Is History Destiny? Resources, Transitions and Child Education Attainments in Canada

Final Report

John Hoddinott, Lynn Lethbridge and Shelley Phipps
Dalhousie University
Strategic Policy
Human Resources Development Canada

December 2002

SP-551-12-02E
(également disponible en français)

The Working Paper Series includes analytical studies and research conducted under the auspices of the Applied Research Branch of Strategic Policy. Papers published in this series incorporate primary research with an empirical or original conceptual orientation, generally forming part of a broader or longer-term program of research in progress. Readers of the series are encouraged to contact the authors with comments and suggestions.

■

This paper is available in French under the title
''Notre avenir est-il dicté par nos antécédents? Ressources, transitions et rendement scolaire des enfants au Canada''
La version française du présent document est disponible sous le titre
« Notre avenir est-il dicté par nos antécédents? Ressources, transitions et rendement scolaire des enfants au Canada »

■

Paper
ISBN: 0-662-33442-6
Cat. No.: RH63-1/551-12-02E

Internet
ISBN: 0-662-33443-4
Cat. No.: RH63-1/551-12-02E-IN

■

General enquiries regarding the documents published by the Applied Research Branch should be addressed to:

Human Resources Development Canada
Publications Centre
140 Promenade du Portage, Phase IV, Level 0
Hull, Quebec, Canada
K1A 0J9

Facsimile: (819) 953-7260
http://www.hrdc-drhc.gc.ca/sp-ps/arb-dgra

Si vous avez des questions concernant les documents publiés par la Direction générale de la recherche appliquée, veuillez communiquer avec :

Développement des ressources humaines Canada
Centre des publications
140, Promenade du Portage, Phase IV, niveau 0
Hull (Québec) Canada
K1A 0J9

Télécopieur : (819) 953-7260
http://www.hrdc-drhc.gc.ca/sp-ps/arb dgra

Table of Contents

List of Tables

List of Figures

Abstract

This paper examines three inter-related issues: whether past levels of resources, context and opportunity structures carry long-term consequences for subsequent child attainments and behaviours; whether shocks, or transition events, alter the path of these outcomes in a positive or negative fashion; and what role can be played by policy interventions in improving these attainments. In doing so, the paper examines the relationship between a child's past, and who she becomes; that is to say, is her history her destiny?

We use three cycles of the National Longitudinal Survey of Children and Youth, fielded in 1994, 1996 and 1998 and analyse development for children in 4 cohorts: 'the babies' (aged 0 to 2 in 1994 and thus 4 to 6 in 1998); 'the pre-schoolers' (aged 4 to 5 in 1994); the 'middle-aged' children (aged 7 or 9 in 1994) and the 'oldest' children (aged 11 in 1994 and thus 15 in 1998). The child attainments studied here are principally from the cognitive domain: school readiness, measured by the Peabody Picture Vocabulary Test (PPVT), and education attainments, measured by tests of mathematics and reading ability. A descriptive analysis presents preliminary evidence of considerable persistence in child attainments over time. For example, we show that children who were 'low birth-weight babies' in 1994 are twice as likely to have PPVT scores in the bottom quintile 4 years later; 39 percent of children who at age 4 or 5 had low PPVT scores were then observed to have reading scores 4 years later which placed them in the bottom 20 percent.

Using a multivariate framework, we then consider the association between household resources and long-term child cognitive attainments. Specifically, we estimate the link between child, maternal and household characteristics in 1994 with attainments in 1998; we then complement this analysis by adding child 1994 attainment as a predictor of 1998 attainment; finally we consider the consequences of transition events such as the development of a child's activity limitation, the acquisition or loss of full or step siblings, changing school or care giver, marriage, divorce and the household moving into or out of poverty.

Results suggest that characteristics of the mother have important associations with long-term child attainments. Mother's education level is particularly important. Household income in 1994 also has consistent associations, although the magnitude of this effect is small. However, we emphasize that the cumulative effect of bundles of household characteristics which often occur together is large. For example, a 'disadvantaged' child' – one with a young mother who has not completed high school living in a household at the 25th percentile of income in 1994 – obtains scores on the mathematics test in 1998 that are 13 to 22 per cent lower than an 'advantaged child' – one with an older mother possessing a university degree living in a household at the 75th percentile of income in 1998. By contrast, apart from presence of an activity limitation, observable child and neighbourhood characteristics, as well as other parental characteristics – including marital status - have little systematic impact on these attainments. Strikingly, none of the observed transition events have a substantive effect on these outcomes.

Finally, multivariate analysis again shows clear evidence of persistence in child attainments over time. This persistence across time is robust to the inclusion of child, parental and household characteristics as well as to estimation techniques that account for potential endogeneity (or 'reverse causation'). Thus, past child attainments have a statistically significant relationship with future achievements and, in many cases, the association is very large. We illustrate these findings mean that past disadvantage can 'snowball' into ever larger problems for the child over time.

A policy implication of these findings is the general importance of directing economic resources to families with young children (which would include supporting educational attainment for parents). Given the strong evidence of persistence in child educational outcomes, it seems very important to do everything possible to set children onto the best possible developmental path early in life. It would also be advisable to identify 'at risk' children (e.g., those with low PPVT scores at school entry) and direct remedial attention to these children as soon as possible.

1. Introduction

Improving the well-being of children is a major Canadian policy concern. In 1989, a unanimous all-party motion called for the elimination of child poverty by the year 2000. Although this goal has not been met, at both the federal and provincial levels, there now exist a variety of initiatives and programs designed to improve the life chances of Canadian children. The design and implementation of these interventions has been informed by cross-sectional analyses of the relationship between *current* conditions of advantage and disadvantage and *current* child attainments.

This paper examines three inter-related issues: whether past levels of resources, contexts and opportunity structures carry long-term consequences for subsequent child attainments and behaviours; whether shocks, or transition events, alter the path of these outcomes in a positive or negative fashion; and what role is played, and can be played, by policy interventions in improving these attainments. In doing so, it examines the relationship between a child's past, and who she becomes; that is to say, is her history her destiny?

We proceed by focusing attention on the determinants of pre-school readiness and subsequent educational attainments. Based on the literature reviewed in chapter 2, our research will focus on the role played by household resources and transition events, controlling for the context in which the child is situated.

2. Literature Review

Child outcomes occur across a variety of spheres – social, physical/health, emotional, language/communication and cognitive/learning. These outcomes may be viewed as a consequence of the contexts in which the child is placed, the resources – family, community, and public – available to her, opportunity structures that become available, or are closed down, over time, and transition events that are child (school entry, puberty) or household (divorce, job loss) specific. In this section, we selectively review the literatures that speak to components of this framework as a means of identifying the outcomes and predictive factors that are included in our analysis.

2.1 Conceptual Overview

Among economists, the most influential model of the determinants of children's well-being is one which focuses on how parental choices affect outcomes for children (see Becker and Tomes, 1986). For example, parents determine the level of economic resources by deciding how much time to spend working for pay (and they have previously decided upon how much education to pursue which is a determinant of their rate of pay). Parents then decide about how economic resources will be used - for adult consumption, for asset accumulation or for investments in children, where investments in children are "expenditures on their skills, health, learning, motivation, 'credentials,' and many other characteristics" (Becker and Tomes, 1986, p. S5). One key prediction of this approach is that children will fare better when families have more resources to invest in them.

More recent work, and reviews of this literature, offer two important modifications to this approach. Behrman (1997) notes that parental investment in children may reflect "reinforcement" or "compensatory" concerns. For example, a parent may read to a child because the child enjoys being read to. Alternatively, a parent concerned about a child's reading ability may spend more time reading to that child. In the former, associations between a parental behaviour (reading to a child) and a child's reading ability may overstate the "true" level of association because it conflates the child's innate interests with a parent's behaviour. In the latter, associations between the parental behaviour and the outcome may understate the "true" level of association because it ignores the inverse relationship between parental actions and (possibly unobservable) child characteristics.

Haveman and Wolfe's (1995) review sees the basic economic framework described above as being one of three factors that affect children's well-being. These are: 1) the choices made by society which will determine the options available to either children or their parents -- what Haveman and Wolfe call "the social investment" and which partly overlaps with HRDC's notion of contexts, resources and opportunity structures beyond the household; 2) the choices made by the parents about both the quality and quantity of resources devoted to children -- "the parental investment;" 3) the choices made by the children themselves. This last component is seen to be most important for older children.

Researchers in other fields offer alternative theoretical perspectives. The 'socialization/role model perspective' which focuses on the important influences of parents, siblings and peers on the development of children's aspirations, values and behaviour (Seltzer, 1994; Jencks and Mayer, 1990). The 'ecological systems' approach, favoured by many developmental psychologists, argues that development occurs throughout life, and that the timing and context of any significant life event (e.g., parental divorce) will modify its impact on that particular individual (e.g., Bronfenbrenner, 1989). Stress theory and coping strategy perspectives argue that a particular stressful event (again, for example, parental divorce) may change a child's equilibrium path of development though the impact of such a stressful event can be mitigated, or not, depending upon parental coping capacities (e.g., McCubbin et al., 1980). As Haveman and Wolfe argue, these psychological and sociological perspectives emphasize environmental/cultural factors rather than the individual choices/characteristics upon which economists focus.

2.2 Household Resources and Socioeconomic Status as Determinants of Child Attainments

Past research with the NLSCY has concluded that children from low-income families generally experience worse outcomes than other children (e.g., Curtis and Phipps, 2000; Kohen, Hertzman and Brooks-Gunn 1998; Ross, Roberts and Scott, 2000; Ryan and Adams, 1998). For example, Ross, Roberts and Scott (2000 -- see also Ross and Roberts, 1999) present evidence from the 1994-95 wave of the NLSCY that "more than one-third of children from low-income families exhibit delayed vocabulary development, compared to only eight percent of children in high-income families (p. 52)." Ryan and Adams (1998) conclude "socioeconomic status has a large and pervasive influence over children's school achievement." Moreover, the finding that children from low-income families have worse outcomes than other children appears to be true for other countries as well (see Phipps, 1999b and 2002; McLanahan and Sandefur, 1994; Duncan and Brooks-Gunn, 1997). However, there is much debate about the *magnitude*, and hence policy significance, of the association between socioeconomic status and child outcomes. Some recent Canadian studies that investigate the link between current household income or poverty status and child well-being find relationships that are small in magnitude or statistically insignificant (for example, see Curtis et al. 2001, Dooley et al. 1998a,b).

Findings of weak or non-existent association between resources such as household income and child outcomes may reflect reliance, occasioned by the use of a single cross section of data, on current income that may be a noisy indicator of long-term household resources, or 'permanent income.' Studies in the United States have argued that permanent income may be a better measure of economic resources in the context of studying the role of income as a determinant of children's well-being (see, especially, Blau 1999, Korenman et al. 1995, Mayer, 1997). The primary data source for this work has been the National Longitudinal Survey of Youth's (NLSY) Mother and Child Supplement that provides very long income histories. At this stage, although there is consensus that permanent income is more important than current income, there is disagreement as to the magnitude of the effects.

Blau (1999) finds only small effects. He claims that the income effects are so small that income transfers to poor families are likely to have very little impact on child development; "*Policies that affect family income will have little direct impact on child development unless they result in very large and permanent changes in income.*" (P.261). Korenman et al. (1995) interpret their results to indicate a 'moderate to large' effect of changes in long-term poverty status on children's cognitive development. Mayer (1997) reviews existing literature and uses several different methodologies and US data sets and concludes that the effect of increases in parental income on child outcomes, ceteris paribus, "is nowhere near as large as many political liberals imagine, neither is it zero as many political conservatives seem to believe" (p. 143). She goes on to say that although the effect on any single outcome may be small, most outcomes seem to be affected by income to some extent, thus increasing income may have a substantial *cumulative* impact.

There are a few Canadian studies examining the links between child outcomes and longer-term measures of family economic status. To date, two longitudinal data sets are available to investigate the association between child health and well-being and socioeconomic status in Canada. Findings using the Ontario Child Health Survey (OCHS) conducted in 1983 and 1987 (OCHS) indicate a consistently significant association between low income or poverty and psychiatric disorders (Offord, Boyle and Jones, 1987), social and educational functioning (Lipman and Offord 1997), and chronic physical health problems (Cadman et al., 1986a) in children. Studies using the longitudinal nature of the OCHS find that changes in income levels are very weakly correlated with changes in the levels of child health (Lipman and Offord 1997 and Lipman, Offord and Boyle 1994). The studies that find a significant relationship between income and child health tend to limit the use of other explanatory variables.

Curtis et al. (2001) investigate the relationship between current low-income and low-average-incomes using the OCHS. The study investigated the presence of emotional problems, cognitive problems, any health problems and an over-all health related quality of life score, the Health Utilities Index Mark 2 (HUI2). As in the Koreman study, children from low-income families suffered from substantially more problems than did children from non-low-income families. For cognitive problems, both current and average low-income were negatively associated, though the effect was larger for average low-income. Curtis et al. (2001) conclude that child outcomes are more strongly related to low-average income than low-current income. Unlike results from the NLSY, they find that the magnitude of income effects to be 'large' for some outcomes.

Curtis and Phipps (2000) re-examine the association between economic resources and children's health and success at school using the 2nd wave of NLSCY data. This again makes it possible to move beyond *current* income and/or poverty status as a measure of the economic resources available to the child. This study also argues that it is possible that the effects of economic resources only appear with a lag, hence it may be that previous year's income is more important than current income. Finally, it is also possible that duration of low-income status is important. These hypotheses are examined with the conclusion that for success at school, it is clearly the longer-term poverty and the two-period average of income that have the largest associations. Economists also argue that stocks of wealth as well as income flows are a vital component of the economic resources available to a family. While the NLSCY does not provide any direct information about family assets, a proxy for home ownership and for the state of repair of the family dwelling are included. Finally, traditional economic reasoning also suggests that, holding income constant, families with more time are better off than those with less. When controls for both housing and available parental time per week are added, results suggest that children who live in owner-occupied housing have better outcomes than children who do not; children who live in housing in need of major repairs have worse outcomes. This represents an additional channel through which economic resources can influence outcomes for children. More hours of parental time available each week, income constant, significantly improves a child's success at school.

In a preliminary study using the 3rd wave of the NLSCY data, Dooley, Lipman and Offord (2001) still find relatively small associations between child emotional/behavioural indicators and 3-period average measures of family income, but somewhat larger associations between these longer-term measures of family income and child cognitive outcomes (math and reading test scores from the 2nd wave of data). They follow Mayers (1997) in examining the hypothesis that families with higher incomes may be able to spend more in ways which are beneficial to the child (specifically, thus far, on recreational programmes, sports, clubs and lessons). While children's participation in recreational programmes has strong associations with family income, their preliminary work suggests that the associations between participation in recreation and other outcomes for children are relatively small. Thus, the preliminary conclusion is that higher income leading to increased 'investment' in recreational programmes is not a particularly important pathway through which income may influence child cognitive outcomes.

To summarize, there is a consensus that children with lower socioeconomic status have poorer outcomes, and this is particularly true for cognitive development and/or schooling attainment measures. There remains disagreement about the magnitude of the associations. There is also a consensus that there are larger associations between longer-term measures of family socioeconomic status and children's attainments.

2.3 The Impact of Transitions on Children's Attainments

Work by Picot, Zyblock and Pyper (1999) using longitudinal microdata from the Survey of Labour and Income Dynamics (SLID) for 1993 and 1994 demonstrates that it is not unlikely that children will face changes, sometimes quite dramatic changes, in family income. These authors conclude that changes in family composition (e.g., divorce/separation or re-marriage of parents) have the largest impact on the probability of a child entering or leaving poverty, respectively. While changing labour market circumstances (e.g., gaining or losing a job) do not have nearly so dramatic an association with the probability of a child changing poverty status, they are much more common. (Phipps (1999a) demonstrates that young children are more likely than any other age group to live in a household in which the head is unemployed.) Picot et al. 1999, conclude that changing family composition and labour market changes are about equally responsible for children moving in and out of poverty in Canada.

There are relatively few Canadian studies that address the consequences of transitions or 'shocks' to family socieconomic status for children's educational attainments. Kohen et.al., (2000) is an exception. They use the 2nd wave of the NLSCY and Structural Equation Modelling techniques to study the consequences of unemployment and/or a drop in family income (of 25 percent or more) on current pre-schooler PPVT scores. Kohen et.al. argue that unemployment can lead to financial hardship and many forms of stress in the household (e.g., parental stress, parental ill health/depression, marital discord) any of which may lead to poorer outcomes for children. These authors present descriptive evidence that the PPVT scores for children in dual earner families were 12 points higher (4/5th of a standard deviation) than children who lived with two parents neither of whom had paid work. This pattern holds in multivariate analysis – that is, children with two unemployed parents (or children in lone-parent households whose parent is unemployed) or children who experience significant drops in family income have worse PPVT scores than other children. The effects of parental unemployment are mediated somewhat, though they remain strong, through maternal mental health, family functioning and, especially, positive parenting behaviour.

Using US data Duncan et. al. (1994) estimate a 'change' model to address the potential problem of unobserved heterogeneity (ie., the possibility that some unobservable factor such as intelligence, drive or motivation may be correlated with both child outcomes and family income). Specifically, they find that the change in IQ measure between ages 3 and 5 has a highly statistically significant relationship with the change in parental income over the same period. This methodology still has the potential problem that whatever *caused* the income change may also have caused the developmental change. Mayer (1997) tests for possible omitted-variable bias by including measures of parental income *after* the outcome in question was measured (hence arguing that the outcome could not have been caused by the income). She finds that the estimated impact of income is much smaller in this case. A problem is that families may well have anticipated future income, and adjusted consumption or other behaviours accordingly (e.g., individuals about to finish medical school).

Another approach to solving the unobserved heterogeneity problem is to use sibling differences, the approach employed by Duncan et al., 1998 using PSID data. Results again suggested that family income is particularly important for cognitive outcomes; that family income is most important during the early years; and that the association between income and child outcomes is non-linear, with income being most important for the lowest income children (see also Smith, et al., 1997).

