



MINISTRY OF EDUCATION

Te Tāhuhu o te Mātauranga

New Zealand

CONTRACT RESEARCH

**THE IMPACT OF FAMILY & COMMUNITY RESOURCES ON
STUDENT OUTCOMES: AN ASSESSMENT OF THE
INTERNATIONAL LITERATURE WITH IMPLICATIONS FOR
NEW ZEALAND**

Authors: Thomas Nechyba; Patrick McEwan, and Dina Older-Aguila
Stanford University

RESEARCH DIVISION

Wāhanga Mahi Rangahau

The Impact of Family and Community Resources on Student Outcomes: An Assessment of the International Literature with Implications for New Zealand¹

Authors: Thomas Nechyba², Patrick McEwan³, and Dina Older-Aguilar⁴
Stanford University

¹ Funding for this research by the Ministry of Education, Wellington, New Zealand is gratefully acknowledged by all three authors.

² Associate Professor of Economics, Duke University; Assistant Professor of Economics, Stanford University; and Faculty Research Fellow, National Bureau of Economic Research, Cambridge, MA (tnechyba@leland.stanford.edu). Part of this research was conducted while Nechyba was on leave at the Hoover Institution (Stanford University) whose hospitality is gratefully acknowledged, as is research support from the US National Science Foundation (under grants SBR-9809269 and SBR-9905706).

³ Ph.D. Candidate, School of Education, Stanford University (pmcewan@leland.stanford.edu).

⁴ Ph.D. Candidate, Department of Economics, Stanford University (dina@leland.stanford.edu).

Abstract

The international literature as we have reviewed it suggests that, while genetics provides an important causal link from parents to child outcomes, the environment a child is raised in represents an equally important link. However, our conception of the environment that is relevant for this process has grown from an early focus on home environments to our present focus on broader environments such as neighborhoods, schools and broadly defined “community groups.” While there is evidence that shared home environments do matter for outcomes measured early in life, the same evidence tends to suggest that such environments become less important as causal explanations for outcomes in adolescence and adulthood. Furthermore, even for outcome measures early in life, certain candidates for home environment factors are increasingly being ruled out as causal factors in explaining differences in many outcomes, factors like family income, differences in parenting styles and certain aspects of family structure. To the extent that they do matter, these factors seem to be more important for children commonly labeled as at-risk, with factors like divorce and single parenthood as well as supervised child care playing a potential role. Parental choices outside the home also have some impact, with choice of school and community potentially important. Parental involvement in schools tends to raise school quality somewhat, even if a great portion of this may be specific to the child whose parents are involved. There is also evidence for some within-class peer effects even if it is difficult to interpret precisely what this evidence means. Finally, while there seems to be increasing consensus around the premise that broader neighborhood effects related to residential, ethnic and cultural communities are important, we still know little about these effects and find them difficult to quantify.

One of the broad themes of the review is that much of the earlier literature was largely correlational in nature, and researchers were often too quick to conclude that observed correlations implied causal relationships. As more sophisticated statistical techniques to control for a variety of possible biases became available, more careful analysis began to suggest that many of these initial results were incorrect. With respect to home environments, the more sophisticated analysis led to a general finding of smaller causal effects in most cases. With respect to schools, some of the most recent evidence is more positive than what was reported from initial correlations, and for neighborhood effects the evidence remains unclear, even if some optimism currently exists that more definitive and larger findings will be facilitated by better data on relevant “neighborhood” groups. In terms of broad trends, we think it accurate to characterize the earlier correlational literature as more eager to find large effects and more willing to over-interpret correlations in the data. In contrast the more recent literature is typically more methodologically careful and more satisfied with modest results. The idea of a “big answer” to the question of why children turn out as they do is largely gone, and new efforts are increasingly aimed at finding smaller but more consistent and reliably estimated effects. We generally view this as progress, even if it has meant letting go of some easy but incorrect policy answers. We also think that the international literature has much to tell New Zealand researchers, and discuss some of these implications in detail. In particular, the international research can aid in the interpretation of current New Zealand specific findings as well as in fine-tuning research on especially unique aspects of New Zealand.

Table of Contents

1	Introduction	9
2	Family/Community Characteristics and Child Outcomes: A Conceptual Roadmap	11
2.1	<i>Family and Parental Characteristics</i>	<i>11</i>
2.2	<i>Parental Choices Within the Household.....</i>	<i>12</i>
2.3	<i>Parental Choices of “Communities”</i>	<i>13</i>
2.4	<i>Parental Choices Regarding School Involvement</i>	<i>14</i>
2.5	<i>Child Choices of Peers</i>	<i>14</i>
2.6	<i>Child/Student Outcomes</i>	<i>15</i>
3	Methodology.....	17
3.1	<i>Specification Issues.....</i>	<i>18</i>
3.2	<i>Identification Issues.....</i>	<i>19</i>
3.2.1	<i>Omitted Variables.....</i>	<i>20</i>
3.2.2	<i>Correlated Unobservables and Measurement Error</i>	<i>20</i>
3.2.3	<i>Selection Bias or Endogenous Group Membership</i>	<i>22</i>
3.2.4	<i>Simultaneity.....</i>	<i>23</i>
3.3	<i>Conclusion.....</i>	<i>24</i>
4	Heredity versus Environment as a Causal Link	27
4.1	<i>Heredity, Race and Social Science</i>	<i>29</i>
4.2	<i>Relevance of Heredity for Social Scientists and Policy Makers</i>	<i>30</i>
4.3	<i>Heritability versus Environment: Evidence from Behavioral Genetics</i>	<i>32</i>
4.3.1	<i>Quantitative Estimates of the Role of Heredity</i>	<i>32</i>
4.3.2	<i>Caveats and Cautions: What Do These Studies Really Tell Us?</i>	<i>37</i>
5	Causal Links of Child Outcomes to Within-Household Choices	41
5.1	<i>Non-Shared Home Environments: Birth Order</i>	<i>42</i>
5.2	<i>Family Structure</i>	<i>44</i>
5.2.1	<i>Sibship Size, Spacing and Configuration.....</i>	<i>44</i>
5.2.2	<i>Divorce</i>	<i>46</i>
5.2.3	<i>Single Parenthood.....</i>	<i>48</i>
5.2.4	<i>Parental Age and Non-Traditional Households</i>	<i>50</i>
5.3	<i>Parental Resource Investments in Children</i>	<i>51</i>
5.3.1	<i>Parental Time Investments: Quantity and Quality</i>	<i>51</i>
5.3.2	<i>Parental Income and Money Investment.....</i>	<i>57</i>
5.3.3	<i>A Brief Note on Maternal Education</i>	<i>62</i>
5.4	<i>Summary.....</i>	<i>62</i>
6	Causal Links Between School Resources and Child Outcomes	63
6.1	<i>Parental Involvement in Schools</i>	<i>64</i>
6.1.1	<i>The Form of Parental Involvement.....</i>	<i>65</i>
6.1.2	<i>Which Parents Choose to Become Involved?.....</i>	<i>66</i>
6.1.3	<i>Effects of Parent Involvement on Outcomes</i>	<i>68</i>
6.2	<i>School Peer Groups.....</i>	<i>69</i>

