Student Performance at School. *The Journal of School Health.* 75(8), 291-295.

Taylor, B. A., Dearing, E., & McCartney, K. (2004). Incomes and Outcomes in Early Childhood. *Journal of Human Resources.* 39(4), 980-1007.

Verbeek, M. (2008). *A Guide to Modern Econometrics*. West Sussex: John Wiley & Sons Ltd.

Wake, M., Hardy, P., Canterford, L., Sawyer, M., & Carlin, J. B. (2007). Overweight, Obesity and Girth of Australian Preschoolers: Prevalence and Socio-Economic Correlates. *International Journal of Obesity.* 31(7), 1044.

Wang, Z., Patterson, C. M., & Hills, A. P. (2002a). Association between Overweight or Obesity and Household Income and Parental Body Mass Index in Australian Youth: Analysis of the Australian National Nutrition Survey, 1995. *Asia Pacific Journal of Clinical Nutrition.* 11(3), 200-205.

Wang, Z., Patterson, C. M., & Hills, A. P. (2002b). A Comparison of Self-Reported and Measured Height, Weight and BMI in Australian Adolescents. *Australian and New Zealand Journal of Public Health.* 26(5), 473-478.

Ware, J. J. E., & Gandek, B. (1998). Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *Journal of Clinical Epidemiology.* 51(11), 903-912.

Watson, N. (ed.) (2009). *HILDA User Manual – Release 7*. Institute of Applied Economic and Social Research, University of Melbourne.

Whitaker, R. C., Wright, J. A., Pepe, M. S., Seidel, K. D., & Dietz, W. H. (1997). Predicting Obesity in Young Adulthood from Childhood and Parental Obesity. *The New England Journal of Medicine.* 337(13), 869-873.

Wilkins, R. (2009). Updates and Revisions to Estimates of Income Tax and Government Benefits. HILDA Project Discussion Paper Series No. 1/09. Institute of Applied Economic and Social Research, University of Melbourne.

Wilkins, R., Warren, D., & Hahn, M. (2009). *Families, Incomes and Jobs, Volume 4: A Statistical Report on Waves 1 to 6 of the HILDA Survey*. Institute of Applied Economic and Social Research, University of Melbourne.

Wooden, M. (2009). Use of the Kessler Psychological Distress Scale in the HILDA Survey. HILDA Project Discussion Paper Series No. 2/09. Institute of Applied Economic and Social Research, University of Melbourne.

Wooden, M., & Watson, N. (2007). The HILDA Survey and Its Contribution to Economic and Social Research (So Far). *Economic Record.* 83(261), 208-231.

Wooden, M., Watson, N., Aguis, P., & Freidin, S. (2008). Assessing the Quality of the Height and Weight Data in the HILDA Survey. HILDA Project Technical Paper Series No. 1/08. Institute of Applied Economic and Social Research, University of Melbourne.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge.

World Health Organisation (WHO) (2000). *Obesity: Preventing and Managing the Global Epidemic: Report of a WHO Consultation.* Geneva.

Zhu, A. (2007). The Effect of Maternal Employment on the Likelihood of a Child Being Overweight. Discussion Paper No. 2007/17, School of Economics, University of New South Wales.

11 Appendix A

11.1 Calculating Permanent Income

Permanent income, for T years, can be calculated for any individual using the procedure outlined below. This procedure is taken from Rodgers and Rodgers (1993, p.37).

Given: savings interest rate in each year, $rs_t \quad \forall t = 1,2,...T$

borrowing interest rate in each year, $rb_t \quad \forall t = 1,2,...T$

and annual income, $y_t \quad \forall t = 1,2,...T$

<u>Step 1</u>: Compute average income as the first approximation of one's permanent income:

$$\bar{y} = \frac{\sum_{1}^{T} y_t}{T}$$

<u>Step 2</u>: Compute saving and borrowing in each year.

$$s_t = y_t - \bar{y} \qquad \forall t = 1,2,...T$$

where $s_t > 0$ for saving and
$s_t < 0$ for borrowing.