It is worth noting that, as argued above, 'coping strategy' and 'stress theory' perspectives suggests that negative life circumstances or vulnerabilities may be offset if parents or children have healthy coping strategies which help them to 'cushion the blows.' In addition to the Kohen et al. 2000, study noted above which emphasizes maternal mental health, family functioning and positive parenting behaviour as mediators of negative economic shocks, some cross-sectional work is helpful in pointing to variables which may help promote resiliency of children who are particularly vulnerable. Jenkins and Keating (1998) emphasize the role played by close relationships, particularly with adults other than parents, as well as sibling and peer relationships, in helping children cope with particularly stressful circumstances. Ross, Roberts and Scott (1998) again emphasize the mediating role of parenting behaviour, in this case for children in lone-parent families. Landy and Tam (1998) once again emphasize parenting practices and social support in helping children cope with multiple-risk situations.

2.4 Critical Periods in Child Development

While economists have paid less attention to the idea that the impact of a negative life event experienced by a child may depend upon *when* in the child's life it occurs (though see Danziger and Waldfogel, 2000), scholars in other disciplines argue that this can be extremely important (Duncan et al., 1998). There is much emphasis placed on what occurs during the 'early years' of development both because this affects biological pathways and because development is a cumulative process so that outcomes/attainments at any age can have important consequences for opportunities and capacities at subsequent ages (e.g., Hertzman, 2000; Mustard, McCain and Bertrand, 2000).

With respect to the importance of 'critical periods,' Duncan and Brooks-Gunn (1997), reporting upon the results of a coordinated analysis by 12 groups of researchers working with 10 different developmental data sets, most of which offered longitudinal measurement of family income as well as measurements of the achievement of children at various point in life, conclude: "That economic conditions in early and middle childhood appeared to be far more important for shaping ability and achievement than were economic conditions during adolescence." This points to the need for the analysis of the determinants of the attainments to be undertaken for separate age cohorts.

3. Overview of Empirical Strategy

As noted in the introduction, this project examines three inter-related issues: whether past levels of resources, contexts and opportunity structures carry long-term consequences for subsequent child attainments and behaviours; whether shocks, or transition events, alter the path of these outcomes in a positive or negative fashion; and what role is played, and can be played, by policy interventions in improving these attainments and behaviours. In this section, we provide some additional detail about the data set we employ – Cycles 1, 2 and 3 of the National Longitudinal Survey of Children and Youth (NLSCY). We then explain how, by exploiting two features of the NLSCY – its longitudinal nature, and the existence of cohorts of children for whom different types of information are available – the research questions noted above are addressed.

The National Longitudinal Survey of Children and Youth (NLSCY) is an ongoing survey of Canadian children designed to help analyze child development and well-being. There are three cycles of data available to date with interviews in 1994, 1996 and 1998 with the expectation of continued biennial interviews until the children reach the age of 25. In addition to the longitudinal file, cross-sectional data are available for each survey year yielding nationally representative results when the sampling weights are applied. In 1994, the sample was of children 0-11 years of age , in 1996 0-13 and in 1998 0-15. Additional children are added to the cross-sectional file each survey year in order to maintain both nationally and provincially representative samples for each year.

The sample for the NLSCY was originally drawn from the Labour Force Survey (a monthly survey by Statistics Canada used to produce labour force information). The survey uses a multistage probability sample where each province is an independent sample. Through stratification, cities, small urban areas and rural areas are broken down into clusters of dwellings from which households are surveyed. From the LFS, households containing children could be selected for the NLSCY. Note that the LFS excludes those living in institutions and on Indian Reserves. In cycle one, 22,831 children were interviewed which included about 5,000 children from households of those in the National Population Health Survey. These children were dropped from the second cycle due to budget constraints. In cycle one, up to four children per household were interviewed but by cycle two, only two children per household were interviewed (for those households with more than two children, those interviewed were randomly selected). Again, this was due to budget constraints. However there was a large increase in the number of children 0-5 year olds interviewed leading to a total sample of 20,025 children in cycle two. For cycle three, no new siblings of children already in the survey were interviewed (as they were in cycle two) but new children selected from the Labour Force Survey households and birth registries increased the sample size to 31,194 0-15 year olds.

Much of the information which we use for our analysis was collected during a personal interview with the ‘person most knowledgeable’ about the child (pmk), generally the child’s mother. In fact, for over 90% of children in the survey, the mother is the pmk. We also use results of tests administered to the child.

In carrying out our research, we focus on four cohorts. Note that for each cohort, the NLSCY has detailed "historical" data (i.e. information on families, neighbourhoods and schools) collected in the first cycle, as well as indicators of attainments as of 1994:

- Children aged 0-2 during cycle 1, the "baby cohort". For these children, there is detailed "attainment" information, for example, on birth-weight, chronic health conditions, and nursing history while infants. The 'youngest children' are aged 4-6 when cycle 3 was fielded and thus will taken the Peabody Picture and Vocabulary Test (PPVT), an indicator of school readiness;

- Children aged 4-5, whose attainments as of 1994 are captured by the PPVT fielded during cycle 1, the "pre-schooler cohort". They are now aged 8-9 when cycle 3 was fielded and will have completed indicator tests on math and reading ability.

- Children aged 7 and 9 in 1994, who completed the math indicator test during cycle 1, the "middle cohort" (there are problems with the 1994 math test for children aged 8 and 10 and hence we exclude these children). Children in our middle cohort are aged 11 and 13 when cycle 3 was fielded and will have completed further indicator tests on math and reading ability in 1996 and 1998.

- Children aged 11 in 1994 and 15 in 1998, the 'oldest cohort.' These children wrote the math indicator test in 1994 and completed subsequent tests (math and reading) in 1996 and 1998.

By using these four cohorts, we can answer:

1. To what extent is past history "destiny"? That is: if in the youngest cohort, conditions of advantage and disadvantage at birth are strongly associated with levels of early school readiness; if, in the pre-schooler cohort, such readiness is associated with subsequent early school success; if in the "middle and older" cohorts early school success is associated with desirable subsequent attainments; and if we assume that these associations are stable over time, this analysis can tell us to what extent the level and distribution of early advantage and disadvantage plays out in attainments observed in early adolescence.

2. What role is played by shocks such as family income loss, changes in household structure (resulting from divorce, re-marriage or other changes to household living arrangements) or significant changes in the child's health? To what extent are the implications for children's attainments of shocks asymmetric? Do negative shocks have adverse effects, but positive shocks have limited, if any advantageous effects?

4. Descriptive Results

4.1 "The Baby Cohort"

We begin our descriptive analysis of child outcomes by analyzing transitions for our youngest cohort of children who were aged 0 to 2 years in 1994 and were thus 4 to 6 years in 1998.[1] As noted above, the only 'educational outcome' we can consider for this cohort is the standardized PPVT[2] score which is regarded as a key predictor of 'school readiness' (Baker, et al., 1993). Our central substantive research question is the extent to which 'history is destiny' – i.e., that poor outcomes earlier in life predict poor outcomes later in life. For our youngest cohort, we compare PPVT scores at age 4, 5 and 6 for children whose birth-weights were less than 2,500 grams (i.e., they were classified as 'low-birth-weight babies' – 6.8 percent of the 3,529 children in this cohort) with PPVT scores for children with 'normal' birth weights. Specifically, we rank children according to their PPVT scores and then divide the population into 5 equal groups (or quintiles). We then compare the percentages of low-birth-weight children in each PPVT score quintile with the percentages for normal birth-weight children (where, by definition, for all children, there would be 20 percent in each quintile). As is evident in Figure 1, low-birth-weight children are twice as likely to have PPVT scores in the bottom quintile (36.1 percent versus 18.7 percent); they are one-half as likely to have PPVT scores in the top quintile (10.5 percent versus 20.6 percent). Thus, a purely descriptive analysis indicates that low-birth-weight children carry with them a disadvantage 4 years later as they prepare to enter school.

---

1 Of course, the sample is at this stage restricted to children who were present in both cycle 1 and cycle 3. See Appendix A for a full discussion of attrition bias issues.

2 The child is shown pictures on an easel and must identify the picture that matches the word the interviewer reads out. Raw scores are simply the number of correct responses. We use scores that are standardized for the age of the child.

Figure 1
Low Birth Weight (Cycle 1) and PPVT Score (Cycle 3)
Ages 0-2 in Cycle 1

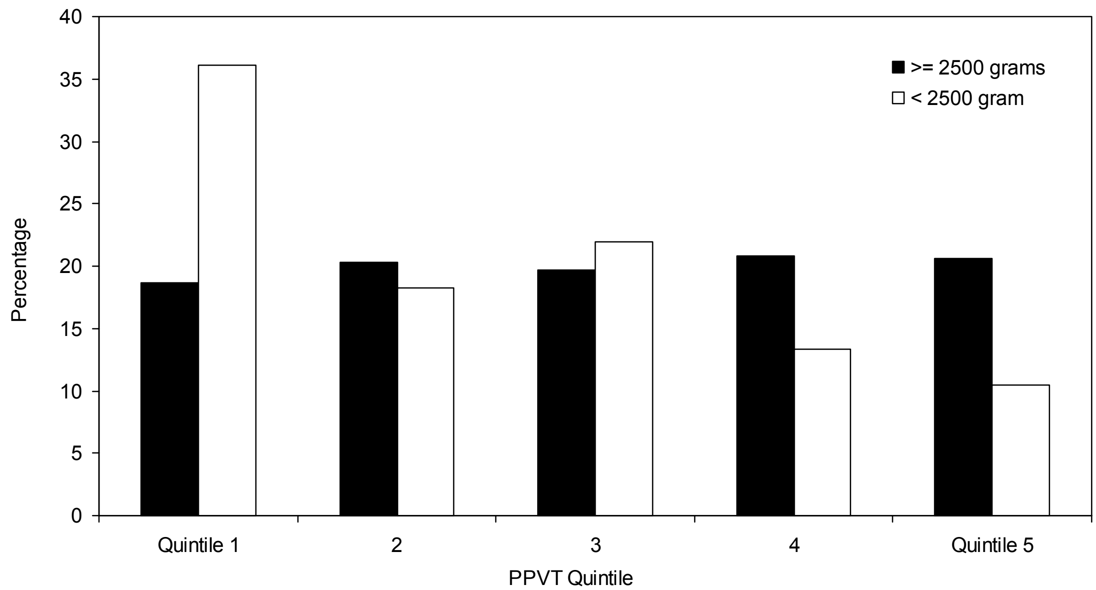
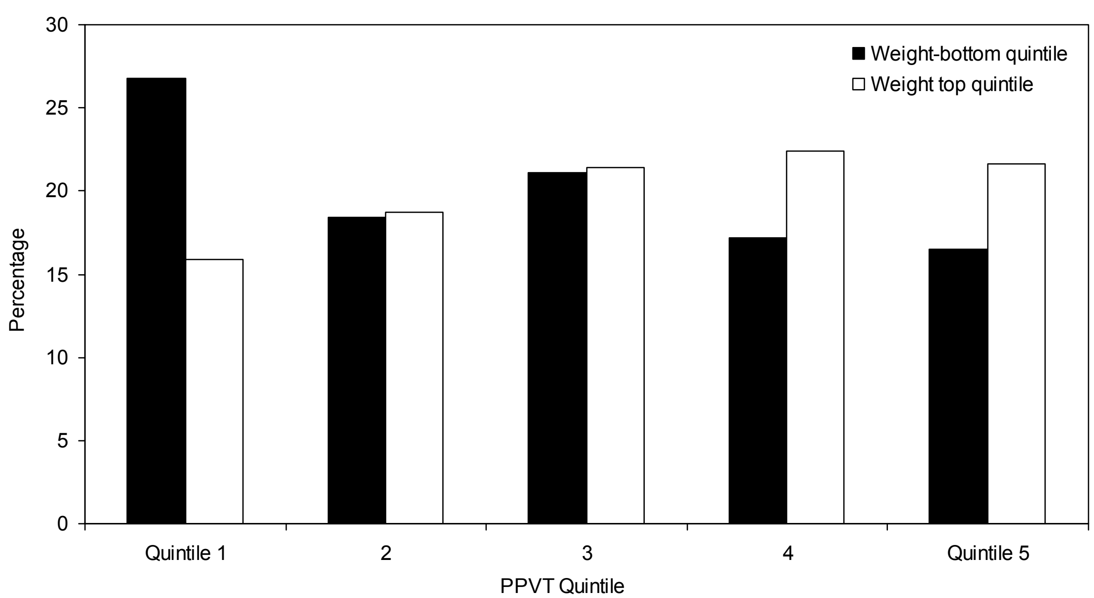


Figure 2
Bottom/Top Birth Weight Quintile and PPVT Quintile (Cycle 3)
Ages 0-2 in Cycle 1



While Figure 1 focuses upon a comparison of children at the very bottom of the birth weight distribution with *all other* children combined, Figure 2 provides a comparison of PPVT scores for children with birth weights in the bottom and top quintiles of the birth weight distribution. In general, being heavier at birth appears to convey little advantage to a child, except insofar as the probability of being at the very bottom of the PPVT score distribution is lower (i.e., 15.9 percent compared to 26.8 percent of children with low birth weights). Thus, the information provided in Figure 1 is the essence of the story here – children with low birth-weights are less 'ready for school' than their counter-parts with birth weights in the normal range.

A second outcome for infants provided in the 1994 NLSCY is a measure of 'activity limitation.' That is, the child's pmk was asked "whether the child has any long term conditions or health problems which prevent or limit his/her participation in school, at play, or in any other activity for a child of his/her age.' In our youngest cohort, only 2.2 percent of 0 to 2 year old children were reported to have activity limitations in 1994. We compare PPVT scores in 1998 for young children with and without activity limitations in 1994. Results are reported in Figure 3 that again suggests considerable persistence of problems. Thus, 48.0 percent of children with activity limitations at age 0 to 2 have PPVT scores in the bottom quintile for 4 to 5 year old children in 1998; only 11.3 percent have PPVT scores in the top quintile. For children without activity limitations, essentially 20 percent are located in each PPVT quintile.

Figure 3
Child is Limited in Activity (Cycle 1), PPVT Score (Cycle 3)
Ages 0-2 in Cycle 1

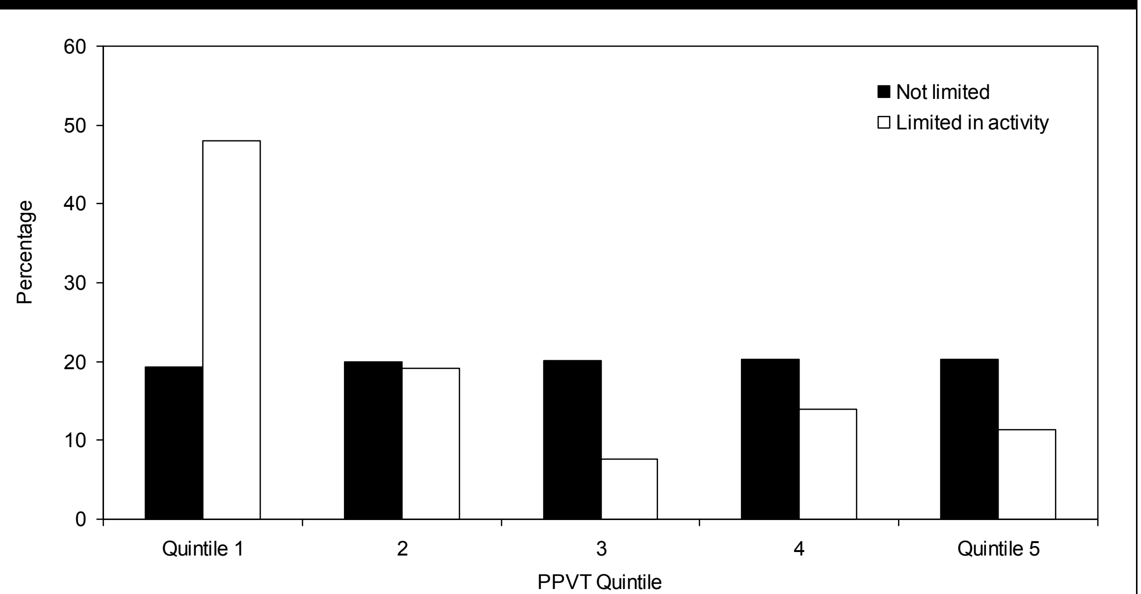
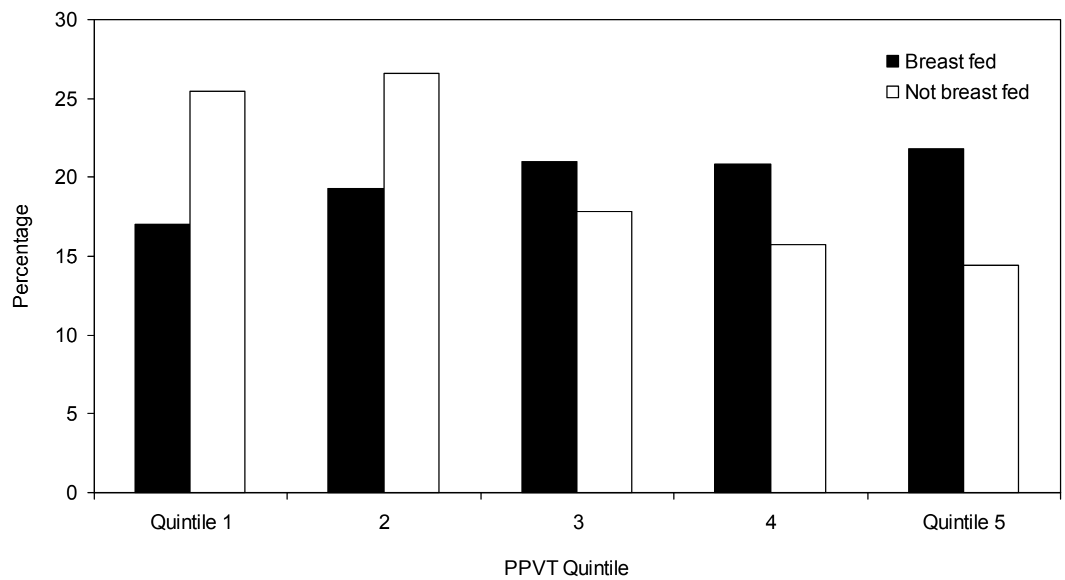


Figure 4 presents the distribution of PPVT scores at age 4 to 5 separately for children who were and were not breast fed as infants (25 percent were *not* breast fed).[3] Children who were *not* breast fed are more likely, at ages 4 and 5, to have PPVT scores toward the bottom of the distribution (25.5 percent in the bottom quintile and 26.7 percent in the second to bottom quintile versus 14.4 percent in the top quintile).