6.2.1	Causal Mechanisms Underlying “Peer-Group” Effects	70
6.2.2	School-Wide Peer Influences.....	72
6.2.3	Ability-Grouping Within Schools.....	75
6.2.4	Ability-Grouping Within Classes	79
6.2.5	Summary of Research on Three Types of Peer Effects	81
6.2.6	A New (Untested) Perspective on Peer Effects	82
6.3	<i>School Inputs</i>	83
6.3.1	School Effects on Achievement.....	83
6.3.2	School Effects on Labor Market Success	86
6.3.3	Summary.....	87
7	Causal Links between Community Resources and Child Outcomes	89
7.1	<i>Mechanisms Linking Neighborhoods to Outcomes</i>	91
7.2	<i>Evidence on the Effects of Culture and Community Resources</i>	92
7.2.1	Effects of Non-Neighborhood Community Resources: Culture, Race and Ethnicity	93
7.2.2	Statistical Evidence on Neighborhood Effects.....	94
7.2.3	Controlled Experiments	99
7.2.4	Summary.....	101
7.3	<i>Residential Choice: How do Families Choose Communities?</i>	101
7.3.1	Residential Segregation by Income	102
7.3.2	Residential Segregation by Race	103
7.3.3	Residential Peer Groups and Education.....	104
7.3.4	Conclusion.....	105
8	Research Implications for New Zealand.....	107
8.1	<i>The Usefulness of International Findings for New Zealand Researchers</i>	108
8.2	<i>Unique Aspects of New Zealand Calling for New Zealand Specific Research</i>	109
8.2.1	Maori and Pacific Island Cultures	110
8.2.2	New Zealand’s Innovative School Reforms	111
8.2.3	Policies Aimed at Revitalizing the Maori Language and Culture.....	112
9	Broader Implications for the Social Sciences	115
9.1	<i>The Use of Statistics and Econometrics</i>	115
9.2	<i>Do We Really Know So Little?</i>	116
9.3	<i>What Next?</i>	117
10	Summary	121
	Tables.....	123
	References	124

1 Introduction

This review provides a critical synthesis of recent empirical research that links the attributes of families and communities to child outcomes. It is intended to establish common results from the economics, education, psychology, and sociology literatures, identifying areas of consensus and controversy as well as areas that remain relatively unexplored. Although we are aware of descriptive and qualitative research that covers similar topics, we note at the outset that such research is beyond the scope of our review. Our primary aim is to evaluate studies from various disciplines that explore the link between family/community resources and child outcomes through statistical analysis of mostly non-experimental data sets, without investigating all that has been written on such links.⁵ While much of this analysis, in part due to data constraints, is U.S. focused, we will attempt to assess to what extent lessons from this literature are consistent with findings from research using New Zealand data and informative for New Zealand policy makers.⁶ Of special concern is the identification of current research findings that may be useful in addressing the education and social policy issues that are relevant to indigenous cultures of New Zealand.

While the empirical literature we seek to review is vast, the most informative studies are much fewer in number. The best research goes beyond merely documenting correlations between family/community resources and child outcomes. By using more sophisticated statistical techniques, they investigate to what extent correlations actually represent causal links. Factors such as socioeconomic characteristics of families, for example, are correlated with a variety of other and often unobserved characteristics of families and communities. This implies that mere correlations of socioeconomic variables with student outcomes are largely uninformative, because the underlying causal explanation for such correlations may be unrelated to socioeconomic characteristics. Unfortunately, the bulk of the literature in the social sciences has concerned itself with identifying broad correlations and over-interpreting these as important causal channels. As a result, we argue in this review that a great deal of “conventional wisdom” in this area is unfounded; much work remains to be done in order to uncover the causal channels that link families and communities to child outcomes. Statistical techniques that allow for a more careful analysis of the kinds of data often used to identify correlations are, however, beginning to be used for precisely the purposes we are concerned with in this review. Therefore, ours is not meant to be an uncritical “meta-analysis” of the hundreds of related studies done in the last decade, which might simply confirm a series of well-known but uninformative correlations. Rather we focus on the few studies that go beyond traditional exercises.

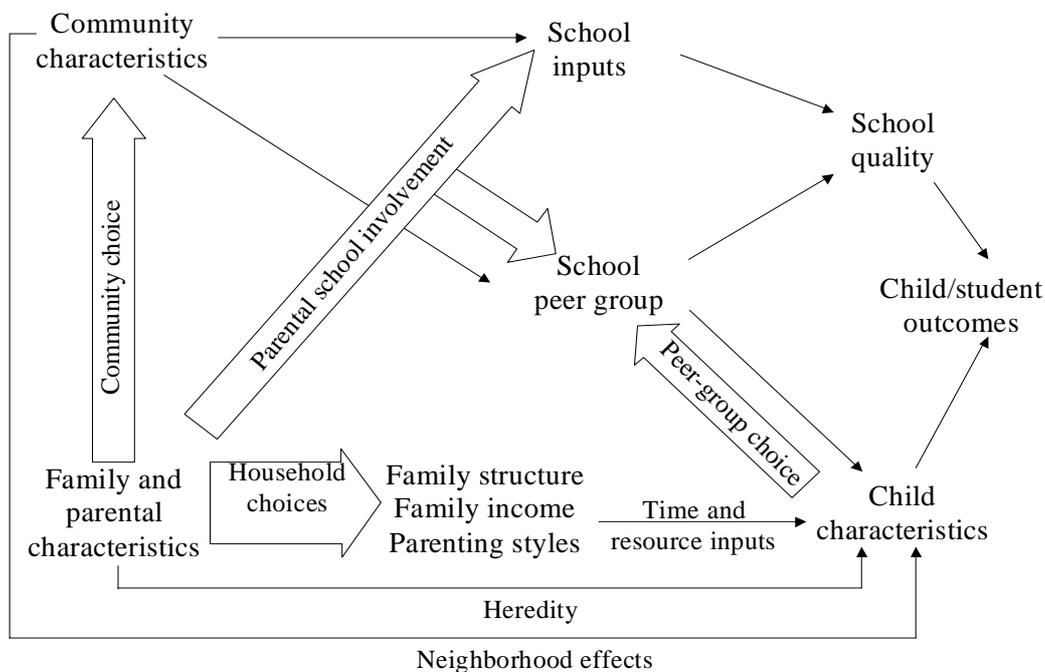
⁵ This does not mean that we see no value in such descriptive and qualitative research. Rather, we find such research thought-provoking and interesting. Insights gained from case studies, for instance, often form the basis for broader quantitative statistical analysis of the kind we investigate in this review, analysis that is aimed at testing the general applicability of such insights. Thus, while we are cognizant of the importance of qualitative and descriptive research, we suspect that it is difficult to ultimately accept inferences and insights from such research without subjecting them to the kinds of rigorous testing made possible through the application of statistical techniques to larger data sets.

⁶ Much of what we include in terms of New Zealand based analysis derives from the Christchurch Health and Development Study, a 20 year longitudinal study that contains some quite unique data absent from other longitudinal data sets (Ferguson, Lloyd et al. 1991; Ferguson, Horwood et al. 1993; Barker and Maloney 1999). Results from this study give us confidence that insights gained from the international literature are often quite applicable to the New Zealand context. We do, of course, recognize that not all such results may generalize because of the proportionately small representation of certain groups (such as Maori) in the Christchurch area compared to areas in the North Island. In addition to the Christchurch studies, more recent and more preliminary data is generated by a new longitudinal study known as “Competent Children”(Wylie, Thompson et al. 1999). Unfortunately this study is not nearly as far along as the longitudinal Christchurch study, and therefore little rigorous analysis has been conducted using its data thus far.