<u>Step 3</u>: Compute the balance at the end of the year if the individual consumes $\bar{y}$ in each year.

$$b_t = s_t + d*(1 + rs_{t-1})*b_{t-1} + (1-d)*(1 + rb_{t-1})*b_{t-1} \qquad \forall t = 1,2,...T$$

where $b_0 = 0$; $d = 1$ if $b_{t-1} > 0$; $d = 0$ otherwise

<u>Step 4</u>: Determine whether the end of year balance in year T (the end of period balance) is acceptably close to zero.

$$b_T \approx 0?$$

For the purposes of this thesis, an end of period balance of less than \$0.10 was taken to be acceptably close to zero.

Step 5: If b_T is not sufficiently close to zero, adjust the savings and borrowings in each year using the formula:

$$s_t = s_t - b_T / T$$

and repeat steps 3 and 4. This iterative procedure (Steps 3, 4 and 5) is repeated until the final end of period balance is acceptably close to zero.

Step 6: Compute permanent income

$$Y_t^* = y_t - s_t \qquad \forall t = 1, 2, \ldots T$$

11.2 Annual Interest Rates

Table 11.1. Indicator Nominal Interest Rates

	Indicator Savings Rates (%)		
Year	Cash management accounts at banks	Bank's term deposits ($10,000)	Average of indicator savings rates
2000-01 Annual Average	3.50	4.76	**4.13**
2001-02 Annual Average	2.39	3.65	**3.02**
2002-03 Annual Average	2.54	3.86	**3.20**
2003-04 Annual Average	2.81	4.14	**3.48**
2004-05 Annual Average	3.17	4.33	**3.75**
2005-06 Annual Average	3.26	4.27	**3.76**
2006-07 Annual Average	4.24	5.10	**4.67**
	Indicator Lending Rates (%)		
Year	Term loans (unsecured)	Credit cards (standard rate)	Average of indicator lending rates
2000-01 Annual Average	11.82	16.46	**14.14**
2001-02 Annual Average	11.29	15.75	**13.52**
2002-03 Annual Average	11.75	16.00	**13.87**
2003-04 Annual Average	11.67	16.31	**13.99**
2004-05 Annual Average	11.90	16.60	**14.25**
2005-06 Annual Average	12.13	16.84	**14.49**
2006-07 Annual Average	12.60	17.54	**15.07**

Notes: The interest rate on cash management accounts is an average of the rate earned on balances totalling $10,000 and $50,000. The interest rate on term deposits is an average of the rate applying to terms of 6 and 12 months. The interest rate on unsecured term loans is an average of prevailing fixed and variable interest rate.
Source: RBA (2009a), RBA (2009b)

12 Appendix B

12.1 Detailed Variable Construction

12.1.1 Permanent RHED Income Excluding Mothers' Wage and Salary

In Section 8.2, the estimated models included permanent RHED income measured net of the mother's wage and salary income. The procedure used to calculate this variable is described here.

This procedure replicates (as closely as possible) the tax model used in HILDA, as detailed in Wilkins (2009) and Watson (2009, pp.58-69).[163] The tax model is required because a separate variable for the taxes payable on the mother's wages and salary income is not available in the HILDA data set.

In each year, the tax liability corresponding to the mother's wage and salary income is calculated as the difference between (a) the income tax payable on the mother's total taxable income and (b) the income tax payable on the mother's taxable income minus the income she received from wages and salary in the relevant financial year. Both (a) and (b) are imputed following the steps outlined in Wilkins (2009), with (b) simply subtracting the mother's wage and salary income from her 'gross income subject to tax'. The calculated tax liability is then subtracted from the mother's gross income from wages and salary to arrive at the mother's after-tax wage and salary income, which is then subtracted from financial year household disposable income to provide a measure of annual RHED income excluding the mother's wage and salary income.[164]

To this measure is added imputed rent on owner-occupied housing; and is equivalised using the modified OECD equivalence scale and expressed in 2006-07 dollars using the CPI.[165] The algorithm described in Appendix A is then used to compute a measure of permanent RHED income (during adolescence) excluding the mother's wages and salary. The procedure to calculate (a) and (b) is described below.