Figure 4
Breast Fed (Cycle 1) and PPVT Score (Cycle 3)
Ages 0-2 in Cycle 1

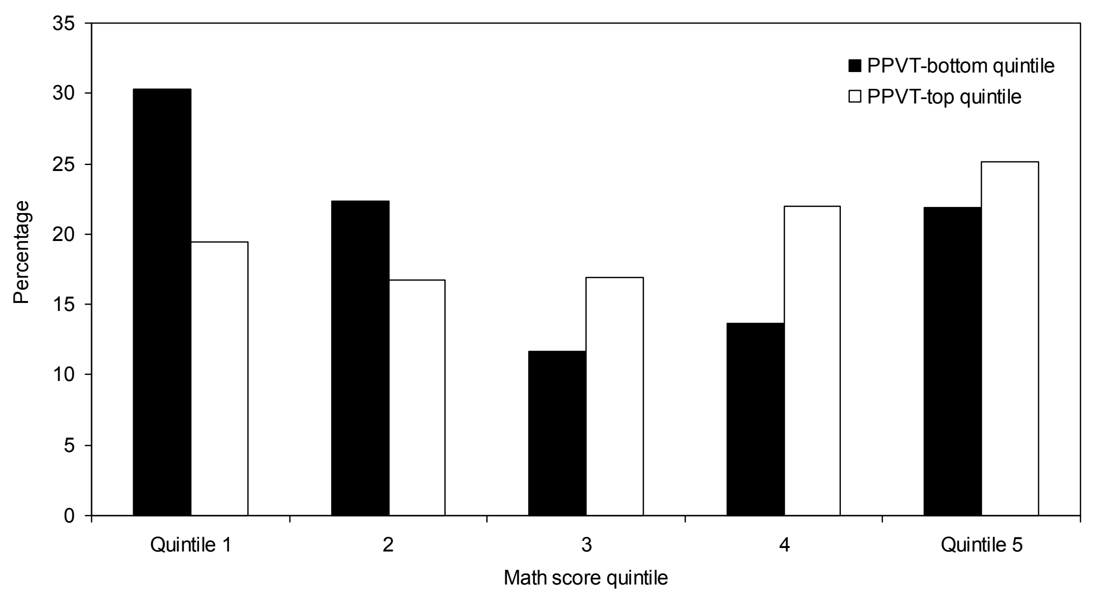


[3] These results do not imply a causal link between breast-feeding and school readiness. For example, children with serious health problems at birth are then less likely to be breast-fed and the health problems rather than the lack of breast feeding are what is important for eventual school readiness. It is also possible that 'breast-feeding' proxies for other parenting attitudes/attributes (e.g., health of mother, need to return quickly to paid work) which may connect to eventual school readiness. Finally, of course, breast-feeding itself may play a direct role, for example by boosting the child's immune system in infancy.

4.2 "The Pre-schoolers"

The second cohort that we study is that of children aged 4 to 5 years in 1994 who are thus aged 8 to 9 years by 1998. For this group, we have two educational outcomes from Cycle 3 to consider.[4] In Figure 5, we use PPVT scores from tests administered to the children in Cycle 1 as an early measure of 'school readiness' and math scores[5] obtained during the Cycle 3 interviews as later measure of educational outcomes. In Figure 5, we compare positions in the distribution of math scores for children who, 4 years earlier, had, respectively, PPVT scores that placed them in the bottom and top quintiles of the PPVT score distribution. It seems clear that children who had low vocabularies in 1994 were relatively much more likely to score badly on the math test administered in 1998 (30.3 percent had math scores in the bottom quintile compared to 20 percent of all children). Children with the best vocabularies in 1994 were also more likely than the average to have the best math test results 4 years later (25.1 percent versus 20 percent for all children). Again, notice that disadvantage in the earlier time period appears to have a larger association with disadvantage in the later time period than advantage in the earlier time period has with advantage in the later time period.

Figure 5
PPVT Score Quintile (Cycle 1) and Math Score (Cycle 3)
Ages 4-5 in Cycle 1



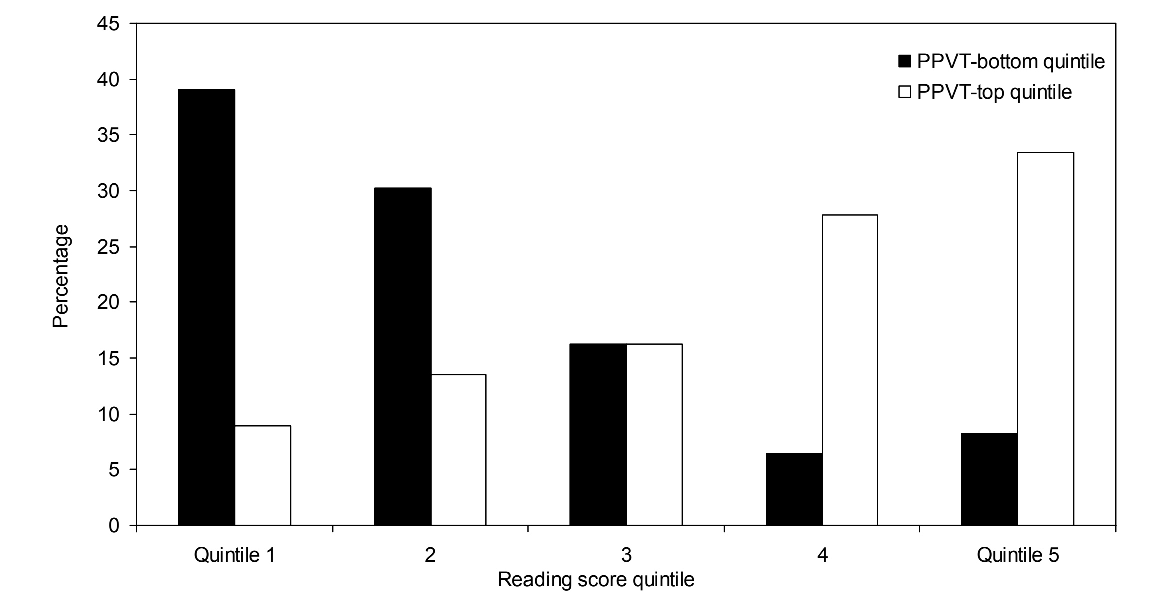
---

4 1,060 children in the pre-schooler cohort had both PPVT scores in 1994 as well as math and reading scores in 1998.

5 Math tests were administered in the classroom, by the child's teacher. The test consisted of 15 questions designed to assess competence with mathematical operations such as addition, subtraction, multiplication and division with whole numbers, decimals, fractions, negatives and exponents. Problem solving involving percentages and the order of operations were also measured. We use scores which have been standardized for child age.

Children in the pre-school cohort also received a reading test[6] in Cycle 3. Hence, we are able to examine more than one measure of educational achievement for this group of children. Figure 6 compares positions in the distribution of reading scores for children with relatively high and low PPVT scores 4 years earlier. Not surprisingly, vocabulary at age 4/5 has a much stronger relationship with reading success at age 8/9 than with math success. 39 percent of children in the bottom quintile of the PPVT score distribution in 1994 were also in the bottom quintile of the reading test score distribution in 1998; only 8.2 percent had moved from the bottom quintile of the PPVT distribution to the top quintile of the reading test distribution. On the other hand, 33.4 percent of children with very high vocabularies in 1994 also scored in the top quintile in their reading test in 1998; only 9.0 percent of those at the top of the PPVT distribution had reading test scores in the bottom 20 percent 4 years later.

Figure 6
PPVT Score Quintile (Cycle 1) and Reading Score (Cycle 3)
Ages 4-5 in Cycle 1



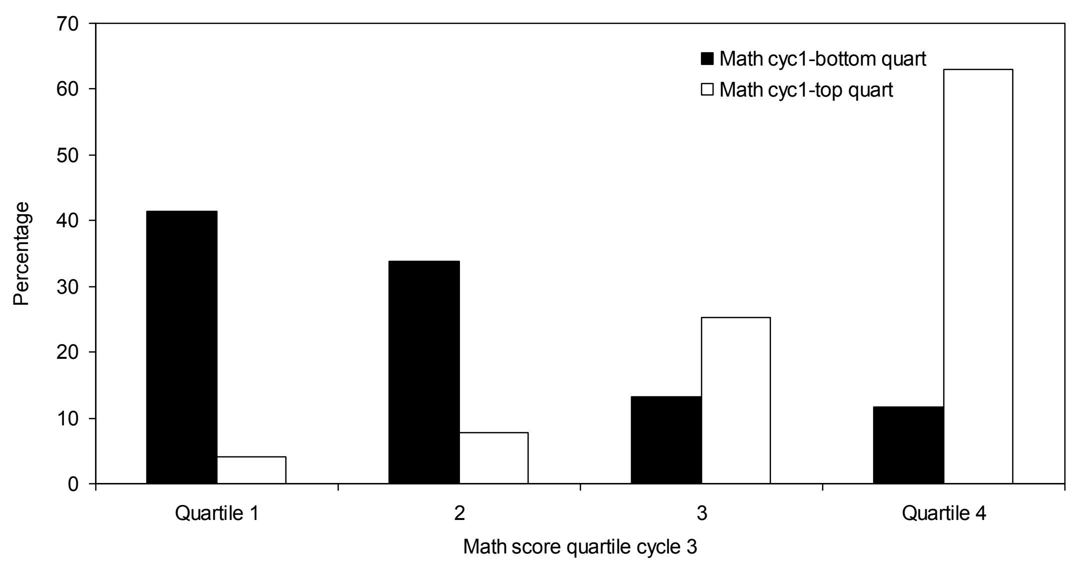
[6] The reading test is also administered in the classroom by the child's teacher. The test is designed to measure basic reading skills such as information recall, analysis of passages, identification of the main idea, interpretation of the main idea, interpretation of various types of writing and critical evaluation. For each grade level, the test consists of four reading passages totalling 20 questions. We use scores standardized for the child's age.

4.3 "The Middle Cohort"

We label children who were aged 7 or 9 years in 1994 the 'middle cohort.'[7] For this group, aged 11 or 13 in 1998, we have math scores from tests written in 1994 as 'earlier' measures of attainment. For 'later' measures of educational attainment, we then have math scores from subsequent tests taken in 1998 as well as reading scores also obtained in 1998. We have 489 children with the necessary information on these variables – a much smaller sample size than was available for the younger two cohorts.

Figure 7 illustrates math scores for Cycle 3 (1998) for children with math scores that were, respectively, in the bottom and top quartiles[8] of the 1994 math scores distribution. Of children who had math scores in the bottom 25 percent in 1994, 41 percent remained in the bottom 25 percent by 1998 while only 11.6 percent had moved up to scores in the top 25 percent of the distribution. On the other hand, of those who scored in the top 25 percent in 1994, only 5.1 percent had fallen to the bottom of the distribution while 63.0 percent remained at the top of the math scores distribution. Again, these descriptive results suggest considerable persistence of child outcomes over time.

Figure 7
Math Score Top/Bottom Quartile (Cycle 1) and Math Score Quartile (Cycle 3)
Ages 7,9 in Cycle 1

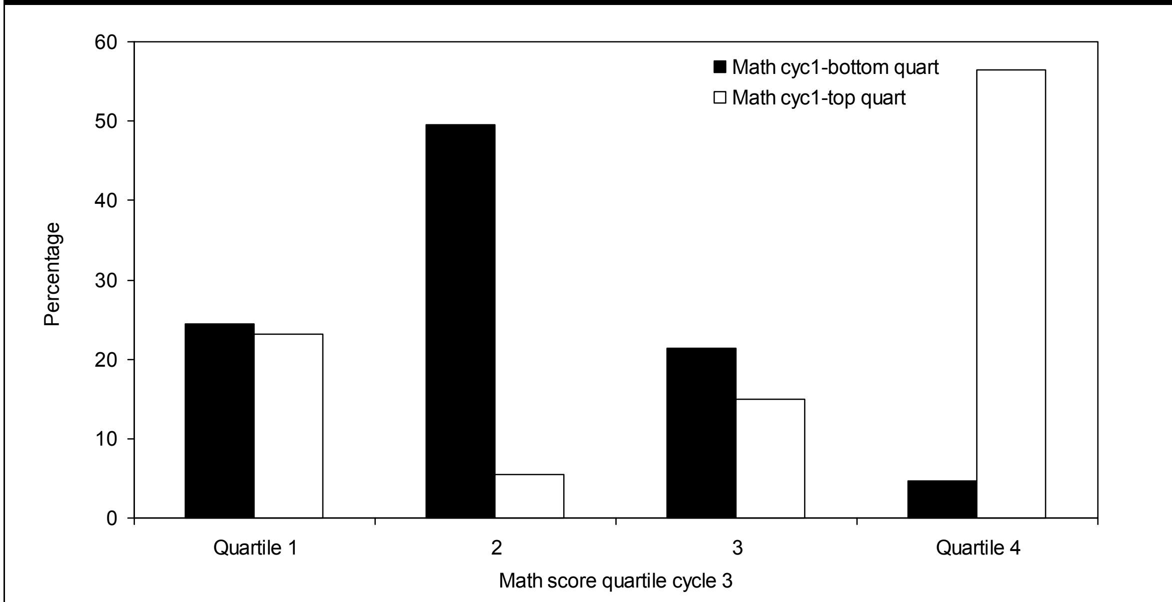


7 Recall that we exclude children aged 8 or 10 years in 1994, as there were problems with the 1994 math test for these children.

8 Given the much smaller sample, we were restricted, in order to ensure confidentiality, to dividing children into 4 groups or 'quartiles,' according to their positions in the over-all distribution of math and reading scores.

Essentially the same exercise is repeated in Figure 8 except that in this case, position in the reading scores distribution in 1998 is used as the 'final' outcome indicator (position in the 1994 math scores distribution is again used as an indicator of initial attainment). In this case, it is very clear that children who had scores among the top 25 percent on the early math test (i.e., in 1994) were very likely to perform among the top 25 percent of children in their subsequent reading tests (i.e., 56.5 percent had reading test scores in the top quartile in 1998). However, other patterns are somewhat less clear. For example, children in the top and bottom quartiles of 1994 math score distributions were equally likely (23.1 percent and 24.5 percent, respectively) to have 1998 reading scores in the bottom 25 percent of the distribution. Children with very low original math scores were particularly likely to have 1998 reading scores in the 2nd quartile (49.5 percent) so that, over-all, 74 percent of children with low initial math scores were in the bottom half of the reading score distribution 4 years later. In comparison, only 28.6 percent of children with high math scores originally had reading test scores in the bottom half of the 1998 distribution.

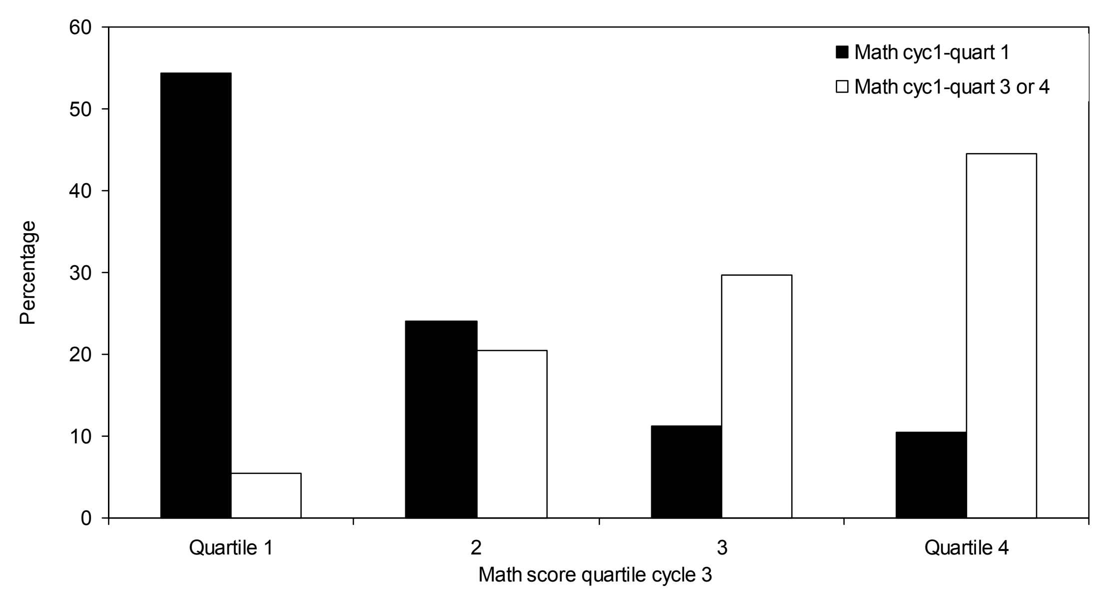
Figure 8
Math Score Top/Bottom Quartile (Cycle 1) and Reading Score Quartile (Cycle 3)
Ages 7,9 in Cycle 1



4.4 "The Oldest Cohort"

The 'oldest cohort' is children aged 11 years in 1994 or 15 by 1998.[9] For this group, we have math score as a measure of attainment in the earlier period and both math and reading scores for 1998. Figure 9 illustrates the connection between earlier and later math scores for our 'oldest' cohort, and persistence is apparent. 54.3 percent of children with scores in the bottom quartile of the 1994 distribution were still located at the bottom of the distribution in 1998; only 10.4 percent of children with 1994 test scores in the bottom 25 percent managed to score in the top 25 percent in 1998. And, children who fared well originally were very likely to score well again (44.5 percent of those who were in the top quartile in 1994 were also in the top quartile in 1998; only 5.4 had fallen to the bottom quartile).