In order to properly interpret findings from different branches of the social sciences, however, it is imperative to begin with a framework that can accommodate the various theories put forth by researchers in different fields and to establish a common language through which to discuss the evidence. Our review therefore begins in Section 2 with an exposition of a conceptual framework that delineates the possible causal channels that might result in the observed and well-documented correlations between family/community resources and child outcomes. Section 3 then describes the methodological issues involved in unraveling precisely what these correlations mean. In particular, it focuses on statistical problems of interpretation that much of the rest of the review will refer back to. Section 4 then begins the substance of the review with a discussion of the importance of heredity relative to environmental factors, while Section 5 investigates the evidence in support of a substantial role for within-household environments. Finally, Sections 6 and 7 search for evidence that environmental factors important for child development may also be found outside the home in schools, neighborhoods and ethnic/cultural groups. While we make reference to issues particular to New Zealand throughout, Section 8 focuses on the implications of our findings for the New Zealand policy environment. Section 9 discusses directions for future research, while Section 10 summarizes and concludes.

Because of the length of this review, some readers who seek a rather quick overview of our basic conclusions will find it frustrating to read through the entire document. With this in mind, we have therefore written a brief synopsis in italics prior to each of the sections of the review. This should enable the casual reader to glance through the document, read the italicized portions to get an overview, and then explore the details of his or her interests. We also note at the outset that every effort has been made to be comprehensive in the major elements of this review, but research in this area is progressing at such a fast pace that we became aware of new studies even in the final weeks of writing this report.

FIGURE 1
A Conceptual Model



2 Family/Community Characteristics and Child Outcomes: A Conceptual Roadmap

This section presents the conceptual framework used to organize the literature review. Summarized in Figure 1, it describes the major causal channels which link family and community characteristics to child outcomes. Family and parental characteristics range from genetic endowments to parental education and occupation. Given these characteristics, families make several choices that are represented in the framework by double-arrows. First, they make decisions within the household that may affect child outcomes; these include, among others, parenting styles, work habits, and financial investments in children. Second, they choose a geographic and social community, which may influence child outcomes via neighborhood effects. Third, they choose whether to become directly involved in school activities. The framework also allows that children (and perhaps parents) may choose a within-school peer group. The composition of the peer group may exert further influence over student outcomes. Ultimately, outcomes such as academic achievement and attainment are influenced by a multitude of factors relating to the families, communities, and schools. The empirical challenge facing researchers is to credibly isolate each of these effects.

Our aim in this section is threefold: (1) to outline a conceptual framework that describes the major channels linking family and community characteristics to child outcomes; (2) to define terms and arrive at a common language with which to discuss studies from very different branches of social science; and (3) to provide the reader with a sense of the challenges faced by empirical researchers who analyze non-experimental data in hopes of ascertaining the underlying causal linkages within this model. Our exposition is intended to be multidisciplinary, encompassing theoretical insights from a variety of social sciences including economics, education, psychology, and sociology. It is our hope that this model will provide a useful roadmap through the diverse literatures we review in the remainder of this paper, and will aid us in evaluating and comparing the various theories contained in the model. In this section, however, we take no particular position on the relative merits of the various causal channels, reserving this discussion for later sections.

The major elements of the general model are outlined in Figure 1. The model finds its origins in literatures across the social sciences in that it allows for the potential impact of family and community characteristics through a large number of causal channels, ranging from hereditary to environmental effects. Linkages that may result from individual choice behavior are indicated by double arrows, while more passive linkages are indicated by single arrows. While we recognize that this distinction is somewhat artificial and involves some broader philosophical judgements,⁷ it is not critical to most of our findings below. We begin our discussion of Figure 1 with the lower left-hand corner, family and parental characteristics.

2.1 Family and Parental Characteristics

Family and parental characteristics include such diverse elements as genetic endowments and abilities, cultural traits, and previously chosen factors such as levels of education and occupations. Some of these could be passively transmitted to the child through a process we denote “heredity”—a process that shapes child characteristics such as cognitive ability and personality traits. Studies of twins and adoptive siblings, for instance, claim to find an important role for heritability in the transmission of personality traits and levels of intelligence (Pedersen 1992; McGue 1993; Plomin 1994), and recent controversial work on racial differences in cognitive skills has inflamed old debates regarding the genetic

⁷ The choice of “community”, for instance, is certainly in part a real choice but is also constrained by such factors as income, culture and race.

component of such differences in the United States (Herrnstein and Murray 1994).⁸ One major problem in ascertaining the hereditary component of various child traits derives from the fact that the same parental characteristics that may be genetically transmitted to children also shape the home environment and could determine parental choices that in turn impact children through what we commonly call environmental channels. All arrows in Figure 1 other than that denoted “heredity” represent such non-hereditary channels.

Given parental characteristics, we assume that parents make three types of choices (denoted by double arrows): within-household choices, school involvement choices, and “community” choices. Each choice may shape the characteristics of the child through other mechanisms, thereby contributing to outcomes such as test scores, school completion, and fertility choices. Furthermore, for reasons delineated below, we allow that children may exercise “choice” in the determination of their peer group, as indicated by the double arrow emanating from “Child characteristics” in Figure 1. As mentioned already, the extent to which these double arrows represent conscious choices unconstrained by the various characteristics of the parents and children making the choices is controversial and beyond the scope of this review, especially as different culturally based conceptions of “family” arise.⁹ Fortunately this is also largely irrelevant for our purposes. All we seek to do here is establish a conceptual framework with which to organize a large and diverse body of literature. With minor alterations, this framework should be able to incorporate a wide range of theories and research questions.

2.2 *Parental Choices Within the Household*

The parental choices most often discussed in relation to child outcomes are those made within the household and represented by the double arrow emanating from “Parental characteristics” in Figure 1. A vast body of literature over the course of this century has investigated the relationship between child outcomes and parental decisions regarding parenting styles, number of children, divorce, single parenthood, early motherhood, financial investment in children, time spent with children, and work habits of parents. Various theories suggest that these choices may have important implications for child development, and these theories are often supported by correlations in the data and powerful anecdotal evidence. Furthermore, these theories have become part of every parenting handbook, with parents becoming increasingly sensitive to how their choices may impact their children’s future.

However, these choices are not made in a vacuum. Rather, they depend in great part on the underlying characteristics of parents, their level of education, their cultural identity, and perhaps even their genetic endowments. Such characteristics may impact children in a variety of ways. If certain personality traits are, for instance, in part genetically influenced, then these traits will influence children both because they are inherited and because they influence the choices parents make regarding the home environment. Furthermore, if a particular parental

⁸ The descriptive statistics of the Maori population (Ministry of Maori Development 1998; Ongley, Carson et al. 1998) may lend themselves to similar analysis. However, as we make clear later, we find no credible evidence for the hypothesis that inter-race differences in cognitive ability are genetic in nature, either for the US or New Zealand.

⁹ Within the New Zealand context, of course, the term “family” has quite a different meaning for traditional Maori than it does in Pakeha settings (Kammerman and Fahn 1997; Smith 1998), and traditional Maori are more likely to emphasize the role of extended family and less likely to think of parents as the crucial decision maker in a child’s life. At the same time, in the past four decades Maori in New Zealand have increasingly urbanized and become disconnected from traditional extended family relationships. Our sense, therefore, is that for many Maori families, parents are indeed increasingly active decision makers on behalf of their children. In any case, the focus in our conceptual framework on the parents as the crucial but constrained decision maker does not exclude the possibility of parents in certain cultural settings making many of these decisions with the extended family or other traditional institutions such as hapu and iwi. As we point out below, such choices could be represented by the arrow labeled “community choice.”

characteristic influences parental choices in one dimension such as parenting style, it is also likely to influence the parents' decisions in other dimensions such as family structure. The underlying characteristics that cause parents to make certain decisions with regard to children may therefore have both direct and indirect effects on their children through a variety of channels, and, as we point out more carefully in Section 3, these effects may be difficult to untangle. Estimation of these effects is further complicated by the possibility that parental choices are also responsive to child characteristics. Different choices may reflect underlying differences in children that will have their own separate impact on outcomes.