First, to impute (a), a measure of *gross income subject to tax* is constructed by summing income from wages and salaries (including workers' compensation), business income, investment income, private pensions and taxable Australian public transfers (Wilkins

[163] The only income tax variable contained in the responding person file is financial year taxes payable on the mother's 'taxable income'. Since taxable income comprises income from wages and salary, investments, businesses, private pensions and public transfers, this variable will overstate the taxes attributable to the mother's wage and salary income, particularly for mothers who receive large amounts of taxable income from sources other than wages and salary.

[164] It is assumed that the mother is non-retired, and hence, total income tax will be calculated as income tax, plus the Medicare levy, less estimated offsets and other credits (Wilkins 2009, p.12).

[165] Negative values are set equal to zero, representing the (very few) cases where the mother's after-tax contributions from wages and salary had previously offset any after-tax losses incurred by the household from unincorporated business or investments.

2009, p.7).[166] Deductions for work-related and other expenses incurred in earning income are then subtracted to obtain a measure of *taxable income*. This follows the procedure outlined in Wilkins (2009, p.7), in which deductions are assumed equal a fixed percentage of gross income that depends on the level of the mother's income. To determine the applicable percentages, data from the Australian Taxation Office (2008, Table 5c) on average deductions as a proportion of income are used for each of 20 income ranges.[167]

To estimate the mother's *income tax payable*, the four standard marginal tax rates are applied to the mother's taxable income. The appropriate tax rates for each wave are taken from Wilkins (2009, p.9). The Medicare levy is calculated at the rate of 1.5 per cent of taxable income for all mothers in the sample, and is added to the mother's income tax payable.[168] Finally, any tax offsets applicable to the mother are subtracted from her income tax liability. These include the Low Income Tax Offset (LITO) and the Dependent Spouse Tax Offset (SPOUTO), which are calculated as described in Wilkins (2009, pp.11-12) using the appropriate thresholds and eligibility criteria.[169]

12.1.2 Grandfathers' AUSEI06 Score

An AUSEI06 score for the youth's maternal and/or paternal grandfather was not provided in HILDA if the grandfather was deceased or if occupation data were unavailable. This was the case for 25 maternal grandfathers and 169 paternal grandfathers in the sample. The author's treatment of missing values is described here.

First, if the grandfather was deceased when the parent was aged 14, the occupational status score was set to zero. If the grandfather was surviving at age 14, but the parent did not know, or did not provide information on the labour force status and/or occupation of their father, then information on the grandfather's highest education level was used to impute a status score. This procedure is based on the concept of "occupational potential" as outlined in McMillan *et al.* (2009, pp.131-132), in which the educational attainments of those not in paid employment (or in this case, not *known* to be in paid employment) can be viewed as "providing them with the potential for entering a range of occupations". An individual's occupational potential is calculated as the average occupational status of employed persons with an equivalent level of education. The relevant imputed AUSEI06 scores are taken from McMillan *et al.* (2009, p.132) based on data from the 2006 Census. These are reproduced in Table 12.1 below.

---

[166] Taxable public transfers are calculated by subtracting Family Tax Benefits A and B, Maternity Allowance, Maternity Payment, the Disability Support Pension and Rent Assistance from public transfer income. These items are not taxable (Wilkins 2009, p.7).

[167] For example, a mother earning between $6000 and $10,000 (gross) in 2006-07 is assumed to claim 9.6 per cent of her gross income as a tax deduction, whilst a mother earning between $70,000 and $80,000 a year claims 4.9 per cent of gross income as deductions.

[168] This simplifying assumption is necessary due to the complexities associated with imputing the Medicare levy. This limitation of this procedure is that it overstates the tax liability of low income mothers who would otherwise qualify for the Medicare levy exemption. However, most of the excess payment is expected to be offset by the low income tax offsets available to such persons.

[169] It is assumed that the mother's in the sample are ineligible for the Senior Australians Tax Offset (SATO), the Pensioner Tax Offset (PETO) and the Mature Age Workers' Offset (MATO).