Figure 9
Math Score Top (2) /Bottom Quartiles (Cycle 1) and Math Score Quartile (Cycle 3)
Ages 11 in Cycle 1

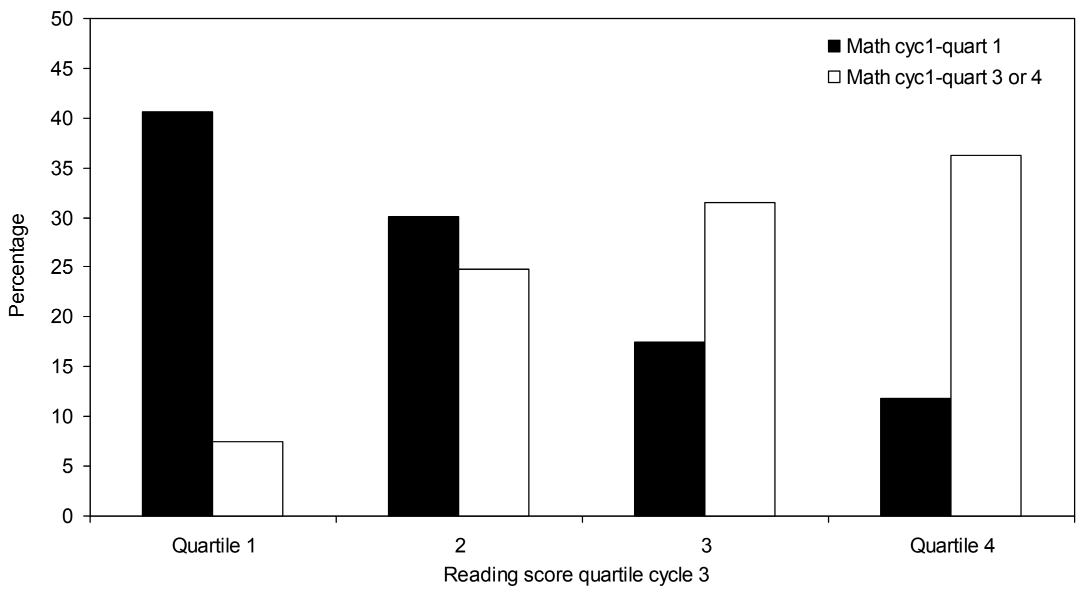


It is also clear in Figure 10 that children with lower initial attainments (as measured via position in the 1994 math score distribution) were likely to have lower 'final' attainments (as measured by position in the 1998 reading test score distribution). Thus, 40.6 percent of children with original math scores in the bottom quartile were also at the bottom of the 1998 reading score distribution; only 11.8 percent scored in the top quartile. In contrast, those who performed well initially were very likely to continue to perform well (36.2 percent of those in the top half of the 1994 math score distribution had reading scores in the top quartile; only 7.5 percent had scores in the bottom 25 percent).

---

[9] Notice that sample size problems are particularly apparent for this group. We have only 306 children with appropriate test scores in both 1994 and 1998 that again necessitates the use of quartiles in order to keep cell sizes large enough to preserve the privacy of respondents.

Figure 10
Math Score Top (2)/Bottom Quartiles (Cycle 1) and Reading Score Quartile (Cycle 3)
Ages 11 in Cycle 1



5. Multivariate Analysis of the Long Term Impact of Household Resources on Attainments

The descriptive data presented in the previous chapter provide some prima facie evidence of the persistence of advantage or disadvantage in attainments over time. In this chapter, we consider a multivariate analysis of the long-term impact of household resources on attainments. Our starting point is an investigation of the associations between child, pmk and household characteristics in 1994 with attainments in 1998. There are two attractions to this approach. First, because we are not using contemporaneous data, these results are less subject to endogeneity bias than studies examining the determinants of current attainments as a function of current characteristics. (For example, children with low attainments may require additional parental care, reducing the time available to earn income. In such circumstances, causality runs from attainments to incomes.) Second, suppose that only *current*, not *past*, household circumstances determine attainments. If this is the case, we would expect, for example, to find no relationship between past incomes and current PPVT, math or reading scores. In other words, examining the links between past characteristics and current outcomes is one way of testing more formally the notion of "history as destiny." In chapter 7, we complement this analysis by including, as an additional characteristics, the child's prior attainments.

Specifically, we estimate the following relationship:

$$Y_{ij} = \alpha + \beta X_{ij} + u_{ij}$$

Where Y_{ij} is the attainment of child i, living in household j, α, β are parameters to be estimated, X_{ij} is a vector of child, pmk and household characteristics and u_{ij} is a white noise disturbance term.

Given the importance placed on differentiating the impacts of household resources by age, we estimate this model separately for our four cohorts. Our dependent variables are indicators of attainments in 1998: the PPVT score for children aged 4-6; and the math and reading scores for children aged 8-9, 11 or 13 and aged 15. We divide our regressors into three broad categories, child, pmk and household characteristics. Child characteristics are age, sex, and and indicator variables denoting quarter of birth, children with activity limitations, number of siblings and birth order. Pmk characteristics are age, sex, education, ethnicity and an indicator variable for suffers a chronic health condition. Household characteristics are province of residence, whether the household is in a rural location, household income per adult equivalent (using OECD equivalence scales) and household income squared.[10] Mean values for these variables for each cohort of children are found in Table 1.

[10] The inclusion of the dummy variables denoting province of residence captures provincial specific effects including curriculum design and resources devoted to education.

Table 1 Descriptive Statistics				
	Children 4-6	**Children 8-9**	**Children 11, 13**	**Children 15**
Dependent variables				
PPVT score	98.6			
Math score		401.3	518.5	631.7
Reading score		223.9	268.6	289.2
Regressors				
Child age (cycle 1)	1.02	4.6	8.0	11.0
Female	48.7%	48.9%	49.2%	51.1%
Born in 2nd quarter	28.3%	24.1%	27.6%	28.0%
Born in 3rd quarter	24.4%	26.3%	26.9%	25.3%
Born in 4th quarter	21.4%	26.0%	18.6%	22.2%
Limited in Activity	2.2%	4.0%	3.8%	5.2%
Siblings in the household	0.98	1.38	1.56	1.44
Child is the eldest child in household	44.8%	42.2%	42.5%	38.7%
Age of the PMK	30.1	33.4	36.5	38.7
PMK is female	97.2%	96.7%	96.0%	96.1%
Lone parent household	14.5%	17.2%	13.7%	15.8%
PMK has less than a high school education	14.8%	14.8%	16.4%	18.8%
PMK has a diploma/certificate	21.8%	21.1%	19.4%	21.6%
PMK has a university degree	17.9%	15.0%	14.9%	15.7%
PMK has a chronic condition	38.5%	40.4%	45.7%	46.5%
PMK is black	1.1%	1.3%	2.1%	1.4%
PMK is North American Indian	4.5%	4.2%	4.0%	4.2%
PMK is Chinese	1.6%	0.8%	1.2%	0.7%
PMK is other race	18.0%	16.5%	16.6%	11.3%
Newfoundland	1.6%	1.9%	2.0%	2.4%
Prince Edward Island	0.4%	0.5%	0.6%	0.6%
Nova Scotia	3.0%	3.1%	2.8%	3.4%
New Brunswick	2.1%	2.4%	2.6%	2.3%
Quebec	23.7%	22.3%	21.6%	22.5%
Manitoba	4.0%	4.1%	3.6%	3.5%
Saskatchewan	3.6%	3.8%	4.1%	4.0%
Alberta	10.9%	10.8%	11.0%	10.6%
British Columbia	12.0%	12.8%	13.4%	10.7%
Rural	16.6%	19.5%	19.3%	21.9%
Adult equivalent income	19449	18214	19392	20223
Adult equivalent income squared	562,779,038	485,111,031	545,515,422	573,313,282
Change variables				
Child developed a long term condition	4.0%	5.5%	6.4%	8.9%
Child no longer has a long term condition	2.7%	4.4%	4.3%	7.1%
Child has more siblings in house	36.7%	14.7%	7.2%	5.3%
Child has less siblings in house	4.4%	7.2%	8.9%	16.1%
Child is now in a step family	6.0%	6.4%	5.4%	4.3%
Child is no longer in a step family	4.3%	5.7%	4.1%	4.1%

Table 1 Descriptive Statistics				
Child changed schools/daycare	55.4%	30.3%	26.6%	21.9%
Child moved	46.5%	38.2%	31.1%	28.6%
PMK divorced	9.4%	7.0%	4.4%	5.4%
PMK married	6.6%	7.0%	4.5%	4.7%
PMK gained education credentials	6.6%	7.8%	7.7%	7.0%
Household poor in 1994 / 1998	16.4%	15.0%	12.7%	9.4%
Household poor in 1994/ not poor in 1998	11.8%	15.7%	12.6%	8.7%
Household not poor in 1994 and poor in 1998	8.2%	7.4%	5.6%	8.7%

In preliminary work, we experimented extensively with this specification. We used different functional representations for incomes, for ages and for education attainments. (For example, we expressed education in terms of years of schooling and different categorical groupings of levels of education attained by the pmk. We used log of incomes and categorical descriptions of the level of household income, including dummy flags for poverty status rather than a continuous income measure.) We experimented with the inclusion of other child, pmk and household variables suggested in the literature, including hours worked outside the home per parent present and home ownership. Following the discussion outlined in chapter 2, we tried average incomes over all three cycles of the NLSCY. We also explored the impact of including neighbourhood characteristics such as: neighbourhood unemployment rate, percentage of adults in the neighbourhood who had not completed high school, proportion of lone mother families in the neighbourhood. Given that we are focussing on educational attainments, we investigated the importance of information from the teacher questionnaires such as teacher characteristics (e.g., level of education, gender, years of experience) and classroom characteristics (e.g., class size; mixed grade class). Using any of these alternative specifications has no meaningful impact on the results reported below and, in general, these variables were not themselves particularly important (and hence in the interests of parsimony, we do not include them in our base specification).

The estimation method for the results reported in Table 2 is weighted least squares, with the weights being those supplied by Statistics Canada to make these data representative given attrition in the NLSCY since 1994. As discussed in Appendix A, these weights eliminate all meaningful attrition bias. In addition, we use the Huber (1967) – White (1980) correction to the regression standard errors to ensure that the results are robust to heteroscedasticity. Lastly, in preliminary work, we experimented with categorical representations of the dependent variable (e.g., dividing scores into 5 rank ordered groups) and estimating for example ordered probits in place of a continuous dependent variable. Again, such an approach does not produce results that are qualitatively different from those reported here.

Table 2
Select Determinants of Child Attainments by Cohort, Weighted Least Squares Estimates

	Children 4-6 PPVT	Children 8, 9 Math score	Children 8, 9 Reading score	Children 11, 13 Math score	Children 11, 13 Reading score	Children 15 Math score	Children 15 Reading score
Child characteristics							
Has activity limitation	-10.87 (3.67)**	15.58 (0.73)	5.00 (0.23)	-1.61 (0.08)	-9.93 (1.35)	-42.65 (2.06)**	-16.59 (1.48)
Is eldest	2.18 (2.28)**	9.12 (1.55)	3.59 (0.90)	21.04 (2.79)**	9.89 (2.96)**	9.28 (0.79)	3.60 (0.64)
Number of siblings	-1.44 (2.98)**	5.38 (1.55)	2.95 (1.19)	8.81 (2.24)**	5.65 (3.25)**	-4.29 (0.56)	0.82 (0.24)
PMK characteristics							
Lone parent	1.13 (0.91)	-10.39 (1.31)	2.87 (0.63)	-6.52 (0.56)	0.71 (0.18)	-12.99 (0.67)	-3.69 (0.56)
Age	0.40 (4.07)**	0.27 (0.51)	0.09 (0.25)	1.41 (1.67)*	0.58 (1.85)*	2.09 (1.71)*	1.14 (2.39)**
Did not complete high school	-3.86 (3.07)**	-16.00 (1.89)*	-24.18 (4.08)**	-14.12 (1.42)	-7.94 (1.67)*	6.04 (0.37)	0.99 (0.12)
Obtained post-high school diploma	1.40 (1.59)	-1.52 (0.25)	2.18 (0.58)	-4.22 (0.53)	-0.76 (0.24)	14.77 (1.01)	1.73 (0.26)
Obtained university degree	3.66 (2.95)**	30.59 (3.57)**	11.32 (1.60)	23.48 (2.04)**	8.40 (2.02)**	63.64 (3.52)**	8.08 (1.44)
Household characteristics							
Income (x 1000)	0.23 (3.87)**	-0.0003 (0.01)	0.52 (1.96)**	1.08 (1.06)	0.37 (1.00)	2.95 (1.99)**	1.02 (1.70)*
Income squared (x 100000)	-0.0009 (1.81)*	-0.0002 (0.56)	-0.00025 (1.07)	-0.00069 (0.44)	-0.00011 (0.22)	-0.0039 (2.10)**	-0.0016 (2.02)**
R2	0.17	0.31	0.17	0.29	0.16	0.25	0.19
Sample size	3084	981	982	872	873	446	445
Mean, dependent variable	98.6	401.3	223.9	518.5	268.6	631.7	289.2

Notes:
1. Absolute value of t statistics in parentheses. 2. Standard errors calculated using Huber (1967) – White (1980) method. 3. * significant at the 10% level, ** significant at the 5% level. 4. Variables included but not reported are child age, sex, quarter of birth, PMK suffers chronic illness, PMK lone parent status, PMK race, province of residence, lives in rural area. 5. Weights are from Statistics Canada (200x) to account for attrition and sample representativeness.

Selected results are reported in Table 2 by cohort. There are several striking features. First, the characteristics listed here have the most well-measured impact on attainments when we restrict attention to the youngest cohort, children who, as of 1998, were aged 4-6 (and thus were 0-2 when the first cycle was fielded). This is consistent with the claim made in our literature review that, generally, household resources and characteristics become progressively less important as children become older. Second, observable child characteristics have some impact on these attainments. Having an activity limitation at the time of first observation is associated with a large reduction in PPVT scores 4 years later, and also has an adverse effect on math and reading scores for children in the oldest cohort. Being the eldest child has a small, positive impact on attainments, but one that is not always well-measured. There is no consistent pattern to the coefficients on the number of siblings. Other child characteristics such as age, sex and quarter of birth never have a statistically significant impact on the attainments examined here.

Education levels of the pmk have a consistent impact on attainments. Relative to the omitted category, pmks who have a high school diploma, children with pmks with less than high school education consistently have poorer attainments; children with careers with university degrees consistently have better attainments. Furthermore, these effects are large in magnitude. Consider a hypothetical example of two children, aged 5 in 1998 (i.e. at the time of the third cycle of the NLSCY) who are identical in all characteristics used in these regressions save the education of their career. The pmk of one child did not complete high school; the pmk of the second child has a university degree. The difference in PPVT scores of these two children is 7.5 points or 7.6%.[11] A similar calculation for older children produces percentage differences in attainments ranging from 3 to 16 per cent. Note too that these results are obtained after controlling for carer marital status, age, and household income levels. Children with older pmk's also tend to obtain higher scores although the effect is not always well measured. Across all attainments, and holding all other characteristics constant, a child with a pmk aged 20 would have scores 4-7% lower than a child with a pmk aged 35. By contrast, being a lone parent has no statistically significant impact on these attainments when controlling for these other child, career, and household characteristics. This finding is consistent with Dooley et.al., 1998a or Curtis and Phipps (2001), for example who find that cognitive outcomes for children living in lone-mother households are not necessarily worse than those for children living in two-parent families, other factors equal.

Incomes appear to have some effect on the outcomes considered here, though the pattern is uneven across children of different ages. Also, the coefficients are difficult to interpret directly because the magnitudes of the dependent variables vary by age and measure. One way of overcoming this is to re-express this relationship as an elasticity. With the exception of the math score for children 8 and 9 years old, this is remarkably small and constant for the youngest three cohorts, ranging in value from 0.03 to 0.05. However, it is three to four times larger for the small sample of children aged 15, that is the oldest cohort. For these young adolescents, the elasticity is 0.15 for the math score and 0.12 for the reading score. The latter implies that a 10 per cent increase in equivalent household incomes is associated with a 1.2 per cent increase in the reading score.

---

[11] The figure of 7.5 points is obtained by adding up the absolute value of the coefficients on "did not complete high school" and "obtained university degree".

Another way of exploring the magnitude of the impact of these characteristics, as measured in 1994, on outcomes measured in 1998, is via calculating the size of the change they induce relative to some base value. To do so, we first calculate the predicted attainment score for a child in a defined "base category". Here, this is a Caucasian boy, born in the first quarter of the year, living in urban Ontario, whose pmk is female, of average age, has completed high school and whose family have the same income as the average family in this sample. We then vary select characteristics and see how this affects the attainment in question.

Results are reported in Table 3 and in Figures 11 and 12. Table 3 reports the impact on attainments, expressed as a percentage change, of varying child, pmk and household characteristics. The striking feature is that, individually, few changes have an impact of any meaningful magnitude on these attainments. The exception to this is changes in the education level of the pmk which tends to have a larger impact. However, looking at these characteristics individually may obscure the fact that in practice, sets of characteristics 'cluster' together. For example, women who have children at a very young age, tend to have less education and live in poorer households. Conversely, older mothers have more education and, all other factors held constant, tend to live in better off households. This observation suggests that considering the cumulative impact of these changes may be more instructive. Figures 11 and 12 provide us with this information. Figure 11 shows the percent change in attainments by cohort relative to the base category for children living in poor households (those at the 25th percentile of adult equivalent incomes) whose pmk has less than a high school education and who is 10 years younger than the mean age for that cohort. The cumulative impacts of these characteristics producing a larger effect, ranging from a fall in attainments of 4 to 12 percent. Figure 12 shows the percent change in attainments by cohort relative to the base category for children living in better-off households (those at the 75th percentile of adult equivalent incomes) whose pmk has a university education and who is 10 years older than the mean age for that cohort. The cumulative impacts of these characteristics producing a larger effect, ranging from an increase in attainments of 6 to 16 percent. There are also large differences when we compare across these two cases. A 'disadvantaged' child – one with a young mother who has not completed high school living in a household at the 25th percentile of income in 1994 – obtains scores on the mathematics and reading tests in 1998 that are 13 to 22 per cent lower than an 'advantaged child – one with an older mother possessing a university degree living in a household at the 75th percentile of income in 1994.