2.3 Parental Choices of "Communities"

The set of parental choices impacting children is not restricted to those made within the immediate household. Parents also choose (1) which residential area to live in; (2) which social communities within neighborhoods to associate with; and (3) which schools to send their children to, thereby choosing the nature and quality of their child's academic upbringing and peer group. Broadly interpreted, each of those choices involves a choice of "community," where the relevant community for child development may be defined by residence, social group, culture, or school. In the case of some more traditional Maori parents in New Zealand, one of the important choices of "community" may involve close identifications with traditional Maori institutions such as hapu and iwi. Such broadly defined choices are represented by the vertical double arrow emanating from "Parental characteristics" in Figure 1.

Community characteristics such as average education or racial composition may then shape child outcomes through a variety of channels. Reviews by Jencks and Mayer (1990) and others recognize that neighborhood peers as well as the parents of such peers might influence a variety of child outcomes. Thus, choices of residential communities may directly impact child development. We denote such effects as "neighborhood effects" in Figure 1, and we recognize that such effects may not be confined to spatial neighborhoods and may extend to social groups chosen by parents. Similarly, parents also choose a community for their children by choosing a school. In many contexts, school choices are intimately linked to residential community choices.¹⁰ School choice (in whatever form) is therefore a kind of community choice that determines the set of potential peers that children will interact with in the school setting. This possibility is represented by the arrow connecting "Community characteristics" and "School peer groups" in Figure 1. Furthermore, the school/community choice determines the types of school inputs a child will benefit from, inputs that are not necessarily restricted to financial inputs. For instance, some communities (parents and non-parents) are deeply involved in their local schools, and this involvement is one way in which community characteristics may translate into better schools (Epstein 1987). Similarly, the wealth and generosity of a community may also determine the level of financial resources available to the school.¹¹ Both financial and non-financial inputs into schools may thus represent important causal links from parents and communities to child outcomes, a possibility represented in Figure 1 by the arrow linking "Community characteristics" to "School inputs." Finally, we

¹⁰ In the U.S. context, eligibility of attending a particular public school is almost solely determined by the residence of the household. Those parents choosing to send their children to public school therefore simultaneously choose schools and neighborhoods. In New Zealand, of course, residentially based school districting is no longer in force. Thus, while parents may still be influenced in their residential location decision by the general availability of quality schools in the vicinity, neighborhood and school choice are no longer synonymous.

¹¹ In many U.S. school districts where funding is partially determined by local governments and funded through local taxation, the link between community characteristics and financial school inputs is often quite direct. In New Zealand, on the other hand, this link is not nearly as visible but nevertheless appears in the form of voluntary parental contributions to public schools (Fiske and Ladd 1999). Such voluntary contributions play an important role in California where funding is similarly centralized (Brunner and Sonstelie 1999).

assume in Figure 1 that school inputs such as levels of community involvement and financial inputs combine with school peer group effects to determine the quality of a particular school.

The potential existence of all these links complicates the process of finding which parental characteristics and choices are actually of causal significance. Parents who choose communities that are conducive to the healthy development of their children are also likely to make important choices at home and may transmit characteristics to their children via heredity. While the documentation of a correlation between neighborhood wealth and positive child outcomes may, therefore, be indicative of important neighborhood effects, it may equally well reflect the fact that parents who chose to live in such neighborhoods have underlying characteristics that translate into good child outcomes through some other channel (such as heredity or within-home environments.) Similarly, such a correlation may arise from the fact that children who are raised in wealthy neighborhoods attend schools in which parents are more involved, and that community involvement in schools is actually the underlying causal link; or that schools attended by children from wealthier neighborhoods are financially richer and thus able to provide a richer intellectual environment for their children.

2.4 Parental Choices Regarding School Involvement

While the previous discussion suggests that communities may impact school quality through both school inputs and peer group composition, the quality of schooling received by a child is unlikely to be solely determined by the parents' choice of community. For example, sociologists have long attempted to ascertain to what extent a parent's direct involvement in a child's school (as distinct from the choice of a school in which average parental involvement is high) impacts her child specifically (e.g., Sui-Chu and Willms 1996). In Figure 1, this parental choice of school involvement can take two forms. First, a parent may directly provide inputs to the school, particularly time spent in the classroom. Second, a parent who is especially involved may be influential in the child's choice of peers within the school, whether explicitly through a choice of classroom for the child in a "tracked" environment, or more subtly through influencing which other children the child spends time with inside the classroom. Harris (1995) argues that children assume dramatically different personalities reinforced through their peers when placed outside the immediate supervision of their parents, which suggests the potential for parental involvement in a child's school to be of particular importance not just to the school but especially to that parent's child. Again, cultural differences may constrain parents from becoming involved in schools and challenging teachers. This seems to be the case among some Pacific Island parents in New Zealand whose culture emphasizes respect for teachers and deference to the authority of professional educators.

The preceding discussion makes clear that parental involvement in schools may have both "public" and "private" benefits, a fact often overlooked in empirical studies. The "public" benefit is one that we discussed in the previous section: the level of general parental and community involvement within a school may be one important school input, and one parent's decision to become involved contributes to school environment for all children. The "private" benefit, on the other hand, may arise if a parent who is involved in her school can benefit her child directly— perhaps by diverting resources such as teacher time to her child, causing her child to focus more on school work, or helping the child select better peers. As before, the possibility of these additional causal channels linking parental characteristics to student outcomes further complicates the task faced by empirical researchers.

2.5 Child Choices of Peers

In our discussion of Figure 1 we have primarily focused on the characteristics of parents and various possible implications of choices made by parents given those characteristics. This has generated a set of causal links that could underlie the many observed correlations between

parental/community characteristics and child outcomes. Some recent influential work by Harris (1995), however, has cast doubt on whether a model based solely on parental choices does not leave out one important decision maker in the child's development: the child herself.

Harris develops a theory of group socialization in which children's peers matter more than parents and home environments. Furthermore, she uses the existing empirical psychology literature to argue that children themselves are quite active participants in the selection of the relatively small peer group that is most relevant for their development. This possibility is incorporated into Figure 1 with the double arrow connecting "Child Characteristics" (assumed to influence child peer decisions) with "School Peer Groups". Thus, even in communities with "good characteristics" on average or in schools with "mostly good peers," a child left to choose on his own may fall in with "bad" peer groups. The negative peer effects might outweigh positive neighborhood or school level peer effects, but without knowing the exact relevant peer group of each child, researchers will have a difficult time separating these effects in large data sets. At the same time, the child's choice of peers may simply reflect underlying unobserved characteristics of the child that impact outcomes independently but are mistakenly attributed to peer effects.

2.6 *Child/Student Outcomes*

Ultimately, the conceptual framework assumes that measured child outcomes like juvenile delinquency, teenage pregnancy, college attendance, and academic achievement¹² result from the combination of the child's characteristics and the quality of schooling he receives. These two ingredients, however, are themselves influenced by many factors, as illustrated in Figure 1 and discussed above. The child's characteristics are determined through heredity, parental behavior, and neighborhood effects, while the quality of schooling results from a combination of school inputs in part determined by the community, by parental involvement, and by peer group effects. Our task in the remainder of the paper is: (1) to get a sense from the literature of which linkages in Figure 1 have the strongest empirical support and also have effects that are of practical significance; and (2) to discuss policy implications of these findings in general as well as for the particular context of New Zealand. Before turning to this empirical evidence, however, we briefly discuss the methodological challenges to inferring causal links from non-experimental data.

¹² While the conceptual framework can be broadly applied to study a variety of child outcomes, our main focus in this review is on cognitive and school outcomes. We will, however, occasionally comment on other types of outcomes when the literatures are sufficiently intertwined.