Table 12.1. Imputed Occupational Potential Scores

Educational attainment	Imputed AUSEI06 score
Left school before completing Year 9	27.0
Completed Year 9	30.4
Completed Year 10	35.7
Completed Year 11	38.6
Completed Year 12	42.8
Post-school Trade Certificate	39.5
Post-school Diploma	56.5
Post-school Degree or Higher	71.6

The parent's reports of the grandfather's highest education and qualification level (at the date of interview) are used to define the grandfather's educational attainment.[170] In cases where information on both employment and education were missing, the occupational status score of the respective grandmother was used. For the paternal grandfather, data was incomplete in 147 cases, and hence the relevant status score was set to zero for these observations and a dummy equalling one was used to indicate missing data.

12.1.3 Unemployment Rate; by Industry Division and State

The unemployment rate (by mothers' industry division and Australian state) is constructed by matching external information on the unemployment rate by industry and state to information on the mother's industry of employment and state of residence in each of the four years over the child's adolescence. First, external data on the number of unemployed persons by industry division of last job (1 digit ANZSIC 2006) and by Australian State from the ABS (Cat. No. 6291.0.55.003, 2009d, Data Cube UQ2) is combined with information on the number of employed persons by industry division (1 digit ANZSIC 2006) and State from the ABS (Cat. No. 6291.0.55.003, 2009d, Data Cube E09), to a produce a highly disaggregated measure of the unemployment rate.[171] In each financial year, the unemployment rate for each industry-State combination is calculated by dividing the number of unemployed persons in each industry-State by the sum of the employed and unemployed persons in that industry-State. This procedure is repeated for financial years corresponding to Wave 2 through Wave 7, but due to data unavailability could not be done for financial year 2000-01.

For employed mothers, the industry is identified according to the 1 digit ANZSIC 2006 division of the mother's current main job. In cases where the mother is not in paid work, the ANZSIC 2006 division of her last job is used, and in cases where information on the

[170] Information was only collected on the highest education level of the child's grandparents at the date of interview, and hence, the assumption is that that the highest education level of the child's grandfather did not increase subsequent to the parent turning 14.

[171] The industries include; Agriculture, Forestry and Fishing; Mining; Manufacturing; Electricity, Gas, Water and Waste Services; Construction; Wholesale Trade, Retail Trade; Accommodation and Food Services; Transport, Postal and Warehousing; Information, Media and Telecommunications; Financial and Insurance Services; Rental, Hiring and Real Estate Services; Professional, Scientific and Technical Services; Administrative and Support Services; Public Administration and Safety; Education and Training; Health Care and Social Assistance; Arts and Recreation Services; Other Services.

current main job or last job is unavailable, the overall unemployment rate of the state is used.

The industry code and State identifier (in HILDA) are used to match each mother to the relevant unemployment rate in a given industry division-state.[172] To remain consistent with the maternal employment measures used in this thesis, a measure of the mother's average exposure to labour market conditions is then constructed by averaging the relevant industry-State unemployment rate over the child's adolescent years (ages 12-15 inclusive), for those years in which the mother was present in the household.

12.1.4 Average Hours Worked by Employed Females

Data on the average hours worked is the average hours worked per week by employed females by industry division (1 digit ANZSIC 2006), major occupation (1 digit ANZSCO 2006) and Australian State/Territory.[173] This data is sourced from the ABS (Cat. No. 6291.0.55.003, 2009d, Data Cube E09), and is matched to the mother's industry and occupation codes and state identifier using the same procedure as that used to match unemployment rate information to the mother. If the mother is employed, the industry division (1 digit ANZSIC 2006) and major occupation (1 digit ANZSCO 2006) of the mother's current main job is used; otherwise the industry and occupation of her last job are used.[174] This procedure yields a measure of the average hours worked by females in the same industry, occupation and State or Territory of the mother for each financial year corresponding to Waves 1 to 7, which is subsequently averaged over the four years that constitute the child's adolescence.

[172] In the instances where the number of employed and unemployed persons in a given industry-State was too small to produce an estimate of the unemployment rate, the overall State unemployment rate is used. This was the case for 20 persons.

[173] The measure of average hours is calculated by dividing the total hours worked by employed persons (within each industry, occupation and state) by the total number of females employed (within that industry, occupation and state).