Table 3
Simulating the Impact of Changes in Household Resources on Child Attainments

	Age 4-6	Age 8-9		Age 11, 13		Age 15	
	PPVT	**Math**	**Reading**	**Math**	**Reading**	**Math**	**Reading**
Base	98.1	400.5	225.7	502.6	261.4	560.08	279.4
Change child							
Has activity limitation	**-11.1%**	3.9%	2.2%	-0.3%	-3.8%	**-7.6%**	-5.9%
Eldest	**2.2**	2.3	1.6	**4.2**	**3.8**	1.7	1.3
Number of siblings	**-1.5**	1.3	1.3	**1.8**	**2.2**	-0.8	0.3
Change PMK							
No high school	**-3.9**	**-4.0**	**-10.7**	-2.8	**-3.0**	1.1	0.4
University	**3.7**	**7.6**	5.0	**4.7**	**3.2**	**11.4**	2.9
Ten years younger than mean	**-4.1**	-0.7	-0.4	**-2.8**	**-2.2**	**-3.7**	**-4.1**
Ten years older than mean	**4.1**	0.7	0.4	**2.8**	**2.2**	**3.7**	**4.1**
Change household							
Income to 25th percentile	**-1.4**	0.1	**-1.2**	-1.1	-0.8	**-1.7**	**-1.0**
Income to 75th percentile	**1.8**	-0.2	**1.4**	1.4	1.1	**1.8**	**0.9**

Notes: 1. Percentages are changes in attainment given change in child, PMK or household characteristic relative to base case. 2. Figures in bold refer to changes based on parameters that are statistically significant at the 10 or 5 per cent level. 3. For the base case, we set all the dummy variables in the base case specification to zero and any continuous variables at the mean for that age group.

Figure 11
Percent Change in Attainments, Children in Poor Households, with Young, Poorly Educated PMKs

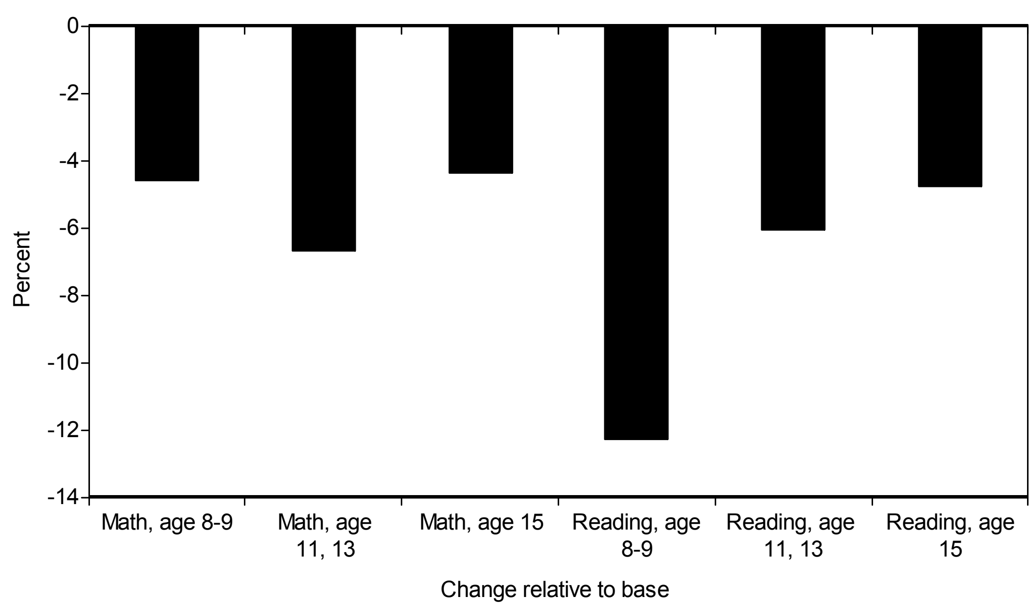
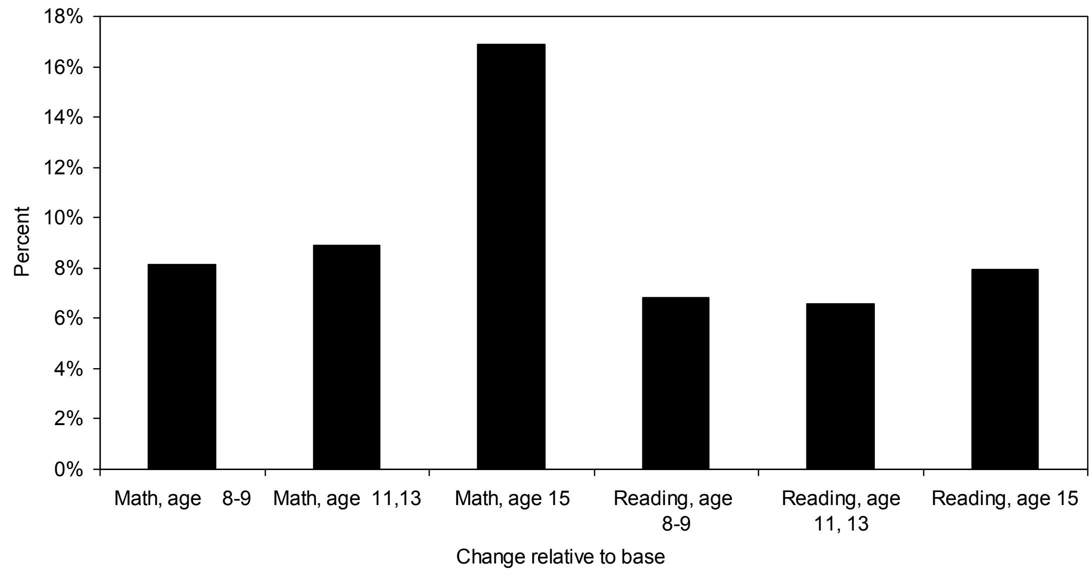


Figure 12
Percent Change in Attainments, Children in Better Off Households with Older, Better Educated PMKs



To conclude, in this chapter we have considered the impact of 'history' as represented by pmk and household characteristics as of cycle 1 (1994) on attainments in cycle 3 (1998). Using multivariate regressions, we find that the impact of these characteristics varies by age of the child. Generally, the education of the pmk tends to have a larger, better measured impact than most other characteristics. Incomes tend to have a statistically significant impact, but one that is relatively small in magnitude, although there is some suggestion in these data that this effect is larger for older children. Larger effects are observed once we recognize that certain characteristics tend to cluster together. We find that children with young, poorly educated pmks living in relatively poor households have attainments 4 to 12 percent below those of the base case; by contrast children with older, better educated pmk's have attainments 6 to 16 percent above the base.

6. Multivariate Analysis of the Importance of Transitions for Child Attainments

Building upon the basic multivariate analysis described in the previous section, we now consider the importance for children's attainments of 'shocks' which may have been experienced during the 1994 through 1996 through 1998 period. The issue of concern is whether important transition events can alter the path of child outcomes established in an earlier period (and we would expect shocks to be most important for the youngest children if they 'bump' the child onto a different developmental path).

For the purposes of this analysis, we identify three kinds of shocks. The first category relates to the child directly: 1) the child developed a long-term health condition at some point between 1994 and 1998 (4.0 percent of the baby cohort; 8.0 percent of the oldest group - see Table 1); 2) the child recovered from a long-term health condition (2.7 percent of the babies; 7.1 percent of the oldest group); the child had siblings added to the household, either biological or step (36.7 percent of the babies; 5.3 percent of the oldest group); 3) the child had siblings leave the household, for example because an older sibling left home or because of a divorce/separation (4.4 percent of the babies versus 16.1 percent of the oldest group); 4) the child now lives in a step-family and did not in 1994 (6.0 percent of the youngest children; 4.3 percent of the oldest[12]); 5) the child no longer lives in a step family, but did in 1994 (4.3 percent of the youngest group; 4.1 percent of the oldest group); 6) the child changed schools or daycare[13] between 1994 and 1996 or between 1996 and 1998 (55.4 percent of the youngest group; 21.9 percent of the oldest group); 7) the child/family moved[14] to a new place of residence (46.5 percent of the youngest children versus 21.9 percent of the oldest).

The second form of transition relates to the pmk: 1) the pmk was divorced or separated between 1994 and 1996 or 1996 and 1998 (9.4 percent of pmk's of the youngest children; 5.4 percent of pmk's of the oldest children); 2) the pmk was married/re-married[15] during the study period (6.6 percent of pmk's of the youngest children; 4.7 percent of pmk's of the oldest children); 3) the pmk completed a higher level of education between 1994/96 or 1996/98 (6.4 percent of pmk's of the youngest group; 8.5 percent of pmk's of the oldest group).

[12] We add this variable to allow for the possibility that the addition of a new biological sibling has different associations with children's attainments than the addition of new step-siblings.

[13] This does not include 'automatic progression' for example from elementary school to junior high.

[14] Any change of residence involving a change of postal code is counted.

[15] This includes legal and common-law marriages.

The final form of shock relates to family finances[16]: 1) the household was poor in *both* 1994 and 1998 (16.4 percent of the youngest children; 9.4 percent of the oldest); 2) the household was poor in 1994 but not poor in 1998 (11.8 percent of the youngest; 8.7 percent of the oldest); the household was not poor in 1994 but poor in 1998 (8.2 percent of the youngest; 8.7 percent of the oldest).

It is clear from the discussion above that many children experience major life changes which could have the potential to alter developmental paths. It is also apparent that the likelihood of experiencing any particular transition varies considerably with the age of the child, with younger children generally being the most likely to experience important life transitions (an exception is that older children are much more likely to have a sibling leave the household).

In Table 4 we report upon a series of regressions which investigate the importance of transitions experienced between 1994 and 1996 or 1996 and 1998 for children's 1998 educational attainments, controlling for initial (i.e., 1994) characteristics of the child, pmk and household. To do this, we add transition variables to the 'base specification' (reported in Table 2 and discussed in the previous section). Regression results for our analysis of 'transitions' are reported in Table 4. To summarize before providing additional detail, we draw two major conclusions from this table: 1) transitions/shocks have relatively unimportant associations with children's educational attainments, after we control for starting places; 2) results reported earlier about which level characteristics matter most for children's educational outcomes are robust to the inclusion of the transition variables.

[16] These financial shock variables do not incorporate 1996. We did also experiment with income loss variables (e.g., income fell by 25 percent or increased by 25 percent) in addition to the 'change in poverty status' variables reported here. Controlling for initial income level, these variables were also rather unimportant.

Table 4
Select Determinants of Child Attainments by Cohort, Weighted Least Squares Estimates – Specifications Include All Shock Variables

	Children 4-6 PPVT	Children 8, 9 Math score	Children 11, 13 Math score	Children 11, 13 Reading score	Children 15 Math score	Children 15 Reading score
Child characteristics						
Has activity limitation	-13.10 (2.28)**	13.54 (0.48)	50.07 (1.27)	15.48 (1.42)	20.37 (0.26)	-29.88 (1.78)*
Is eldest	2.30 (2.11)**	7.66 (1.07)	29.26 (3.59)**	13.09 (3.44)**	0.33 (0.02)	-2.59 (0.49)
Number of siblings	-1.43 (2.65)**	5.21 (1.36)	12.17 (2.56)**	7.79 (3.81)**	-15.38 (1.78)*	-6.21 (1.92)*
PMK characteristics						
Lone Parent	0.72 (0.38)	-29.44 (2.92)**	5.27 (0.32)	2.09 (0.34)	6.40 (0.27)	-6.73 (0.80)
Age	0.33 (2.92)**	-0.02 (0.03)	0.84 (0.91)	0.63 (1.73)*	0.44 (0.29)	0.69 (1.20)
Did not complete high school	-4.13 (2.88)**	-28.71 (2.98)**	-15.28 (1.37)	-9.97 (1.65)*	4.46 (0.23)	-1.78 (0.18)
Obtained post-high school diploma	1.70 (1.79)*	0.80 (0.11)	-1.44 (0.16)	1.98 (0.53)	16.97 (0.91)	-2.60 (0.33)
Obtained university degree	4.53 (3.23)**	32.25 (3.23)**	36.43 (3.07)**	9.09 (2.06)**	66.58 (3.14)**	11.96 (1.96)*
Household characteristics						
Income (x 1000)	0.16 (2.16)**	-0.87 (1.20)	1.48 (1.11)	-0.18 (0.28)	-1.64 (0.71)	-0.32 (0.41)
Income squared (x 100000)	-0.00005 (0.90)	-0.0009 (1.11)	-0.001 (0.72)	0.0006 (0.69)	0.002 (0.79)	-0.00007 (0.07)
R2	0.19	0.40	0.40	0.26	0.33	0.26
Sample size	2227	622	611	612	314	314
F-score – all change variables	1.73**	2.49*	1.63*	2.89**	0.81	2.56**
F-score – all child change variables	1.77*	1.24	1.59	3.71**	0.63	1.70*
F-score – all PMK change variables	1.73	5.01**	2.38*	1.02	0.69	1.99
F-score – all change variables – specification includes ONLY change variables and an indicator of poverty in 94/98	5.80**	1.55*	2.41**	2.99**	1.11	2.63**

Notes:
1. Absolute value of t statistics in parentheses. 2. Standard errors calculated using Huber (1967) – White (1980) method. 3. *significant at the 10% level, ** significant at the 5% level. 4. Variables included but not reported are: child age, sex, quarter of birth, PMK suffers chronic illness, PMK lone parent status, PMK race, province of residence, lives in rural area. 5. Weights are from Statistics Canada (200x) to account for attrition and sample representativeness.

In more detail, then, Table 4 reports coefficients for selected key variables from the base specification, this time estimated after *all* transition variables have been added to the model. Models are again estimated separately for different age cohorts, and the estimation method is again weighted least squares. Basic conclusions noted above remain valid when we consider the set of regression results with all transition variables added. Thus, education level of the pmk is extremely important for educational outcomes of children and this holds for children at all ages studied. Family income continues to play a role, though the size of the effect is relatively small and the variable is in some cases statistically insignificant. The child's own health status matters for educational outcomes. Finally, number of siblings and birth order play a significant role, though living in a lone-mother family is usually statistically insignificant.

We do not report regression results for all transition variables included in our estimated models, because most of them are statistically insignificant. Of course, individual transitions were occasionally statistically significant, but few clear patterns emerge with which to tell a story about these results. (Perhaps the only exception is that having a pmk who completed a higher level of education between 1994 and 1996 or 1996 and 1998 is associated with better educational outcomes for children in 3 of 7 cases.)

We do, however, report a series of F statistics for tests of the hypotheses that *all* transition variables taken together add nothing to the estimated base model, that all child-related transitions add nothing to the model and that all pmk-related transitions add nothing to the model.[17] F-statistics for each of these variants are reported at the bottom of Table 4. In many cases, we cannot reject the hypothesis that the set of transitions adds nothing to the base specification. Not surprisingly, if we exclude all the level variables and include only the transitions, then the estimated importance of these variables is somewhat greater (see F statistics in bottom row of Table 4). However, the estimated model with 'only transitions' explains much less of over-all variation in observed current outcomes than the base model with 'only 1994 levels.' Of particular note, being poor in both 1994 and 1998 has a large negative association with child attainments when we do not control for the initial level of income.

To conclude, although we find that children, especially younger children, are quite likely to experience major life transitions, it does not appear, once we have controlled for starting places, that these transitions play a very important role in altering development paths.[18] A possible explanation for this finding is that since initial conditions play a very important role in determining the probability that a child will experience any given shock (e.g., a pmk with low education is more likely to become unemployed; a young pmk is more likely to move or divorce/remarry), it is hard to separate out the two.

---

17 We also experimented with including only child transitions and only pmk transitions.

18 Given this conclusion, we did not go on to pursue the question of which factors might help to mediate or reinforce the consequences of shocks in a child's life as we had originally planned.

7. The Impact of Earlier Attainments on Subsequent Attainments

In chapters 5 and 6, we considered the impacts of 'history' – as represented by characteristics at the time of cycle 1 – and transitions – as represented by shocks or changes between cycles 1 and 2, 2 and 3 and 1 and 3, on child attainments. In this chapter, we complement that analysis by considering another dimension of 'history', namely the impact of a child's earlier attainments on what she subsequently achieves. We have already had a glimpse of this in the analysis presented in chapter 4 where we saw that there were correlations between attainments measured in cycle 1 and attainments in cycle 3. However, while suggestive, such analysis is hardly conclusive because such correlations may be driven by some third factor that is not taken into account in that bivariate analysis.

In this chapter, we use a multivariate analysis of the relationship between earlier attainments, as measured in cycles 1 and 2, and attainments as measured in cycle 3. That is, we estimate relationships of the following form:

$$Y_{ij} = \alpha + \beta X_{ij} + \gamma Z_{ij} + u_{ij}$$

Where Y_{ij} is the attainment of child i, living in household j; α, β and γ are parameters to be estimated, X_{ij} is a vector of child, pmk and household characteristics, Z_{ij} is a measure of earlier attainments and u_{ij} is a disturbance term. Before continuing, however, we must note the following concern. Z_{ij} is not an exogenous variable; in fact it is an outcome determined by child, pmk and household characteristics in earlier periods as well as unobserved characteristics that are absorbed into u_{ij}. As a result, it is reasonable to expect that $E(Z_{ij}u_{ij}) \neq 0$. For example, children with a greater interest or aptitude for reading (an unobservable characteristics that is absorbed into u_{ij}) will, holding all other factors constant, have better attainments as measured both by Y_{ij} *and* Z_{ij}. Consequently, all parameter estimates are vulnerable to bias.