3 Methodology

Readers of empirical research on children's outcomes should be aware of several factors that can undermine our ability to infer causal relationships between two variables in non-experimental data. To begin with, the way a research question is posed can itself affect the relationships that we are able to uncover. For example, a model may include assumptions on: the level at which peer effects take place (an entire school versus a grade or class); the characteristics of a group that matter (mean ability versus the range of abilities); and the shape of the relationship (whether income matters over the entire range of values or whether there is a threshold beyond which higher incomes do not affect outcomes).

Within a given model, estimates of an input's direct effect will be biased if that input is related to unobserved forces that also affect outcomes. For example, if an important determinant (such as parental income) of the dependent variable is left out of a study, its impact may be incorrectly attributed to an included, correlated variable (such as ethnicity). When evaluating study findings, it is important to consider what factors affecting outcomes are not included in the model and how these factors are related to the variable under study. A similar problem can arise when group characteristics are good estimates of individual characteristics. Even in the absence of a causal relationship, average group traits will be correlated with outcomes if they act as a proxy for poorly measured individual characteristics. Another important source of bias is the process by which individuals are assigned to groups or chose their own traits. If groups and traits are not randomly assigned, we should ask if the mere fact that someone has chosen or been assigned to a particular group or trait says something about them and their expected outcome separate from any independent effect of the group they belong to. Finally, if group outcomes and individual outcomes (or explanatory and dependent variables in general) determine each other in a recursive fashion, the estimation of such factors as peer effects can become clouded. If the reciprocal relationship is not accounted for we will be unable to tell the extent to which group outcomes drove observed individual outcomes versus the extent to which an individual impacted the group outcome. These potential limitations should be kept in mind when evaluating the claimed findings of any study.

There are numerous suggestive correlations between family characteristics, community characteristics and children's educational and life outcomes. However, it is very difficult to identify the extent to which these correlations are driven by causal relationships. In studies of family, peer, and community effects we face the general problem of research in the social sciences: the absence of controlled experiments. We are usually unable to manipulate the variables of interest (such as family size, divorce status or ethnicity) and simply measure the ensuing changes in outcomes. Instead, most of the literature we review uses existing variation in population characteristics and outcomes to learn about the effects of different inputs into children's education. Researchers compare the outcomes of two groups of children that have different amounts of a given characteristic, say mother's education, and make statistical "controls" for other key determinants of the outcome. Such analysis often makes use of regression analysis, where the correlation between a "dependent variable" (such as child test scores) and various "independent variables" (such as family background, school characteristics, etc.) is analyzed. Such statistical methods are helpful in attempts to single out the factors that may matter to outcomes, that are related to differences in outcome levels or that can be used to predict achievement and attainment levels. We must take care, however, not to blindly interpret correlations found in data analysis as evidence of causal relationships. In fact, most of the empirical work in this field is severely compromised by errors that bias estimates of the direct effects of the independent variables that are studied.

The goal of this section is to explain to the reader the problems commonly encountered in these studies, to discuss potential statistical remedies, and to prepare the reader to evaluate the effectiveness with which the existing studies have addressed these issues. We begin with a

discussion of statistical model specification and move on to a presentation of issues involved in identifying the direct, causal effect of a specific variable.

3.1 *Specification Issues*

It is first important to understand that the way in which a research question is framed or modeled will contain many implicit assumptions about the phenomenon under study. These assumptions often remain unstated in the studies but may be important to ascertain whether a particular interpretation given by the author is correct. In order to distinguish between stated research goals and the questions a paper is truly able to ask and answer, it is necessary to take a closer look at the exact statistical equations that are estimated. The variables that are included or excluded, the ways they are measured, and the forms in which they enter the model will all affect our interpretation of the data and the relationships we will be able to detect. This section will highlight some issues involving the specification of the statistical model that are particularly relevant to the study of neighborhood and peer effects, as well as some more general specification issues.

The first step in studying the impact that a social group's members have on one another is to define our notion of a group. As Tienda (1991) discusses, the empirical study of groups is complicated by our lack of understanding of the exact way in which peer effects operate and at what social level these effects take place. Is the larger community important (entire neighborhoods and schools), or just those people with whom a family or individual has close interactions (such as church and other social groups or small friendship groups)? Unfortunately, we cannot use statistical relationships between group characteristics and member outcomes to locate the relevant group since, as we will show in the section on identification, a measurable "group effect" may be present even when there is no true effect of group membership on individual outcomes.

Once a particular group level is selected, we are often limited in our ability to study this group by data availability. Some theories of peer effects stress the importance of small groups, as few as three or five children (Harris 1995), but data at this level of detail are very rare. The characteristics of larger groups may be insufficient to answer questions about the effects of close peers. For example, within two schools that differ greatly in mean student ability levels, children in both schools may still associate only with children of similar ability. Using available data, we will be unable to learn anything about the effect of a close friend's ability on individual attainment. Furthermore, data on groups are often collected with reference to geographic and political boundaries (such as counties or school districts), which may be poor measures of relevant social groupings (such as religious or ethnic communities). The conceptual model of Figure 1 explicitly allows for a broad conception of "community," ranging from spatial definitions to ethnic or cultural ones. However, data often limit our ability to empirically test which notion of community is most relevant.

Related to the issue of group definition is the selection of the relevant measure of group characteristics. While many studies take the mean of group outcomes as a summary of group influence, the variation in or range of characteristic values may also influence member outcomes. This limitation will affect both our estimates of group effects and the very nature of the questions we can study.¹³ For example, a study that only uses the mean ability of a class group would be incapable of asking whether classes with a large spread of abilities are more difficult to teach and lower children's overall attainment.

Model specifications often include implicit assumptions regarding the effect on outcomes of different combinations of characteristics. For example, a study may implicitly assume that the

¹³ Glewwe (1997) illustrates how empirical estimates of peer effects can be quite sensitive to empirical decisions such as these.

effect of parental education on student outcomes is the same for all income levels. Group effects in particular are often assumed to affect outcomes in such a manner, with no moderating effect of individual traits on group effects (Duncan, Connell et al. 1997). To allow for more complex relationships between the determinants of outcomes, the equations estimated can include interaction terms (terms in which a variable of primary interest is multiplied by indicators of other characteristics) or separate regressions may be calculated for different groups. For example two regressions (one for boys and one for girls) may be estimated to find the effect of single-sex schools on self-confidence, and to explore whether this effect differs based on gender.

Another common restriction is the assumption of a linear relationship between the explained and explanatory variables. In other words, a model may implicitly assume that the impact on outcomes of an increase in some factor (such as income or school inputs) is the same over the entire range of that factor. This will mask the effect of increases that vary for different values of the variable. For example, increases in parental income may be very important for children below some threshold level, but have little effect on children from middle and higher income families. Including squared (or higher order) terms in regression equations might better enable us to pick-up such relationships.

Finally, even within a given “true” model of outcome determination, we may estimate different equations to learn about either the direct or indirect effects of a particular variable. For example, imagine that in addition to any within-home effect of parental education on children’s achievement, parental education also has the indirect effect of increasing school quality because parents with higher incomes are better able to exert pressure on local educators and school boards. Including measures of school inputs would help to isolate the direct effect of differences in parental education. This would be useful in measuring the relative importance of parental and school inputs. However, if we are interested in the larger advantage held by children with highly educated parents, holding teacher effort and class inputs constant may underestimate the full benefit.