[174] If information is unavailable on the last job of the mother, or if the mother has never engaged in paid employment, the average hours worked by females in the relevant State/Territory is applied.

13 Appendix C

13.1 Jackknife Standard Errors

Standard errors of the estimates reported in Chapter 6 are calculated using the Jackknife methodology described by the Melbourne Institute (Hayes 2008, pp.3-4). The process involves calculating each estimate a number of times using the sets of replicate weights provided in the HILDA data and measuring the variability of these 45 estimates around the estimate calculated using the 'main' weight. That is

$$SE(\hat{x}) = \sqrt{\frac{R-1}{R}\sum_{j=1}^{R}(\hat{x}_j - \hat{x})^2} \tag{13.1}$$

where $\hat{x}$ is the estimate based on the 'full' sample computed using the 'main' weight and $\hat{x}_j$ is the estimate calculated from the sub-sample that is obtained when the R^{th} set of replicate weights are used (Rodgers *et al.* 2009, p.304). The HILDA dataset provides 45 replicate weights $(R = 45)$ for each type of weight on the dataset. Therefore, using the main weight and the set of 45 replicate enumerated person longitudinal weights, the standard errors are calculated as:

$$SE(\hat{x}) = \sqrt{\frac{44}{45}\sum_{j=1}^{45}(\hat{x}_j - \hat{x})^2} \tag{13.2}$$

To test whether $\hat{x}$ is significantly different to a second sample estimate, $\hat{y}$, the $\hat{x}$ in Formula 4.2 is replaced with the estimate of the *difference* between the two sample statistics, $\hat{d} = \hat{x} - \hat{y}$;

$$SE(\hat{d}) = \sqrt{\frac{44}{45}\sum_{j=1}^{45}(\hat{d}_j - \hat{d})^2} \tag{13.3}$$

The relevant test statistic calculated as;

$$z = \frac{\hat{d}}{SE(\hat{d})} \tag{13.4}$$

which is approximately normally distributed for large samples (follows z distribution).

The estimates in Section 6.1 are also compared to the published estimates from other surveys. As the difference between two survey estimates is subject to sampling error, a standard error of the difference between two independent estimates can be calculated using:

$$SE(\hat{x}-\hat{z}) = \sqrt{\left[SE(\hat{x})\right]^2 + \left[SE(\hat{z})\right]^2} \quad (13.5)$$

where $\hat{z}$ is the estimate from the 'benchmark' survey and $SE(\hat{z})$ is the standard error of the estimate. The HILDA sample and the benchmark sample are independent samples, and hence, the covariance between $\hat{x}$ and $\hat{z}$ equals zero. The difference between the estimates is significantly different to zero if the standard normal statistic;

$$z = \frac{\hat{x}-\hat{z}}{SE(\hat{x}-\hat{z})} \quad (13.6)$$

exceeds the critical value at conventional levels of significance.