Our results are reported in Tables 5 and 6a to 6f For the purposes of comparison, the first column replicates the core findings of Table 2, which we call 'specification (1)'. Next, we report the results obtained by treating the earlier attainments as exogenous. In specification (2), we include the attainment measured in cycle 1. For children aged 8 and 9 in 1998 (i.e. children 4 or 5 at the time of cycle 1), this is the PPVT score. For all older children, the first stage attainment is the math score. In specification (3), the earlier attainment is the test score from cycle 2. In the case of reading, this is the reading score obtained during cycle 2 and in the case of mathematics, it is the math score obtained during cycle 2. An attraction of this approach is that we can compare the impact of other characteristics, such as pmk education and household income, once these earlier attainments are taken into account. We then introduce two changes to this specification. First, we treat earlier attainments as endogenous, using two different sets of instruments. The first set of instruments are the 'shocks' and 'transition' events that occur between

1994 and 1996. These are the development, or loss of a child's activity limitation, the acquisition or loss of full or step siblings, changing school or care giver, pmk's marrying, divorcing, moving or gaining an educational diploma and the household moving into or out of poverty. Second, we include as an additional set of regressors the same set of shock and transition variables, but for the period 1996 to 1998 (i.e. changes between cycles 2 and 3). Lastly, note that because we do not have an 'attainment' for children aged 0-2 at the time of the start of the NLSCY, we restrict our attention to the three older cohorts. Even with this, there are four new specifications for three cohorts, each with two attainments. Consequently, we describe our results thematically.

Table 5
PPVT Score for Children 4-6

	(1)	(2)
PMK characteristics		
Age	0.40 (4.07)**	0.41 (3.97)**
Did not complete high school	-3.86 (3.07)**	-3.75 (3.02)**
Obtained post-high school diploma	1.40 (1.59)	1.28 (1.44)
Obtained university degree	3.66 (2.95)**	3.26 (2.60)**
Lone Parent	1.13 (0.91)	1.24 (1.01)
Household characteristics		
Income (x 1000)	0.23 (3.87)**	0.23 (3.71)**
Income squared (x 100000)	-0.00009 (1.81)*	-0.00009 (1.58)
Earlier attainments		
Low birth weight		-4.57 (2.38)**
Mean, dependent varaible	98.6	

Notes:
(1) Specification reported in Table 2.
(2) Specification reported in Table 2 *plus* PPVT score (treated as being exogenous) from round 1.

Table 6a Math Score for Children 8, 9					
	(1)	(2)	(3)	(4)	(5)
PMK characteristics					
Lone parent	-18.14 (0.84)	-5.56 (0.64)	-22.89 (2.27)**	-16.19 (1.02)	-
Age	0.27 (0.51)	0.37 (0.64)	0.69 (0.82)	0.77 (0.63)	-
Did not complete high school	-16.00 (1.89)*	-13.37 (1.50)	-17.99 (1.45)	-21.95 (1.27)	-
Obtained post-high school diploma	-1.52 (0.25)	-4.57 (0.73)	-11.42 (1.23)	-2.48 (0.20)	-
Obtained university degree	30.59 (3.57)**	21.94 (2.43)**	34.51 (2.79)**	19.79 (1.17)	-
Household characteristics					
Income (x 1000)	-0.0003 (0.01)	-0.0007 (0.02)	-0.205 (0.45)	1.25 (1.02)	-
Income squared (x 100000)	-0.0002 (0.56)	-0.00001 (0.35)	-0.0002 (0.60)	-0.0003 (1.67)*	-
Earlier attainments					
PPVT from cycle 1	-	0.47 (2.20)**	-	-	-
Math score from cycle 2	-	-	0.38 (5.07)**	0.58 (2.92)**	0.72 (6.37)**
F test on instruments for earlier attainments	-	-	-	2.91**	3.75**
Mean, dependent variable	401.3				

Notes:
(1) Specification reported in Table 2; (2) Specification reported in Table 2 *plus* PPVT score (treated as being exogenous) from round 1.
(2) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 2.
(3) Specification reported in Table 2 *plus* math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996.
(4) Math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996 as well as child, PMK and household characteristics observed as of 1994.

Table 6b **Reading Score for Children 8, 9**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
PMK characteristics					
Lone parent	2.87 (0.63)	8.35 (1.92)*	-0.05 (0.01)	11.05 (0.96)	-
Age	0.09 (0.25)	0.21 (0.75)	0.18 (0.32)	0.34 (0.44)	-
Did not complete high school	-24.18 (4.08)**	-19.69 (3.63)**	-18.74 (1.54)	-5.51 (0.39)	-
Obtained post-high school diploma	2.18 (0.58)	-0.22 (0.06)	-4.36 (0.84)	0.55 (0.07)	-
Obtained university degree	11.32 (1.60)	-1.14 (0.21)	9.65 (1.24)	-0.006 (0.00)	-
Household characteristics					
Income (x 1000)	0.52 (1.96)**	0.44 (2.00)**	-0.205 (0.63)	1.31 (1.34)	-
Income squared (x 100000)	-0.00025 (1.07)	-0.0008 (0.45)	-0.0002 (0.68)	-0.0002 (1.14)	-
Earlier attainments					
PPVT from cycle 1	-	0.84 (6.04)**	-	-	-
Reading score from cycle 2	-	-	0.62 (6.79)**	1.02 (3.79)**	0.92 (5.43)**
F test on instruments for earlier attainments	-	-	-	4.63**	5.75**
Mean, dependent variable	223.9				

Notes:

(1) Specification reported in Table 2.
(2) Specification reported in Table 2 *plus* PPVT score (treated as being exogenous) from round 1.
(3) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 2.
(4) Specification reported in Table 2 *plus* math score (treated as being endogenous) from round 2 *plus* child, PMK and household ‘shocks’ experienced between 1996 and 1998. Instruments are child, PMK and household ‘shocks’ experienced between 1994 and 1996.
(5) Math score (treated as being endogenous) from round 2 *plus* child, PMK and household ‘shocks’ experienced between 1996 and 1998. Instruments are child, PMK and household ‘shocks’ experienced between 1994 and 1996 as well as child, PMK and household characteristics observed as of 1994.

Table 6c Math Score for Children 11, 13					
	(1)	(2)	(3)	(4)	(5)
PMK characteristics					
Lone parent	-6.52 (0.56)	21.31 (1.29)	12.39 (1.15)	5.12 (0.39)	-
Age	1.41 (1.67)*	2.64 (2.94)**	0.47 (0.62)	0.30 (0.33)	-
Did not complete high school	-14.12 (1.42)	0.84 (0.06)	-2.39 (0.21)	-0.88 (0.07)	-
Obtained post-high school diploma	-4.22 (0.53)	-1.82 (0.16)	8.27 (1.20)	4.20 (0.46)	-
Obtained university degree	23.48 (2.04)**	27.89 (1.75)*	43.08 (3.50)**	41.74 (3.20)**	-
Household characteristics					
Income (x 1000)	-1.08 (1.06)	0.78 (0.65)	1.61 (1.81)*	2.02 (1.28)	-
Income squared (x 100000)	-0.0007 (0.44)	-0.0005 (0.22)	-0.002 (1.47)	-0.002 (1.37)	-
Earlier attainments					
Math score from cycle 1	-	0.43 (5.69)**	-	-	-
Math score from cycle 2	-	-	0.54 (8.28)**	0.30 (0.82)	0.82 (7.00)**
F test on instruments for earlier attainments	-	-	-	1.59*	8.73**
Mean, dependent variable	518.5				

Notes:
(1) Specification reported in Table 2.
(2) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 1.
(3) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 2.
(4) Specification reported in Table 2 *plus* math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996.
(5) Math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996 as well as child, PMK and household characteristics observed as of 1994.

Table 6d Reading Score for Children 11, 13					
PMK characteristics					
Lone parent	0.71 (0.18)	-0.20 (0.03)	2.86 (0.85)	5.05 (1.03)	-
Age	0.58 (1.85)*	1.04 (2.49)**	0.03 (0.13)	0.18 (0.50)	-
Did not complete high school	-7.95 (1.67)*	-7.12 (1.04)	3.04 (0.59)	-1.84 (0.28)	-
Obtained post-high school diploma	-0.77 (0.24)	1.52 (0.29)	1.77 (0.56)	3.71 (0.96)	-
Obtained university degree	8.40 (2.02)**	4.76 (0.84)	9.57 (2.32)**	9.61 (2.32)**	-
Household characteristics					
Income (x 1000)	0.37 (1.00)	-0.44 (0.76)	0.42 (1.10)	0.64 (1.28)	-
Income squared (x 100000)	-0.00012 (0.22)	0.001 (1.28)	-0.0005 (0.99)	-0.0006 (1.02)	-
Earlier attainments					
Math score from cycle 1	-	0.14 (3.77)**	-	-	-
Reading score from cycle 2	-	-	0.51 (10.85)**	0.20 (1.39)	0.43 (5.52)**
F test on instruments for earlier attainments	-	-	-	2.48**	10.99**
Mean, dependent variable	268.6				

Notes:
(1) Specification reported in Table 2.
(2) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 1.
(3) Specification reported in Table 2 *plus* reading score (treated as being exogenous) from round 2.
(4) Specification reported in Table 2 *plus* reading score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996.
(5) Reading score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996 as well as child, PMK and household characteristics observed as of 1994.

Table 6e Math Score for Children 15					
	(1)	(2)	(3)	(4)	(5)
PMK characteristics					
Lone parent	-12.99 (0.67)	-6.95 (0.43)	2.26 (0.15)	-11.81 (0.52)	-
Age	2.09 (1.71)*	0.49 (0.35)	0.38 (0.35)	1.72 (1.13)	-
Did not complete high school	6.04 (0.37)	-18.16 (1.07)	19.71 (1.42)	23.31 (1.22)	-
Obtained post-high school diploma	14.77 (1.01)	2.04 (0.15)	16.58 (1.01)	20.47 (1.13)	-
Obtained university degree	63.64 (3.52)**	30.82 (1.60)	34.53 (1.76)*	59.88 (2.61)**	-
Household characteristics					
Income (x 1000)	2.95 (1.99)**	2.17 (1.57)	0.29 (0.22)	1.28 (0.74)	-
Income squared (x 100000)	-0.0039 (2.10)**	-0.003 (1.76)*	-0.0004 (0.24)	-0.0013 (0.57)	-
Earlier attainments					
Math score from cycle 1	-	0.62 (6.94)**	-	-	-
Math score from cycle 2	-	-	0.74 (12.77)**	0.13 (0.48)	0.86 (5.67)**
F test on instruments for earlier attainments	-	-	-	2.48**	4.17**
Mean, dependent variable	631.7				

Notes:
(1) Specification reported in Table 2.
(2) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 1.
(3) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 2.
(4) Specification reported in Table 2 *plus* math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996.
(5) Math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996 as well as child, PMK and household characteristics observed as of 1994.

Table 6f Reading Score for Children 15					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
PMK characteristics					
Lone parent	-3.69 (0.56)	1.11 (0.14)	-0.00 (0.00)	-2.68 (0.36)	-
Age	1.14 (2.39)**	1.18 (2.44)**	0.35 (0.68)	0.61 (0.96)	-
Did not complete high school	0.99 (0.12)	-8.26 (0.63)	2.76 (0.28)	2.39 (0.21)	-
Obtained post-high school diploma	1.73 (0.26)	0.65 (0.11)	-2.63 (0.45)	-3.51 (0.52)	-
Obtained university degree	8.08 (1.44)	-2.79 (0.49)	-6.94 (1.25)	-0.16 (0.02)	-
Household characteristics					
Income (x 1000)	1.02 (1.70)*	0.96 (1.69)*	0.46 (0.70)	-0.29 (0.49)	-
Income squared (x 100000)	-0.0016 (2.02)**	-0.002 (2.14)**	-0.0006 (0.60)	0.0002 (0.18)	-
Earlier attainments					
Math score from cycle 1	-	0.13 (3.33)**	-	-	-
Reading score from cycle 2	-	-	0.57 (7.88)**	0.22 (0.81)	0.53 (2.83)**
F test on instruments for earlier attainments	-	-	-	1.12	4.56**
Mean, dependent variable	289.2				

Notes:
(1) Specification reported in Table 2.
(2) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 1.
(3) Specification reported in Table 2 *plus* math score (treated as being exogenous) from round 2.
(4) Specification reported in Table 2 *plus* math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996.
(5) Math score (treated as being endogenous) from round 2 *plus* child, PMK and household 'shocks' experienced between 1996 and 1998. Instruments are child, PMK and household 'shocks' experienced between 1994 and 1996 as well as child, PMK and household characteristics observed as of 1994.

One area of interest is whether pmk and household characteristics have any additional impact over and above any relationship that they have to an attainment at an earlier period. To see whether this is the case, we compare the results of specification (1) – where earlier attainments are excluded – with the results of specifications (2) through (4). The general pattern across all six tables is the following. A number of pmk and household characteristics, as measured during cycle 1, have an impact on attainments as measured in cycle 3 when an earlier measure of attainment is *excluded*. However, once we include the measure of earlier attainment, the impact of these characteristics begins to diminish. Most strikingly, when we compare specification (4) – where we treat earlier attainments as endogenous - with that of (1), the impact of earlier pmk and household characteristics all but disappears. The only meaningful impact that persists is the cases where the pmk has obtained a university degree, which continues to raise math scores where children are 11, 13 or 15. Although the quadratic on income in Table 6a and possession of a university degree in table 6d are both statistically significant, the magnitude of both effects is, at less than 3 per cent, trivial.

Next, we consider the impact of attainments measured four years previously on subsequent attainments, controlling for the same set of child, pmk and household characteristics described in chapter 5. However, in the absence of any viable instruments, we treat these earlier attainments as exogenous. In all six specifications, we find a statistically significant relationship. Further, the magnitudes of some of these associations are large. As an example, consider Table 6e which reports the impact of the math score obtained when the child was 11 on her math score at age 15. To see this, first note that the coefficient on the earlier math score is 0.62 meaning that for every additional point obtained on the math attainment in 1994 was associated with a 0.62 point increase on the score in 1998. Next, recall that in Table 3, we simulated the impact of various changes in pmk and household characteristics on attainments. Consider the case of two children who are otherwise identical in terms of observable characteristics save the education of their pmk. Table 3 tells us that the child whose pmk did not complete high school would have a math score at age 11 that was 2.8% or 14 points below the base. The child whose pmk has a university degree would have a math score at age 11 that was 4.7% or 24 points above the base. Thus, these children are predicted to have, at age 11, a difference in math scores of 38 points. At age 15, the difference in their attainments – assuming that they are exposed to exactly the same shocks and transition events – is due to both the direct effect of pmk education (the coefficients –18.16 and 30.82 respectively) plus the impact of this pmk characteristic on the earlier math score. The difference in their predicted math scores is the sum of the absolute values of the pmk education characteristics (18.16 + 30.82) *plus* the difference in predicted math score in cycle 1 – 38 times 0.62. Collectively, this produces a difference in predicted attainments at age 15 of 72.5 points, a 13 per cent difference relative to the mean.

Also instructive is the comparison between the results for prior attainments, as measured in cycle 1, with those obtained in cycle 2. Because of problems associated with the reading test in cycle 1, we can only do a 'like for like' comparison for the math scores for children aged 11 and13 and 15 in 1998. In both cases, as is shown by comparing specifications (2) and (3) in Tables 6c and 6e, the coefficients on the math score are higher when we use the more recent prior attainment.

Our next step is to consider the potential impact of the endogeneity of prior attainments. Our first attempt at addressing this is found in specification (4). Recall we use the representations of shocks and transitions between 1994 and 1996 described above as instruments. We want to compare these results with those from specification 3. Doing so across all six tables produces an ambiguous set of results. In two cases - math score for children 8, 9; reading score for children 8, 9 – the effect on instrumenting is to increase the size of the parameter estimate on the earlier attainment. But in the other four cases, instrumenting causes the parameter estimate to fall by considerable magnitudes and to become statistically insignificant. It would seem, therefore, that for older children there is no impact of earlier attainments on subsequent attainments once the endogeneity of the former is taken into account.

Such a conclusion, however, is premature. Bound, Jaeger and Baker (1995) show that when instruments have poor statistical power, that the use of two stage least squares is unlikely to correct for endogeneity bias and, even more grieviously, lead to parameter estimates that are downwardly biased. One way of examining the explanatory power of these instruments is to note the results of a F test on their joint significance in the first stage regression that predicts the endogenous variable. These are reported in the last row of Tables 6a through 6f. Echoing the findings reported earlier about the relative unimportance of transition variables, these F-statistics are all remarkable small, well below the target value of 8 to 10 suggested by Bound, Jaeger and Baker. Rectifying this problem with the instruments requires that we find additional variables with explanatory power in the first stage regression (i.e. variables that determine the earlier attainment) that can be plausibly excluded from our second stage regression (i.e. do not affect the later attainment). Recall that we have already noted that pmk and household characteristics, as measured in 1994, have little impact on attainments in 1998, once a measure of earlier attainments is taken into account. This suggests that these characteristics might also serve as plausible instruments.