When approaching an empirical study of the kind we review in much of this paper, it is therefore vital to critically examine the way the study frames or specifies the research question. What kind of effects are included? How are different aspects of the problem measured? Are effects assumed to be linear only? And are we measuring an effect that is truly the effect we are interested in (for policy purposes, for example)? Many times researchers will claim that they are making no particular assumptions but will simply “let the data speak for themselves” by including many variables in a regression. As we hope to emphasize throughout this review, however, there is no such thing as “letting the data speak for themselves.” The very fact that a regression model has been specified and particular measurements have taken place implies a variety of underlying assumptions that must be delineated and questioned. Once a model has been specified and the researcher is comfortable with the specification, the issue of inferring causation from correlations arises. We now turn to a discussion of difficulties commonly encountered in estimating such true causal effects of a single variable on a given outcome.

3.2 Identification Issues

Identification refers to the ability to separately determine the independent effect of each factor used to explain the variation in the outcome we are studying. In many cases it is technically possible to estimate the effect of a particular factor, but that estimation is biased, either upward or downward. Much of the literature we will review is plagued by potential biases that limit our ability to attribute causal significance to correlations in the data.

Often, problems in identification can be understood as issues of **endogeneity**, or the correlation of an explanatory variable with unknown factors also affecting outcomes. It is

impossible to completely specify all the factors that determine a given outcome; there will always be some unknown portion of variation across individuals. This does not pose a problem so long as what is unknown is unrelated to the variable whose effect we want to identify. If, however, unknown factors are correlated with explanatory variables, we may incorrectly attribute the effect of the unobserved factor to included factors. We will examine four common sources of endogeneity: omitted variables, measurement error and correlated unobservables, selection bias and simultaneity.

3.2.1 *Omitted Variables*

Omitted variable bias is often created when important determinants of the dependent variable are simply left out of the regression. When the independent variable of interest (the variable whose effect we are trying to estimate) is correlated with a variable that has been left out, the estimated effect of the included variable will include some of the impact of the omitted variable. The size and direction of the bias will depend on the nature of the relationship between the excluded and included variables and their relationships with the dependent variable.

For example, consider a study that attempts to estimate the effect of divorce on children without measuring family income and mobility. The omitted variables in this regression, low family income and frequency of relocation, are both correlated with divorce; children from divorced families often experience a decline in family income and are forced to change residences. If higher incomes also increase outcomes, children of divorced parents will tend to have lower outcomes simply because they tend to have lower family income. Similarly, children of divorced parents will tend to have lower outcomes if higher mobility leads to lower outcomes. Both of these effects will strengthen the estimated negative relationship between divorce and children's outcomes, although it is possible that an omitted variable could also lead to a weakening of any negative effect of divorce. If there were a third factor that was positively related to divorce, but tended to improve children's outcomes, such as income or maternal education, exclusion of this factor would tend to make divorce seem more beneficial to children.

In the above example it is clear that one potential solution is to gather more information on the factors influencing the measured outcome. Anecdotal evidence and conventional wisdom may identify omitted explanatory variables that are related to the variable whose impact we are attempting to measure. Unfortunately many important variables such as child ability, school quality, and time spent on education within the home, may be exceedingly difficult to measure.

When evaluating findings on the impact of particular variables it is therefore important to consider the following questions: what factors affecting the dependent variables are not included in the regression? Are these factors correlated with variables that are included in the regression? And finally, in what direction is this likely to bias the estimated impact of the included variables?

3.2.2 *Correlated Unobservables and Measurement Error*

Correlated unobservables refers to instances in which the size and direction (either positive or negative) of the unexplained variation in outcomes tends to be the same within groups. An example would be if all the children within a class tended to either over- or under-perform on a test given what we know about them. This common component may incorrectly be interpreted as the effect of overall group outcomes on individual student outcomes (i.e., a peer

effect may be found where none exists). This is due to the fact that members of a group will all tend to be either unexpectedly high achieving or low achieving giving the impression that students in high achieving classes do unexpectedly well because of the exposure to high ability peers.

Within this example, a potential source of correlation might be teacher skill level, or any other classroom input that the students in a class all share. If the model is estimated without a measure of teacher skill, given the observed measures of child characteristics (such as family background or ability), all students in a class will tend to have unexpectedly high scores when teacher skill is high and low scores when teacher skill is low. Because of teachers' ability to raise test scores, a classroom with high mean test scores is more likely to have a good teacher than a classroom with low mean test scores. Therefore, a child in a class with high mean test scores is more likely to have a highly skilled teacher and an unexpectedly high test score.¹⁴ A positive relationship between group achievement and student achievement would exist even if classmate ability has no effect on a student's performance.

Errors in the measurement of explanatory variables can also create endogeneity problems.¹⁵ If an individual's characteristic is not perfectly measured and the included group mean provides a good additional measure of the characteristic then some of the effect of the individual trait on outcomes will be associated with group values of the trait.

To illustrate this effect, consider a model of the effect of mean neighborhood income on children's college attendance rates. If neighborhoods are composed of people with similar lifetime incomes, neighborhood income may be a good proxy for individual family resources. Measurement error can be introduced if annual incomes are all we observe, but lifetime resources are the true determinants of attendance probabilities. For example, younger parents are likely to have an annual income that is below their long-run annual earnings while older parents with the same lifetime earnings are likely to have higher incomes, above this long-run average. But children from families with the same lifetime earnings will be equally likely to attend college even if their parents' income seems to be very different. Neighborhood income will be positively related to college attendance (even after considering individual family income) because children in high-income neighborhoods are themselves more likely to be from high-income families even if they do not seem affluent based on their parents' current earnings.

Household resources, educational attainment and student ability are three commonly examined determinants of outcomes that are often thought to be both poorly measured and closely related to group measures. Snapshots of annual income are usually the only available measure of family resources even though long-run income may be more important; educational attainment is measured by the years of schooling completed with no adjustment for the quality of education received; and innate student ability is very difficult to measure. And yet, there are plausible reasons that individuals will tend to live and attend school with people with similar measures of these three traits. Particular attention should be paid to measurement error bias created by these three variables. Again, the question should be asked if community characteristics are correlated with outcomes because of an independent effect on outcomes, or because they provide additional information about the true value of individual characteristics.

¹⁴ Endogeneity is present because the explanatory variable, mean score, is correlated with an unobserved determinant of outcomes, teacher skill.

¹⁵ This is in addition to the familiar bias created by measurement error; error in the measurement of an independent variable will lead to an underestimation of the magnitude of its effect on the dependent variable.

One potential statistical solution to these sources of bias is the estimation of “fixed effects” models. Fixed effects estimations allow for the possibility that each group has its own unobservable impact on outcomes. Instead of examining the effect of individual characteristics on outcomes, the effect of a deviation in a variable from the group mean on the deviation in the independent variable from its group mean is estimated. For example, to control for the fact that some schools may be better than others in unobservable ways (and that raw measures of the returns to years of education may be confused with returns to quality of education if different quality schools tend to encourage students to stay for different lengths of time) a researcher could study how much more than one’s own schoolmates an extra year of schooling completed relative to the mean number of years completed at one’s own school. This is in contrast to the typical regression of total earnings on total years of education. Common components of the error term (or correlated unobservables) will be the same for each individual in a group and will disappear when the differences in outcomes from the mean are calculated. By observing how relative outcomes move in relation to changes in relative individual characteristics, we can estimate the effect of these characteristics even if they are correlated with a common unobservable term. The cost of this method is that we will be unable to use it to study the effect of mean group characteristics on individual members since these will also vanish when differences are taken. Another disadvantage of this method is that it often requires more data than are available.