14 Appendix D

14.1 Variable Definitions

Table 14.1. Variables Used in Analysis

Variable	Definition
	Key Variables
ln(perminc)	Log of permanent RHED income
ln(perminc_m)	Log of permanent RHED income less mother's wage and salary income
years PT	No. years mother employed part-time during adolescence
years FT	No. years mother employed full-time during adolescence
	Basic Child, Parental and Household Characteristics
female	=1 if female
age	Age (in years)
indigenous	=1 if Aboriginal or Torres Straight Islander
mcob_Oceanic	=1 if mother is other Oceanic born
mcob_NW Euro	=1 if mother is North or West Europe born
mcob_SE Euro	=1 if mother is South or East Europe born
mcob_Asia	=1 if mother is Asia born
mcob_Other	=1 if mother is 'other' born
supp	=1 if supplemented observation (BMI measured in Wave 6)
ownmum	=1 if mother is child's own mother
	Parental Education and SEIFA
med_Yr 12	=1 if Yr 12 is mother's highest education completed
med_Cert	=1 if a trade certificate is mother's highest education completed
med_Uni	=1 if University is mother's highest education completed
fed_Yr 12	=1 if Yr 12 is father's highest education completed
fed_Cert	=1 if a trade certificate is father's highest education completed
fed_Uni	=1 if University is father's highest education completed
SEIFA_disad	=1 if lowest three deciles of SEIFA (IRSAD)
SEIFA_adv	=1 if highest three deciles of SEIFA (IRSAD)
	Supplementary Variables
age mother	Age of mother (in years)
age mother ^ 2	Age of mother (in years) squared
years married	Number of years mother legally married during adolescence
years PT (dad)	No. years father employed part-time during adolescence
years FT (dad)	No. years father employed full-time during adolescence
years not in HH (dad)	No. years father not present in household during adolescence
years nonresp (dad)	No. years father non-responding but present in household
	Proxy Variables
BMI mother	Mother's BMI
BMI father	Father's BMI
transfers (units)	Average transfer income : gross income ratio
mental health	Mother's SF-36 mental health score (0-100)
mum smoke	=1 if mother smokes
dad smoke	=1 if father smokes
prebirth	=1 if mother worked in 12 months prior to birth of the child
mmemploy	=1 if mother's mother was in paid employment when mother was aged 14
bothparents	=1 if mother's own mother and own father present when mother was 14
math	Mother's self-rated mathematical skills relative to an 'average' adult (0=very poor, 10=very good)
reading	Mother's self-rated reading skills relative to an 'average' adult (0=very poor, 10=very good)
GST	Mother not comfortable working out amounts like the GST (1=strongly agree, 10=strongly disagree)
numbers	Mother reports to be good with numbers and calculations (1=strongly disagree, 10=strongly agree)

15 Appendix E

15.1 Instrumental Variables Tests

15.1.1 Testing for Weak Instruments

The traditional rule of thumb, as suggested by Staiger and Stock (1997), is that the instruments are weak if the F statistic for joint significance of the instruments in the first-stage regression is less than 10. As this rule of thumb is not sufficiently conservative if there are many overidentifying restrictions, the first-stage F statistic can also be compared against the critical values proposed by Stock and Yogo (2005), which are functions of the number of endogenous regressors (m) and the number of instruments (k_2). The critical values relevant to the specifications estimated in this analysis are presented in Table 15.1.

Table 15.1. Stock and Yogo (2005) Critical Values for Weak Instruments

	Permanent Income (Table 8.1-8.3) $[m=1,\ k_2=4]$	*Maternal Employment* (Table 8.6-8.8) $[m=2,\ k_2=2]$
10% maximal IV relative bias	10.27	—
10% maximal IV size	24.58	7.03

Note: Critical values are for the F statistic associated with the hypothesis that the coefficients on instruments in the first-stage regression are jointly equal to zero. Critical values for the ‘maximal IV relative bias’ test not available for maternal employment specifications since the test requires at least two overidentifying restrictions.

The first test assumes that 10 per cent is the desired maximal bias of the 2SLS estimator relative to the OLS; whilst the second test assumes that 10 per cent is the largest acceptable size for a 5 per cent Wald test of the joint significance of the endogenous regressors in the structural equation (Equation 7.6).[175] In both tests,

[175] The first test address the concern that the estimation bias of 2SLS resulting from the use of weak instruments can be large, and possibly even exceed the bias of OLS (Cameron and Trivedi 2009, p.190). The second test addresses the concern that weak instruments can result in size distortions of Wald tests on the parameters in small samples.

the null hypothesis is that the instruments are weak.[176] Using IV-probit, the first-stage F statistic is from an equivalent 2SLS-linear probability model.

15.1.2 Testing for Instrument Exogeneity

In an *overidentified* ($k_2 > m$) equation, a test of overidentifying restrictions can be used to test whether the instruments are suitably independent of the error process.

The test statistic, known as the J-statistic, is calculated as $J = mF$ where F is the homoskedasticity-only F-statistic testing the hypothesis that $\boldsymbol{\xi}_1 = \boldsymbol{\xi}_2 = 0$ in;

$$\hat{\varepsilon}_i^{2SLS} = \mathbf{x}_i \boldsymbol{\xi}_1 + \mathbf{z}_i \boldsymbol{\xi}_2 + e_i \tag{15.1}$$

where $\hat{\varepsilon}_i^{2SLS}$ is the estimated 2SLS residuals and e_i is a disturbance. Under the null hypothesis that *all* instruments are exogenous, J is distributed as $\chi^2_{k_2-m}$ in large samples, where $k_2 - m$ is the "degree of overidentification" (Stock and Watson 2007, p.444).[177]

The heteroskedasticity-robust variant of the J statistic (known as Hansen's J) is computed using Generalised Method of Moments (GMM) and also distributed as $\chi^2_{k_2-m}$ (Stock and Watson 2007, p.735).[178] Using IV-probit, the Hansen's J statistic is calculated using an equivalent 2SLS-linear probability model.