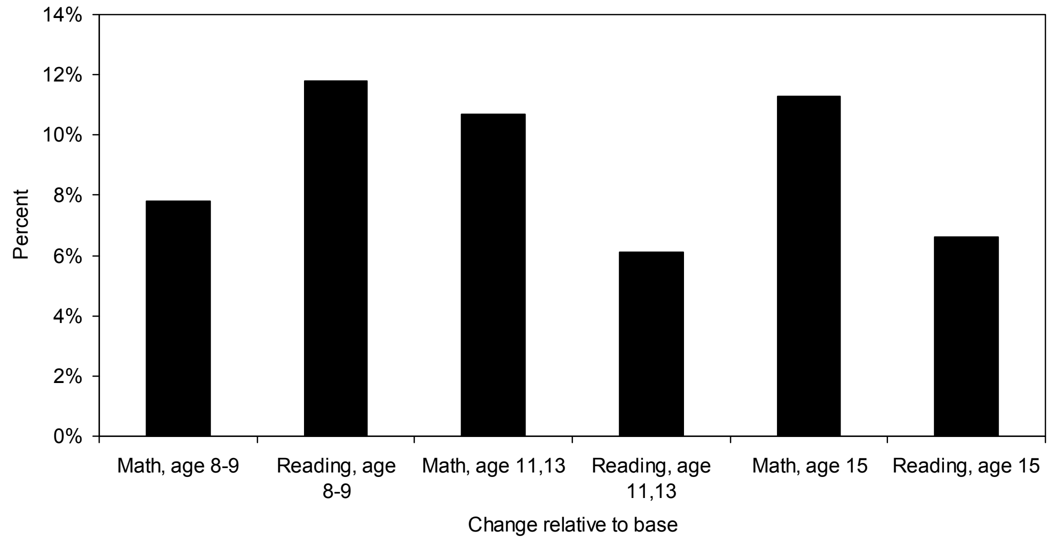
Accordingly, Tables 6a to 6f, contain the results of a second attempt at estimating a two stage least squares regression (what we will call specification (5)). The attainment measured in cycle 2 is treated as endogenous with child, pmk and household shocks experienced between 1994 and 1996 as well as child, pmk and household characteristics observed as of 1994 used as instruments. Additional determinants of attainments in cycle 3 are the child, pmk and household shocks and transitions that occur between 1996 and 1998. There are four findings across these six tables that should be noted. First, the F test on instruments rises in value, suggesting that we are doing a better job of predicting the endogenous variable (although these values are not always as high as we might like). Second, with this improved set of instruments, we once again find a statistically significant relationship between earlier and later attainments. Third, in four cases – all math attainments as well as the reading score for children aged 8, 9 in 1998 – the parameter estimate rises in value compared to specification (3) where the same measure of earlier attainment is treated as being exogenous. In the other two cases, the parameter estimates are basically the same. This indicates that failing to account for the endogeneity of earlier attainments leads us to underestimate their correlation across time. Fourth, the magnitudes of these associations generally tend to be larger for math scores than for reading scores.

Another way to understand the importance of past attainments (i.e., 'history') for current attainments is to use our estimated regression models to calculate the change in current attainments which would be predicted for some given change in past attainments. Results of such an exercise are reported in Table 7, which illustrates for each specification reported in Table 5 and Tables 6a-6f (and for each cohort of children), the predicted consequences of increasing earlier test scores by 15 percent. (The exception is for the 'baby cohort' for whom we can only predict the implications of having been a low-birth-weight baby versus not.)

Table 7 Simulating the Impact of Changes in Prior Attainments (percent change from the base)							
	Age 4-6	**Age 8-9**		**Age 11, 13**		**Age 15**	
	PPVT	**Math**	**Reading**	**Math**	**Reading**	**Math**	**Reading**
Child was low birth weight	-4.7%	-	-	-	-	-	-
Increase cycle 1 attainment	-	1.8%	5.6%	4.9%	2.9%	8.3%	3.3%
Increase cycle 2 attainment	-	4.1%	8.3%	7.3%	7.1%	10.4%	8.5%
Increase cycle 2 attainment (instrumental variable spec)	-	7.8%	11.8%	10.7%	6.1%	11.3%	6.6%
Notes: This simulation uses specifications (2) in table 5 and (2), (3) and (5) in tables 6a-6f to predict attainments. In each case the simulation is to increase the prior attainment by 15%. For the age 4-6 group, the simulation was to make the base case child a low birth weight baby.							

While results for all estimated models are reported in Tables 6a-6f, we focus the discussion on simulation results corresponding with Specification 5 since we have argued in the previous section that this is our preferred model. Consider, first, the cohort of children who were aged 4 or 5 in 1994 and hence 8 or 9 in 1998. Figure 13 illustrates that for a 'base child' (i.e. all change variables set to 0 and the continuous variable, the previous attainment, set at the mean), a 15 percent increase in 1996 math score causes a 7.8 percent increase in 1998 math score; a 15 percent increase in 1996 reading score causes an 11.8 percent increase in 1998 reading score. For children aged 9 and 11 in 1998, a 15 percent increase in 1996 math score is associated with a 10.7 percent increase in 1998 math score; a 15 percent increase in reading score is associated with a 6.1 percent increase in 1998 reading score. Finally, for our oldest cohort of children aged 15 in 1998, a 15 percent increase in 1996 math score is associated with an 11.3 percent increase in 1998 math score; a 15 percent increase in 1996 reading score is associated with a 6.6 percent increase in 1998 reading score.

Figure 13
Percent Change in Attainments When Cycle 2 Attainment Increased by 15%



However, even these simulations do not entirely depict the full extent to which ‘history’ can matter for child attainments. To further illustrate the possible ‘snow-balling’ of past events into current outcomes, consider the following thought experiment which compares the development of two otherwise identical boys who were both 7 years old in 1994. Suppose the first boy is our average ‘base case’ (i.e., he is a Caucasian boy, born in the first quarter of the year, living in urban Ontario with a pmk of average age who has completed high school ; family income is the average for this sample). The other is ‘disadvantaged’ insofar as he comes from a family with income at the 25th percentile, his mother is 10 years younger than the average and has less than a high-school education. Using our base model for the ‘middle cohort’ (i.e., that reported in Table 2), we would predict that by the time both children are 11 years old, the disadvantaged child will have a math score which is 6.7 percent lower (469 versus 502.6). (We have already illustrated this point in Figure 11). Now, if assume that the disadvantaged child remains disadvantaged relative to the base, and *as well*, has by age 11 less developed math skills, then both factors will contribute to limiting further development to age 15. Suppose we go on to predict for the now 11-year old children, the difference in their expected math scores by age 15 (using Specification 2 from Table 5[19]). Our calculations suggest that the difference in math scores would increase to 9.3 percent, of which 5.4 percent could be attributed to the direct effect of the continued disadvantage (ie., lower income and younger, less well-educated pmk) and an additional 3.9 percent could be attributed to the indirect consequences of that disadvantage through the persistence of problems accumulated from the past – the ‘snowball’ effect.

---

19 We now plug in the lower math score predicted at age 11 and continue to assume the same elements of disadvantage. Notice that this procedure assumes that the regression estimates would be stable across cohorts and over time.

We have reviewed many results in this chapter, so it is helpful to conclude by summarizing our principal findings. History *is* destiny. Earlier attainments are causally associated– both in terms of statistical significance as well as magnitude – with subsequent attainments. This causal relationship tends to be larger when we account – albeit imperfectly – for the endogeneity of these earlier attainments and also tend to be larger for the math scores. pmk and household characteristics have little impact on subsequent attainments over and above their impact on earlier attainments.

8. Conclusion

The title of this paper asks the question: "Is History Destiny?". Do past levels of resources, contexts and opportunity structures carry long-term consequences for subsequent child attainments? The conclusion from this study suggests that the answer is yes. Using three cycles of the National Longitudinal Survey of Children and Youth, fielded in 1994, 1996 and 1998 to examine three inter-related issues: the impact of past levels of resources, contexts and opportunity structures on subsequent school readiness and education attainments; whether shocks, or transition events, alter the path of these outcomes in a positive or negative fashion; and what role is played, and can be played, by policy interventions in improving these attainments. Resources include characteristics of the person most knowledgeable of the child (most often the mother) such as age, education and health status, characteristics of the household such as income levels, neighbourhood characteristics and location. Transition events include the development, or loss of a child's activity limitation, the acquisition or loss of full or step siblings, changing school or care giver, marriage, divorce and the household moving into or out of poverty.

Age and education of the person with greatest contact with the child, and hence most knowledgeable, as measured in 1994 has a long-term effect on school readiness, as measured by the PPVT, and tests of mathematics and reading ability measured four years later. Household income in 1994 also affects these outcomes in 1998, although the magnitude of this effect is small. The cumulative effect of these characteristics is large. For example, a 'disadvantaged' child – one with a young mother who has not completed high school living in a household at the 25th percentile of income in 1994 – obtains scores on the mathematics test in 1998 that are 13 to 22 per cent lower than an 'advantaged child – one with an older mother possessing a university degree living in a household at the 75th percentile of income in 1994. By contrast, apart from presence of an activity limitation, observable child and neighbourhood characteristics, as well as other parental characteristics – including marital status - have little systematic impact on these attainments. Strikingly, none of the observed transition events have a substantive effect on these outcomes.

There is evidence of persistence in attainments over time. Cross-tabulations indicate that children in the lowest – or highest quintile of attainments – as of 1994 are highly likely to remain in that group when measured again in 1998. In the context of multivariate analysis, this persistence across time is robust to the inclusion of child, parental and household characteristics as well as estimation techniques that account for the endogeneity of these attainments.

Appendix A

Examining the potential impact of attrition bias

In the NLSCY, attrition may arise for two reasons. The first relates to the possibility that some children will have not been traced because the household moved, or the household was located but refused, or was unable to continue in the survey. The second source of bias relates to the absence of full information on certain attainments, correlates or predictor variables. To see why this attrition may be a source of concern, consider the "middle cohort", children aged 7 or 9 in 1994. In the first cycle, 69% of these children lived in homes owned by their parents. However, of the children successfully re-interviewed in cycle 3, 77% of these children lived in homes owned by their parents in 1994. If attrition were purely random, this proportion would remain unchanged between cycles 1 and 3. However, it would appear that children in rented accommodation – or other forms of housing not owned by residents themselves – are less likely to be re-interviewed, presumably because such families are more mobile. Suppose that, holding other factors constant, children in families that are mobile experience disruptions to learning. Failing to take this into account could mean that any subsequent analysis will be based on a sample that omits those children who have been more likely to experience disruptions to learning.

In order to understand the magnitude and likely impact of such potential biases, it is necessary to understand how the NLSCY operates, the magnitude of attrition and the mechanisms already in existence to address this bias.

The NLSCY is designed to follow a representative sample of Canadian children from infancy to adulthood. The first cycle of data collection was carried out in the winter and spring of 1994/95. As Table 1 indicates, this produced a sample of 22,831 children. As is clear from comparing the proportion of respondents in the "master file" with the proportion of children under 15 in the 1996 census, sample proportions by province do not match the proportion of children actually living in each province. Smaller provinces are over-represented in the sample and larger provinces under-represented. This reflects several deliberate decisions; most notably the need for a sufficiently large sample by province so as to produce reliable estimates and the request by the Government of New Brunswick for a supplementary sample (in cycle 2, Statistics Canada, 1999).

The second cycle was implemented in the winter and spring of 1996/97. The third cycle was fielded in the fall of 1998 to June 1999. Sample size falls between cycles 1 and 3 for several reasons. First, cost considerations led to the decision to drop approximately 5000 children (Knighton, Claveau, Laflamme and Michaud, 2000; Statistics Canada, 2001). Second, in cycle 1, up to four children were selected per household. In order to reduce respondent burden, this was reduced to a maximum of two children (Statistics Canada, 2001). Additionally, respondents are allowed to determine whether their information is to be held only by Statistics Canada or made available to other government departments and this sub-sample is called the "shared file."

All these features make examining attrition in the NLSCY rather tricky. Mindful of this, Table A1 brings out a number of important features. First, as a consequence of both purposive sampling decisions and sample attrition, there have been significant reductions in the number of children for whom data are available. Just under two-thirds (63.8%) of the children who appear in the shared file for cycle 1 also appear in cycle 3. Second, there are marked provincial variations in attrition, ranging from a loss of 44.8 per cent of the original sample in Prince Edward Island to 31.1 per cent in Quebec. Third, the ratios of attrition reported in the master and shared files differ. On average, attrition is six per cent higher in the shared file than it is in the master file. However, there are considerable variations in this ratio across provinces, with attrition in the shared file for Newfoundland, Nova Scotia, New Brunswick and Saskatchewan being considerably higher than this average.

Table A1
Sample Attrition Between Cycle 1 and Cycle 3 by Province

Province	Number of children responding, cycle 1, master file	Number of children responding, cycle 1, shared file	Proportion of sample, cycle 1, master file	Proportion of sample, cycle 1, shared file	Proportion of children under 15 from 1996 census	Number of children interviewed for cycle 3, master file	Number of children interviewed for cycle 3, shared file	Proportion of sample remaining in cycle 3, master file	Proportion of sample remaining cycle 3, shared file	Proportion of children attrited by cycle 3, master file	Proportion of children attrited by cycle 3, shared file	Ratio of attrition in shared file to attrition in master file
Newfoundland	1,232	1,176	5.4%	5.5	1.9	846	729	5.6%	5.3	31.3%	38.0%	1.21
Prince Edward Island	764	735	3.4	3.4	0.5	434	406	2.9	3.0	43.2	44.8	1.04
Nova Scotia	1,532	1,437	6.7	6.7	3.1	1,085	960	7.2	7.0	29.2	33.2	1.14
New Brunswick	1,426	1,327	6.3	6.2	2.5	958	818	6.4	6.0	32.8	38.4	1.17
Quebec	4,065	3,859	17.8	18.0	23.4	2,845	2,658	18.9	19.4	30.0	31.1	1.04
Ontario	6,020	5,618	26.4	26.2	37.7	3,762	3,455	25.1	25.4	37.5	38.5	1.03
Manitoba	1,789	1,673	7.8	7.8	4.2	1,114	989	7.4	7.2	37.7	40.9	1.08
Saskatchewan	1,878	1,782	8.2	8.3	3.9	1,257	1,130	8.4	8.3	33.1	36.6	1.11
Alberta	2,185	2,069	9.5	9.6	10.5	1,420	1,392	9.5	10.2	35.0	32.7	0.93
British Columbia	1,940	1,779	8.5	8.3	12.5	1,284	1,150	8.6	8.4	33.8	35.4	1.04
Total	22,831	21,455				15,005	13,687			34.3	36.2	1.06

Notes:
1. Results from NLSCY “master file” are drawn from Statistics Canada (2000, 2001).
2. Results from 1996 census data are taken from Statistics Canada (1998).
3. Remaining results are derived from NLSCY “shared file”.

Documentation on the NLSCY made available by Statistics Canada in the fall of 2001 indicates that weights are available to take these factors into account. "The NLSCY weighting strategy is based on a series of cascaded adjustments applied to a basic (or initial) weight. Conceptually, the basic weight of each child is approximately equal to the inverse of the child's probability of selection" (Statistics Canada, 2001, p. 48). A second adjustment factor accounts for non-response since initial sample selection. A third adjustment is based on the homogeneous response group (HRG) method. This "involves an attempt to consolidate those individuals with the same propensity to respond. These groups are formed using the characteristics for each child reported in Cycle 1" (Statistics Canada, 2001, p. 49) and from this an additional correction factor is generated. Lastly, a further weighting is applied to ensure consistency between survey estimates and known demographic characteristics by province, sex and age. Although not stated in Statistics Canada (2001), a further weighting factor is applied to take account of differential rates of attrition observed in the master and shared files.

Given the availability of these weights, the key question is whether they eliminate all meaningful attrition bias. To examine this, we adopt the following approach. We observe the sample of children interviewed in cycle 1. Subsequently, they sub-divide into two groups: those who attrite and those who remain in the sample. As Fitzgerald, Gottschalk and Moffitt (1998) and Alderman et al (2000), explain if the weighting schemes described above eliminates all sources of attrition bias, there should be no statistically significant difference in the characteristics of children in both groups. Tables A2, A3 and A4 explore this issue in the context of province of residence, household and child characteristics respectively.

Table A2
Province of Residence for Children Appearing Only in Cycle 1 and Children Appearing in Cycles 1 and 3

Province	Proportion of children under 15 from 1996 census	Proportion of children in cycle 1 only, shared file	Proportion of children in cycles 1 and 3, shared file	T statistic on differences in proportions
Newfoundland	1.9	1.7	1.9	0.97
Prince Edward Island	0.5	0.6	0.5	0.54
Nova Scotia	3.1	2.9	3.1	0.48
New Brunswick	2.5	2.5	2.5	0.22
Quebec	23.4	18.4	23.8	8.63**
Ontario	37.7	42.6	38.0	6.30**
Manitoba	4.2	4.1	3.9	0.73
Saskatchewan	3.9	3.6	3.7	0.19
Alberta	10.5	10.1	10.6	1.13
British Columbia	12.5	13.5	12.1	2.67**

Notes:
1. Results from 1996 census data are taken from Statistics Canada (1998).
2. Remaining results are derived from NLSCY "shared file".
3. Proportion of children in cycle 1 only is weighted by cycle 1 cross-sectional weights.
4. Proportion of children in cycle 1 and 3 is weighted by cycle 3 longitudinal weights.
5. ** Statistically significant at the 95% confidence level.

Table A3 **Household Characteristics Measured in Cycle 1, for Children Appearing Only in Cycle 1 and Children Appearing in Cycles 1 and 3**			
Household characteristic	**Mean value for children appearing in cycle 1 only, shared file**	**Mean value for children appearing in cycles 1 and 3, shared file**	**T statistic on differences in proportions**
Lone parent	17.5%	15.3%	4.07**
PMK female	87.7%	92.5%	11.16**
PMK < high school	18.0%	16.4%	2.84**
PMK has high school	44.8%	47.0%	2.87**
PMK has certificate/diploma	19.8%	20.6%	1.35
PMK has university degree	17.4%	16.0%	2.46**
PMK in labour force in past 12 months	67.3%	70.6%	4.83**
PMK has had no occupation in past 12 months	32.8%	29.4%	4.93**
Equivalent household income	17549	18235	3.64**
Household is poor	30.7%	26.4%	6.41**
Total number of siblings	1.5	1.3	13.09**
Household size	4.5	4.3	9.20**
Home owned by residents	68.8	71.9	4.50**

Notes:
1. Results are derived from NLSCY “shared file”.
2. Proportion of children in cycle 1 only is weighted by cycle 1 cross-sectional weights.
3. Proportion of children in cycle 1 and 3 is weighted by cycle 3 longitudinal weights.
4. ** Statistically significant at the 95% confidence level.