3.2.3 *Selection Bias or Endogenous Group Membership*

A third form of endogeneity comes about from the process by which groups and traits are either assigned or chosen by individuals. This specific form of endogeneity is called **selection bias**. The mere fact that an individual is present in a group contains information beyond that summarized in the observed characteristics of that individual. Selection bias is extremely common in all empirical research and the potential for this bias should be considered in evaluating any non-experimental empirical study.¹⁶ In a sense, selection bias leads us to attribute characteristics of an individual to the group they belong to instead of accurately seeing how the characteristics of the individual affected group choice. As an extreme example, consider the fact that many of the people who visit drug dealers subsequently use drugs. This is not because of the negative influence of the dealer on the people who come to see him or her, but reflects the fact that people who visit drug dealers have most likely already decided to use drugs.

If the process by which sorting takes place generates correlations between the error variable and the variable of interest, our estimate of the effect of that variable will be biased. Imagine that children in a study of test scores are sorted by their school into high, medium, and low-ability groups. In this example, ability is the unobserved determinant of test scores, and sorting creates a relationship between group scores and individual, unobserved ability. Even if median class ability had no effect on children’s scores, children with high achieving classmates would also tend to be high achievers. Interpreting this relationship causally would lead us to believe that the presence of high achieving classmates increases an individual child’s score. In fact, the assignment of a child to a high ability class is itself a sign that she is of high ability and would tend to have high test scores regardless of who her classmates were.

Researchers frequently deal with the problem of endogenous group membership by explicitly modeling the selection process. In addition to an equation modeling outcome generation, there would also be a set of equations determining the probability that any individual joins or is assigned to a group. The presence of an individual in a group which, for that individual, is

¹⁶ The effect of selection bias in the estimation of peer effects, in particular, is illustrated by Evans, Schwab and Oates (1992). They find that when parental selection of peer groups is accounted for, much of the impact of neighborhoods and peers on teenage pregnancy vanishes in their sample.

thought to be a low probability choice, provides the researcher with information about the unobserved characteristics of the individual.

For example, if we were studying the effectiveness of private schools, we would have to account for the fact that families that chose private schools may value education more and may help their children more outside of school; the benefit of this support would be incorrectly credited to private education. To correct this bias we could model parents' decisions between private and public schools. Certain families will be more and less likely to send their children to private schools. For example, since private schools are costly, higher income families will be more likely to chose private education. If we observe a low probability family sending their child to a private school we may assume that the family is likely to have an unusually high dedication to education. Intuitively, we could then correct for the selection bias in this model, by including our best estimate, for each family, of the unknown dedication to education. This "dedication" estimate should absorb the influence of parental support and lead to a better estimate of the direct effect of school type.

Other solutions to the selection bias problem include the collection of further data to directly capture the information implicitly contained in group choice, and the use of "instrumental variables". In instrumental variables the predicted group type is included in the regression rather than the observed group. Basing this predicted value on variables that are not self-selected, purges the group measure of its relationship to unknown qualities of the individual. To illustrate, instead of using the school a child actually attends to identify private vs. public sector we could predict private school attendance using family religion, or the proximity of the nearest private school.

Selection bias is extremely common and should be considered whenever groups are not randomly assigned (which, in non-experimental data, is almost always). The reader should ask what characteristics led to the choice of a group by individuals and how these characteristics might also affect the outcome we are measuring.

3.2.4 *Simultaneity*

Simultaneity problems occur when the explanatory variables in the study are themselves determined along with outcomes. Price and quantity are the classic examples of two variables that both effect each other. It is hard to say that a price change led to certain change in the amount of product sold when the two variables are determined **simultaneously** in the market. The failure to consider this simultaneity can lead to confusion as to why we may see that the quantity of a product sold is sometimes highest when prices are also high (think of the year's hottest toy or other faddish products). A simple-minded regression of quantity on price could lead us to believe that high prices increase demand. The mistake here is the failure to account for the fact that prices are themselves affected by quantity and high demand has led to high prices rather than the other way around. Similar problems can be found in the education literature, especially in the presence of peer effects. In this case, the overall quality of institutions such as school and community not only shape children's outcomes, but are in turn determined by the children who attend.

In the presence of peer effects, an individual student's test score is affected by the scores of her classmates but also affects them itself. An exceptionally bright child would have unexpectedly high outcomes and would also improve the performance of her classmates. In turn, a class of high achievers would have high average outcomes and would also lead to unexpectedly high performance by any individual class member. Estimation of peer effects will be clouded by our inability to state the extent to which a child's score is high because her

class is highly capable and the extent to which her classmates' scores are high because she is present.

Simultaneity will not only limit our ability to determine the effect of the explanatory variable that is determined within the system. It will also limit the ability to estimate the effect of variables that are unaffected by the process determining individual outcomes. For example, own parents' education is not caused by a child's test score, nor is mean class parental education, but we may be unable to separately identify the effects of these variables on outcomes.¹⁷

Using our test score example, we can imagine a simple model in which a child's score is a function of her parent's education, the average education of her classmates' parents, and the average achievement (as measured by test scores) of her classmates. For technical reasons, we will be unable to identify an equation with all three variables. In a sense, average achievement is redundant since it is determined by average and individual parental education and these underlying forces are already captured by the included education measures. In the simplified equation that contains only parental and average education levels, the estimated impact on child performance of parents' education will capture more than the direct effect on their own child's test score. Increased parental education will improve class scores in general, which will then feed back to improve the score of the individual child. This feedback through the simultaneously determined variable complicates estimation of the direct effect of own parent's education.

When reading studies on the impact of peer effects, family inputs and community inputs into educational outcomes, the reader should ask which independent variables in the model are not fixed by forces outside the model and what the researcher has done to address these issues. In general, extreme caution should be used in inferring causal relationships simply because a given variable tends to move up or down in tandem with measures of childhood outcomes.

3.3 Conclusion

Many of the statistical issues presented are complex, and a full understanding of the methodological concerns relevant to any given study would be impossible to anticipate and communicate in a broad introduction such as this one. We think, however, that there are a few straightforward questions that all readers should be able to take to studies conducted in this area, questions that will enable them to uncover the broader issues that may be relevant to a particular study. Questioning should begin with the very model itself. The reader should examine the equations estimated to determine the extent to which they reflect the theoretical ideas being considered. In particular, do the **variables** measured truly capture the forces of interest; is there room for the **interaction** of different factors, and can the model capture **threshold effects**? The reader should then ask a few key questions aimed at uncovering the potential for biases in the estimation of the effects of individual variables: (1) are there large categories of factors likely to affect outcomes that have been **omitted** from the model? (2) Are there unknown factors, **other than peer effects**, that might cause the outcomes of a group to **move up or down together**? (3) Do group characteristics independently affect outcomes (is there a true peer effect) or do they simply provide a good **substitute** for poorly measured individual characteristics that contribute to differences in outcomes? (4) Are groups or inputs **randomly assigned**? If not, are members of different groups likely to differ systematically in ways that affect outcomes but that are difficult to observe? (5) Finally, are all the explanatory variables in the model determined outside the system under study and **independent** of the process that determines outcomes?

¹⁷ See Manski (1993) and Moffitt (1998) for a more detailed discussion of the identification problems generated by simultaneity.

Statistics in the social sciences is a tricky business with many potential pitfalls. Those who are not used to the issues that might arise in such analysis may easily place too much credence in assertions that are made in statistical analysis, assertions that often rest on unstated assumptions which, when stated, undermine the most common interpretations. Throughout this section, we have tried to give some indications as to what kinds of problems to watch out for and what questions to keep in mind when critically examining work involving statistics in the social sciences. We think these problems will become clearer as they become more contextualized with the specific examples of literatures we review throughout the rest of this paper. Furthermore, we hope that our review provides some evidence not only of the existence of such statistical pitfalls but also of their potential importance. In many cases, long standing results of large and statistically significant correlations become smaller and less significant as better techniques are applied. In other cases, effects that previously were not identified in the data are found through the application of such techniques. Each of these scenarios represents important progress.