15.1.3 Testing for Regressor Endogeneity; 2SLS

Consider the structural model (Equation 7.4) with additional variable(s) $\hat{\mathbf{v}}_i$, representing the vector of fitted residual(s) from the first-stage regression(s) (Equation 7.5);

---

[176] When there is more than one endogenous regressor (Tables 8.6-8.8), the test statistic is instead the minimum eigenvalue of a matrix analogue of F statistic that is defined in Stock and Yogo (2005, p.84).

[177] The idea of the J-statistic is that, if the overidentifying restrictions hold, u_i will be uncorrelated with the instruments. This is commonly known as the Sargan test of overidentifying restrictions.

[178] This statistic is preferred given it allows observations to be correlated within groups when cluster-robust standard errors are estimated (STATA 2009)

$$y_{1i} = \alpha + \mathbf{y}_{2i}\boldsymbol{\beta}_1 + \mathbf{x}_i\boldsymbol{\beta}_2 + \boldsymbol{\rho}\hat{\mathbf{v}}_i + u_i \tag{15.2}$$

The test of exogeneity of $\mathbf{y}_{2i}$ using 2SLS is a test of $\mathrm{H}_0{:}\boldsymbol{\rho} = 0$, which has a null hypothesis of exogeneity.[179] In the case of multiple endogenous variables, $m > 1$, the correlation of each variable with the error on the structural equation ($\mathrm{H}_0{:}\ \rho = 0$) can be separately tested (Cameron and Trivedi 2009, p.184).[180]

15.1.4 Testing for Regressor Endogeneity; IV-Probit

Using IV-probit, a test of the null hypothesis of exogeneity of y_2 is equivalent to a test of $\mathrm{H}_0 : \rho = 0$ using an asymptotic t test. This is because $\rho = 0$ if u_1 and u_2 are uncorrelated (Cameron and Trivedi 2009, p.467; Wooldridge 2002, p.476).

[179] This test is also known as the Durbin-Wu-Hausman (DWH) test.

[180] For independent homoscedastic errors, this test is asymptotically equivalent to a Hausman test comparing the IV and OLS estimates of the potentially endogenous regressors (Wooldridge 2002, p.118). However, provided robust variance estimates are used, the test is also robust to heteroscedasticity.

16 Appendix F

16.1 Estimates of Maternal Employment Controlling for All Sources of Income

This Appendix presents the estimates of Specifications (3) and (4) (and sibling-differences) in Tables 8.6-8.8 with controls for permanent RHED income *including* mothers' wage and salary income. The results of Specifications (1) and (2), which do not control for permanent RHED income, are identical to those presented in Tables 8.6-8.8.

Table 16.1. Estimates of Maternal Employment, Controlling for All Sources of RHED Income
DV = BMI

	(3)		(4)		
	OLS	*2SLS*	*OLS*	*2SLS*	*Sibling-Difference*
years PT	-0.263*	-0.167	-0.250	-0.273	-0.200
	(0.157)	(0.403)	(0.162)	(0.386)	(0.611)
years FT	-0.094	-0.663	-0.141	-0.574	-0.799
	(0.170)	(0.509)	(0.175)	(0.490)	(0.860)
Parents' Education	X	X	X	X	—
Income + Other	X	X	X	X	—
Proxy Variables			X	X	—
First-stage F: PT	—	37.588	—	41.917	—
First-stage F: FT	—	31.929	—	30.212	—
F (test of exogeneity)	—	1.450	—	0.613	—
Prob > F	—	0.235	—	0.542	—
χ^2 (test of exogeneity: PT)	—	2.087	—	0.719	—
Prob > χ^2	—	0.149	—	0.397	—
χ^2 (test of exogeneity: FT)	—	2.864*	—	1.275	—
Prob > χ^2	—	0.091	—	0.259	—
Adjusted R^2	0.027	-0.015	0.079	0.062	0.018
F	1.630**	1.580**	2.310***	2.330***	1.200
Prob > F	0.014	0.020	0.000	0.000	0.289
N	866	866	866	866	231