Table A4 **Child Characteristics Measured in Cycle 1, for Children Appearing Only in Cycle 1 and Children Appearing in Cycles 1 and 3**			
Child characteristic	**Mean value for children appearing in cycle 1 only, shared file**	**Mean value for children appearing in cycles 1 and 3, shared file**	**T statistic on differences in proportions**
All children			
Age	5.8	5.6	4.02**
Child is female	48.9%	48.7%	0.27
Children 0-2			
Birth weight	3400	3392	0.45
Child had low birth weight	4.6%	6.3%	2.45**
Child is limited in activity	3.1%	2.4%	1.48
Child was breastfed	74.9%	75.2%	0.19
Children 4-5			
PPVT score (standardized)	98.2	99.6	2.35**
Children 7-9			
Math score (standardized)	374.9	373.4	0.33
Children 11			
Math score (standardized)	492.2	505.7	2.72**

Notes:
1. Results are derived from NLSCY “shared file”.
2. Proportion of children in cycle 1 only is weighted by cycle 1 cross-sectional weights.
3. Proportion of children in cycle 1 and 3 is weighted by cycle 3 longitudinal weights.
4. ** Statistically significant at the 95% confidence level.

Table A2 explores the distribution of children across provinces. Data found in the column labeled "proportion of children in cycle 1 only" are weighted on the cross-sectional weights applied to the cycle 1 data. Data found in the column labeled "proportion of children in cycles 1 and 3" are weighted using the method described above. For reference, also included are the proportions of children under the age of 15 as reported in the 1996 census (Statistics Canada, 1999). In the four Atlantic Provinces, all three proportions are approximately equal as are the proportions for the three Prairie Provinces. However, even with these weighting adjustments, significant differences are found in Ontario and Quebec. A statistically significant difference is also found for British Columbia, though the magnitude of the difference is small. In particular, children who have attrited from the shared file after cycle 1 appear to come disproportionately from Ontario.

Table A3 reports selected household characteristics for children appearing only in cycle 1 and those appearing in both cycles 1 and 3. Even after weighting these data, children are less likely to remain in the NLSCY if they do reside in a lone-parent household, if the person most knowledgeable (pmk) about them is male, if the pmk has not completed high school, or has a university degree, if the pmk has had no occupation in the past 12 months and if they come from a household considered poor. They are more likely to remain if the pmk has a high school diploma, if the pmk has been in the labour force in the past 12 months, and if the home is owned by its residents. Children are less likely to remain in the survey if they come from larger families or have more siblings. They are more likely to remain in households with higher incomes, adjusted using standard OECD equivalence scales.

Table A4 examines whether attainments observed in cycle 1 differ between those children who drop out of the NLSCY and those who remain. As before, the figures reported in this table are based on weighted data. For children aged 0-2 years, those who had low birth weights (below 2500g) were slightly more likely to remain in the sample. Children aged 4-5 years and at age 11 are more likely to remain in the sample if, on average, they obtained higher scores on the PPVT and math test respectively. However, while these differences are statistically significant, the magnitudes of these differences are quite small.

Given that we continue to observe some differences in characteristics despite the application of these weights, it is useful to explore the extent and magnitude of this possible bias further. We again following the methodology laid out in Fitzgerald, Gottschalk and Moffitt (1998) and Alderman et al. (2000), estimating a probit regression. The dependent variable equals one if the child is present in both cycles 1 and 3, 0 if the child is only present in cycle 1. We are particularly concerned as to whether initial attainments, such as scores on the first round of standardized tests, are associated with differential probabilities of attrition.

Selected results of these probits are reported in Table A5. We focus on two sets of regressors that may affect attrition, initial attainments and household income. These regressions control for a wide variety of other household characteristics, including location, and characteristics of the person most knowledgeable about the child (age, sex, education, occupation). Strikingly, unlike the univariate results reported above, in a multivariate setting none of these characteristics has a consistent impact on the likelihood of attrition. While the comparison of mean incomes suggested that children from poorer households were less likely to remain in the survey, this does not carry over to the multivariate results. For older children, initial attainments have no impact on the likelihood that the child remains in the sample, but this is not true for younger children.

Table A5
Selected Results from Probits on Likelihood of Attrition

	Children 0-2	Children 4-5	Children 7, 9	Children 11
Low birth weight	0.238 (0.092)**			
Standardized PPVT score		0.003 (0.0015)**		
Standardized math score			-0.0004 (0.0006)	0.001 (0.0008)
Household equivalent income (‘000s of dollars)	0.006 (0.003)	-0.006 (0.000)	0.003 (0.007)	-0.005 (0.00001)
Household equivalent income squared (‘000s of dollars)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)

Notes:
1. Results are derived from NLSCY “shared file”.
2. Proportion of children in cycle 1 only is weighted by cycle 1 cross-sectional weights.
3. Proportion of children in cycle 1 and 3 is weighted by cycle 3 longitudinal weights.
4. ** Statistically significant at the 95% confidence level.
5. Regressors included but not reported are: province of residence, characteristics of the PMK (age, sex, education and occupation), and selected household characteristics (number of siblings, whether home is owned).

Hence, our last step is to explore the magnitude on attrition of these initial attainments. To do so, we set out a “base case” for each age group in which we predict the probability of remaining in the cycle based on the coefficients generated by the probits, the mean values for all continuous variables and setting all dummy variables equal to zero. We then re-estimate these probabilities after increasing each attainment by 10 per cent. This produces the following results:

	Probability of remaining in NLSCY
Children aged 0-2	
Base case probability	0.652
Child had low birth weight	0.736
Children aged 4-5	
Base case probability	0.657
Increase mean PPVT score by 10%	0.668
(Note: Coefficient not statistically significant at 95% confidence level.)	
Children aged 7, 9	
Base case probability	0.597
Increase mean math score by 10%	0.591
(Note: Coefficient not statistically significant at 95% confidence level.)	
Children aged 11	
Base case probability	0.500
Increase mean math score by 10%	0.527
(Note: Coefficient not statistically significant at 95% confidence level.)	

From these results, we conclude:

- There has been a considerable reduction in the size of sample available for analysis between cycles 1 and 3. On average, less than two-thirds of the original sample remains.
- The rate of attrition in the shared file is higher than in the master file.
- Even after applying weights that adjust for the sampling distribution of the NLSCY and rates of attrition by group characteristics, significant differences persistent between the characteristics of children, and their families, who remain in the NLSCY and those who are only found in the first cycle.
- Multivariate analysis, however, indicates that these differences are either not statistically significant or of minimal magnitude. The only marked difference that remains is that children with low birth weight have a significantly higher probability of remaining in the NLSCY.

Bibliography

Alderman, H., J. Behrman, V. Lavy and R. Menon, (2001), “Child health and school enrollment: A longitudinal analysis”, *Journal of Human Resources,* in press.

Alderman, H., J. Behrman, H. Kohler, J. Maluccio and S. Watkins, (2000), “Attrition in longitudinal household survey data: Some tests for three developing country samples”, mimeo, World Bank, Washington D.C.

Baker, Paula; Canada Keck; Frank Mott; Stephen Quinlan. 1993. “NLSY child Handbook – Revised Edition – A Guide to the 1986 – 1990 National Longitudinal Survey of Youth Child Data” Center for Human Resource Research, The Ohio State University. Columbus, Ohio.

Becker, Gary and Tomes, Nigel. (1986). “Human capital and the rise and fall of families.” *Journal of Labor Economics.* 4, No. 3, pt. 2 (July): S1-S39.

Becketti, S., W. Gould, L. Lillard and F. Welch, (1988), “The panel study of income dynamics after fourteen years: An evaluation” *Journal of Labor Economics*, vol. 6, pp. 472-492.

Behrman, J., (1996), “The impact of health and nutrition on education” *World Bank Research Observer*, vol. 11, pp. 23-37.

Behrman, J. (1997). “Intrahousehold distribution and the family” in *Handbook of Population and family economics* ed. By M. Rosenzweig and O. Stark, Amsterdam: North Holland.

Blau, David M. (1999). “The Effects of Income on Child Health and Development.” *Review of Economics and Statistics*, 812: 261-276.

Bound, J., D. Jaeger, and R. Baker, (1995), “Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak”, *Journal of the American Statistical Association*, vol. 90, pp. 443-450.

Boyle, Michael and Ellen Lipman. (1998). “Do Places Matter? A Multilevel Analysis of Geographic Variations in Child Behaviour in Canada.” Strategic Policy. Applied Research Branch. Human Resources Development Canada. Working Paper.

Bronfenbrenner, Urie. (1989). “Ecological Systems Theory.” *Annals of Child Development,* 6, 187-249.

Cadman, D., Boyle, M.H. Offord, D.R. Szatmar, P., Rae-Grant, N.I., Crawford, J. Byles, J. (1986) “Chronic illness and functional limitation in Ontario children: findings of the Ontario Child Health Study.” *Canadian Medical Association Journal*, 135: 761-767.

Curtis, Lori, Martin Dooley, Ellen Lipman and David Feeny. 2001. "The Role of Permanent Income and Family Structure in the Determination of Child Health in the Ontario Child Health Study" *Journal of Health Economics*, 10:4, 287-302.

Curtis, Lori and Shelley Phipps. 2000. "Economic Resources and Children's health and Success at School: an Analysis with the National Longitudinal Survey of Children and Youth." Applied Research Branch, Strategic Policy, HRDC, W-01-1-4E.

Danziger, Sheldon and Jane Waldfogel. 2000. "Investing in Children: What Do We Know? What Should We Do?" London School of Economics. CASE paper 34, Centre for Analysis of Social Exclusion. .

Davidson, R. and J. MacKinnon. 1993. *Estimation and inference in econometrics* (New York: Oxford University Press).

Deaton, D. (1997). *The analysis of household surveys* (Baltimore: Johns Hopkins University Press).

Dooley, M.D., Lipman, E.L., (1996) "Child psychiatric disorders and poor school performance: the roles of family type, maternal market work and low income." In: Towards the XXIst Century: Emerging Sociodemographic Trends and Policy Issues in Canada. Proceedings of a Symposium of the Federation of Canadian Demographers.

Dooley, M.D., Lori Curtis, Ellen Lipman and David Feeny, (1998a) "Child Behaviour Problems, Poor School Performance and Social Problems: The Roles of Family Structure and Low Income in Cycle One of the National Longitudinal Survey of Children and Youth." In Labour Markets, Social Institutions, and the Future of Canada's Children. Edited by Miles Corak, Statistics Canada.

Dooley, M.D., and Lori Curtis (1998b) "The Role of Income, Family Structure and Parental Market Work in the Determination of Child Health in the National Longitudinal Survey of Children and Youth." Canadian International Labour Network Working Paper, McMaster University.

Dooley, Martin, Ellen Lipman and Dan Offord. 2001. "Consumption Indicators, Family Income and Child Outcomes. Paper presented at the 2001 meetings of the Canadian Economics Association. Montreal.

Duncan, G.J. and Brooks-Gunn, J. 1997. (eds). *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.

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

Duncan, G.J., Brooks-Gunn, J., & Klebanov, P.K. (1994). Economic deprivation and early childhood development. *Child Development* 65, p. 296-318.

Fitzgerald, J., P. Gottschalk and R. Moffitt, 1998, "An analysis of sample attrition in panel data" *Journal of Human Resources*, vol. 33, pp. 251-299.

Glewwe, P., H. Jacoby and E. King, (2001) "Early childhood nutrition and academic achievement: A longitudinal analysis", *Journal of Public Economics*, in press.

Haveman, Robert and Barbara Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, 33 (December): 1829-1878.

Huber, P (1967) "The Behavior of Maximum Likelihood Estimates Under Non-Standard Conditions" in Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability. Berkeley: University of California Press.

Hertzman, Clyde. 2000. "The Case for an Early Child Development Strategy." ISUMA*: Canadian Journal of Policy Research*, 1:2, 11-18.

Janus, Magdalena and Dan Offord. 2000. "Readiness to Learn at School." ISUMA*: Canadian Journal of Policy Research*, 1:2, 71-75.

Jencks, Christopher and Susan Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighbourhood." In Laurence Lynn, Jr., and Michael G.H. McGeary (eds) *Inner City Poverty in the United States*. Oxford: University Press.

Kohen, Dafna, Hertzman, C. and Brooks-Gunn, J. 1998. "Neighbourhood Influences on Children's School Readiness." Strategic Policy. Applied Research Branch. Human Resources Development Canada. Working Paper.

Jenkins, Jenny and Daniel Keating. 1998. "Risk and Resilience in Six- and Ten-Year-Old Children. Discussion Paper No. W-98-23E. Ottawa: Applied Research Branch. Strategic Policy. Human Resources Development Canada. .

Knighton, T., J. Claveau, G. Laflamme and S. Michaud. 2000. Cycle 3 response rate report. NLSCY research paper, Statistics Canada.

Kohen, Dafna, Hertzman, C. and Michele Wiens. 1998. "Environmental Changes and Children's Competencies." Strategic Policy. Applied Research Branch. Human Resources Development Canada. Working Paper.

Kohen, Dafna. Clyde Hertzman, Jeanne Brooks-Gunn, Barry Foer and Miriam Linver. 2000. "Effects of Unemployment and Economic Instability on Parents and Children." Draft.

Korenman, Sanders, Janet Miller and John Sjaastad (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.

Landy, Sarah and Kowk Kwan Tam. 1998. "Understanding the Contribution of Multiple Risk Factors on Child Development at Various Ages." Working Paper No. W-98-22E. Ottawa: Applied Research Branch. Strategic Policy. Human Resources Development Canada.

Laroche, Mireille. 1998. "In and Out of Low Income." *Canadian Social Trends*, 50: 20-24.

Lefebvre, Pierre. and Philip Merrigan. 1998. "Effects of Parental Work on Developmental Outcomes of Young Children." Department of Economics. University of Montreal. Working Paper.

Lipman, E.L. and Offord, D.R. 1994. "Disadvantaged children." In *The Canadian Guide to Clinical Preventive Health Care.* The Canadian Task Force on the Periodic Health Examination, Minister of Supply and Services Canada, Ottawa, Ontario, (Cat. No. H21-117/1994E), pp. 356-368.

Lipman, E.L., and Offord, D.R. 1997. "Psycholosocial morbidity among poor children in Ontario. In G. Duncan and J. Brooks-Gunn (eds), *Growing Up Poor*. New York: Russell Sage Foundation.

Lipman, E.L., Offord, D.R., Boyle, M.H. 1994. "Economic disadvantage and child psychosocial morbidity." *Canadian Medical Association Journal*, 151: 431-437.

Mayer, Susan. 1997. *What Money Can't Buy*: *Family Income and Children's Life Chances*. Cambridge: Harvard University Press.

McCubbin, et.al., (1980). "Family Stress, Coping, and Social Support: A Decade Review," *Journal of Marriage and the Family*, 42:4, 855-71.

McLanahan, S. & Sandefur, G. (1994). Growing up with a single parent, Cambridge, MA: Harvard University Press.

Mustard, Fraser, Margaret Norrie McCain and Jane Bertrand. 2000. "Changing Beliefs to Change Policy: the Early Years Study." ISUMA*: Canadian Journal of Policy Research*, 1:2, 76-79.

Offord, D.R., Boyle, M.H., Jones, B.R., (1987) "Psychiatric disorder and poor school performance among welfare children in Ontario." *Canadian Journal of Psychiatry*; 32:518-525.

Phipps, Shelley. 1999a. The Innis Lecture. "Economics and the Well-Being of Canadian Children." *The Canadian Journal of Economics.*

Phipps, Shelley A. (1999b) *An International Comparison of Policies and Outcomes for Young Children* Canadian Policy Research Network Study No. F/05. Ottawa: Renouf Publishing Company.

Phipps, Shelley. (2002). "The Well-being of Young Canadian Children in International Perspective: A Functionings Perspective." *The Review of Income and Wealth*, 48:2, 493-515.

Picot, G. M. Zyblock, W. Pyper. 1999. "Why Do Children Move Into and Out of Low Income: Changing Labour Market Conditions or Marriage and Divorce?" Research Paper Series. No. 132 Statistics Canada. Analytical Studies Branch. .

Ross, David and Paul Roberts. 1999. *Income and Well-being: A new perspective on the poverty debate.* Ottawa: Canadian Council on Social Development.

Ross, David. Paul Roberts and Katherine Scott. 1998. “Mediating Factors in Child Development Outcomes: Children in Lone-parent Families.” Discussion Paper No. W-98-#E Ottawa: Applied Research Branch. Strategic Policy. Human Resources Development Canada.

Ross, David. Paul Roberts and Katherine Scott. 2000. “Family Incomes and Child Well-being” *ISUMA*, 1:2 (Autumn), pp. 51-54.

Ryan, Bruce A. and Gerald R. Adams. 1998. “Family Relationships and Children’s School Achievement: Data from the National Longitudinal Survey of Children and Youth” . Discussion Paper No. W-98-13E. Ottawa: Applied Research Branch. Strategic Policy. Human Resources Development Canada.

Seltzer, Judith (1994), “Consequences of marital dissolution for children,” *Annual Review of Sociology*, 20, pp. 235-266.

Shorrocks, A. (1978) “The measurement of mobility” *Econometrica*, vol. 46(5), pp. 1013-24.

Smith, Judith, Jeanne Brooks-Gunn and Pamela Klebanov. 1997. “ Consequences o Gorwing Up Poor for Young Children.” In G. Duncan and J. Brooks-Gunn (eds), *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.

Statistics Canada. 1999a. Annual Demographic Statistics 1998 Cat. No. 91-213-XIB.

Statistics Canada. 1999b. National Longitudinal Survey of Children and Youth – Cycle 2, 1996-1997, User’s Guide.

Statistics Canada. 2000. National Longitudinal Survey of Children and Youth, overview of survey instruments for 1998 – 1999 data collection, cycle 3.

Statistics Canada. 2001. National Longitudinal Survey of Children and Youth – Cycle 3, 1998-1999, Primary Users Guide.

White, H (1980) “ A Heteroscedasticity-Consistent Covariance Matrix and a Direct Test for Heteroscedasticity” Econometrica, Vol. 48, pp. 817-838.