Notes: see note on Table 8.6. Specification (3) and (4) are the same as Specification (3) and (4) of Table 8.6, except that permanent RHED income includes mothers' wages and salary income. The sibling-difference specification is also the same as Table 8.6, with mothers' wage and salary income included in permanent RHED income.

Table 16.2. Marginal Effects of Maternal Employment, Controlling for All Sources of RHED Income
DV = Overweight

	(3)		(4)		
	probit	*IV-LPM*	*probit*	*IV-LPM*	*Sibling-Difference*
years PT	-0.009	0.012	-0.014	-0.007	-0.001
	(0.015)	(0.040)	(0.015)	(0.043)	(0.059)
years FT	-0.002	-0.037	-0.013	-0.045	-0.101
	(0.016)	(0.048)	(0.017)	(0.057)	(0.081)
Parents' Education	X	X	X	X	—
Income + Other	X	X	X	X	—
Proxy Variables			X	X	—
First-stage F: PT	—	37.588	—	41.110	—
First-stage F: FT	—	31.929	—	30.175	—
F (test of exogeneity)	—	1.292	—	0.568	—
Prob > F	—	0.275	—	0.576	—
χ^2 (test of exogeneity: PT)	—	2.310	—	1.058	—
Prob > χ^2	—	0.129	—	0.304	—
χ^2 (test of exogeneity: FT)	—	2.194	—	1.023	—
Prob > χ^2	—	0.139	—	0.312	—
Pseudo R^2	0.044	—	0.085	—	0.053
Adjusted R^2	—	-0.017	—	0.021	—
Wald χ^2	44.230	—	91.090***	—	13.340
Prob > χ^2	0.113	—	0.000	—	0.272
F	—	1.370*	—	2.010***	—
Prob > F	—	0.082	—	0.000	—
Log likelihood	-475.954	—	-455.767	—	-143.322
N	866	866	866	866	231

Notes: see note on Table 16.1. The dependent variable equals one if the youth is overweight *or* obese; and zero otherwise.

Table 16.3. Marginal Effects of Maternal Employment, Controlling for All Sources of RHED Income
DV = Obese

	(3)		(4)		
	probit	*IV-LPM*	*probit*	*IV-LPM*	*Sibling-Difference*
years PT	-0.014**	-0.011	-0.011**	-0.018	-0.035
	0.007	0.026	0.005	0.026	(0.026)
years FT	-0.010	-0.035	-0.010*	-0.028	-0.085**
	0.007	0.032	0.006	0.029	(0.037)
Parents' Education	X	X	X	X	—
Income + Other	X	X	X	X	—
Proxy Variables			X	X	—
First-stage F: PT	—	37.588	—	41.917	—
First-stage F: FT	—	31.929	—	30.212	—
F (test of exogeneity)	—	0.777	—	0.195	—
Prob > F	—	0.460	—	0.823	—
χ^2 (test of exogeneity: PT)	—	1.164	—	0.073	—
Prob > χ^2	—	0.281	—	0.787	—
χ^2 (test of exogeneity: FT)	—	1.537	—	0.343	—
Prob > χ^2	—	0.215	—	0.558	—
Pseudo R^2	0.106	—	0.200	—	0.142
Adjusted R^2	—	-0.003	—	0.042	
Wald χ^2	71.430***	—	115.150***	—	22.120
Prob > χ^2	0.000	—	0.000	—	0.024
F	—	1.160	—	1.150	—
Prob > F	—	0.245	—	0.222	—
Log likelihood	-201.239	—	-180.128	—	-70.766
N	866	866	866	866	231

Notes: see note on Table 16.1. The dependent variable equals one if the youth is obese; and zero otherwise.