We therefore now turn to a review of the literature investigating the empirical support for the various causal links in Figure 1. Again, the fundamental problem in inferring causation from statistical correlation (even when control variables are used) is that the data used by social scientists usually do not arise from controlled experimental settings. In the rare circumstances when they do, the analysis becomes significantly more simple as it does for some studies we are about to review. In particular, Section 4 begins with the broad question of whether the recent literature has shed any light on the degree to which heredity rather than environmental channels are responsible for linking the two sides in the figure. Because of the unique nature of the data used to analyze this question, data that is almost like data from a controlled experiment, many of the methodological issues do not arise in this literature. We will, however, refer back to these issues frequently in Sections 5 through 7 where such experiments are harder to come by.

4 Heredity versus Environment as a Causal Link

This section examines evidence on the role of hereditary factors in determining cognitive ability, achievement, and other child outcomes. The exact portion of outcome variation that can be explained by genetic factors may seem to be of limited policy relevance; policy makers cannot affect genetic endowments and certainly would not abandon attempts to increase children's outcomes even if some of the underlying difference is due to heritable traits. Research into the genetic links between parents' attributes and children's outcomes is necessary, however, for three main goals: (1) to determine the magnitude of what we know and don't know about the determinants of children's success; (2) to estimate the potential effectiveness of policy interventions; and (3) to even approach accurate estimation of the impacts of the environmental channels that policy can affect.

At the outset, we note that links between genetics and inter-group (or racial) differences in intelligence are exceedingly tenuous. For example, there is evidence that IQ is partly determined by genetic factors, and that racial group differences exist on IQ measures. However, it would be a logical fallacy to infer that genetic differences among races are directly determining IQ differences. Despite the apparent controversy of examining the links between heredity and child outcomes, it is nonetheless imperative to do so. By completely omitting a consideration of heredity, empirical research is likely to either over- or under-state the effects of certain environmental factors. In turn, the policy advice based on such research will inevitably be misleading.

Much of the best research in this area relies upon samples of twins, siblings and adopted children. If genetics plays an important role in determining outcomes, then we would expect, for example, similar outcomes across twins who are raised apart, despite their varying environments. In contrast, differing outcomes would imply a greater role for the home environment. Similar inferences can be drawn when comparing siblings and adopted children raised in the same environment, or when comparing various other pairs of children where the degree of genetic variation is known. A review of this kind of evidence suggests four conclusions. First, heredity seems to play a large role in the formation of characteristics such as personality traits and cognitive abilities, with estimates of this genetic effect accounting on average for approximately 50 percent of differences across cognitive abilities in children. Second, genetic effects on scholastic achievement are somewhat similar to genetic effects on cognitive ability or IQ. Third, heredity is of increasing importance relative to childhood environment as individuals transition from childhood into adulthood and beyond. Finally, a surprisingly small fraction of differences in adulthood (0-10 percent) can on average be causally attributed to shared home environments, at least within the ranges of families studied.

There are some caveats to the interpretation of these findings. For example, some studies are based on small samples, and some are based on samples in which twins volunteer to participate, implying possible biases. Despite these critiques, it is nonetheless striking just how consistent the results are across many studies that use different samples and methods. The rather consistent findings on hereditary influences might still be misinterpreted in several ways, however. For example, the findings do not necessarily imply that policy will be unable to improve outcomes for disadvantaged children, whether the disadvantage is genetic or environmental. Similarly, it would be unwarranted to broadly conclude that "parent don't matter." To provide one illustration of this point, the existing research applies mainly to a range of "average" parents, parents whose households may differ greatly in terms of income, parenting styles, family structure, etc. but where extreme neglect or attention or extreme poverty are not typically observed. For parents outside this range, it is quite feasible that shared home environments do explain differences in outcomes. Furthermore, parental efforts may well matter in an absolute sense without differences among them accounting for differences in outcomes. For example, if the large majority of parents provide healthy

environments for their children, even if they do so in very different ways, then their efforts may be important without accounting for large differences. Finally, while these studies are certainly informative, their precise meaning remains unclear so long as the way in which environments and genes interact is not better understood.

The multitude of correlations between parental and child characteristics has led to the fairly common practice among social scientists of using “family background” as a control variable when attempting to isolate particular empirical phenomena. Consequently, statements such as “holding fixed family background, variable x is found to contribute significantly to child outcome y” are ubiquitous. The mechanism responsible for this statistical importance of family background, however, can be of two particular types. Family characteristics may be genetically transmitted from parents to children (as, for example, eye color), or they may be transmitted through the environment that children with particular family backgrounds have in common. As argued in Section 2, the “environment” that is relevant may include features of the home, the neighborhood, the school or some other aspects of relevant “communities.”¹⁸

This tension between nature (heredity) and nurture (environment) has a long and not always distinguished history in social science.¹⁹ Given the correlation of genetic and environmental factors—high IQ parents tend to have higher incomes, for instance—it has been difficult to isolate heredity from environment, and often social scientists have implicitly assumed heredity to play no role at all, a practice sometimes dubbed the “sociological fallacy.” The most persuasive research in this direction is that involving monozygotic twins (whose genes are, by definition, identical), fraternal twins (who share 50% of genes) and/or adoptive children (whose genes are not derived from adoptive parents). Several striking conclusions emerge from this literature. First, heredity seems to play a large role in the formation of characteristics such as personality traits and cognitive abilities, with estimates of this genetic effect accounting on average for approximately 50 percent of differences across cognitive abilities in children. Second, genetic effects on scholastic achievement seem to mirror but are not identical to genetic effects on cognitive ability or IQ. Third, contrary to intuitive speculation, heredity is of increasing importance relative to childhood environment as individuals transition from childhood into adulthood and beyond. Finally, while the research on twins and adopted children leaves roughly half of the differences in personality and cognitive ability to be explained through environmental factors, a surprisingly small fraction can on average be causally attributed to shared home environments, at least within the ranges of families studied.

We caution the casual reader that each of these four conclusions is easily subject to overstatement and misinterpretation and that the relevance of these findings for both researchers and policy makers requires careful thought. In Section 4.1 we therefore say a few words about the controversy surrounding the issue of heredity in the social sciences. In particular, we address the common misconception that a finding of genetic explanations for inter-personal differences such as intelligence or personality imply that inter-group differences are also genetically based. Then, in Section 4.2, we explore the potential danger of ignoring heredity in an attempt to be uncontroversial and discuss briefly the kinds of research

¹⁸ The inability of social scientists to identify what exactly is represented by “family background” has long been recognized. In one influential study of two decades ago, for instance, Scarr and Weinberg (1978) use a sample of adoptive and biologically related families with children in their late teens and early twenties to isolate genetic influences on child IQ from parental environment. Their conclusion, surprising at the time, supports the hypothesis that much of the “family background” effect is genetic.

¹⁹ Much of the early work on this topic, especially as it relates to genetic transmission of intelligence, was heavily tainted by racist and otherwise ideologically driven agendas, and little scientific credence can therefore be given to this work which will go uncited here. More recent speculations concerning a large genetic component to differences in IQ scores for different races (Jensen 1969; Herrnstein and Murray 1994) remain largely unsupported by the evidence (Sowell 1994; Heckman 1995; Brooks-Gunn, Klebanov et al. 1996).

and policy bias potentially introduced by ignoring genetic factors. Section 4.3 then proceeds with a more thorough review of the evidence in support of the four main conclusions cited above.

4.1 *Heredity, Race and Social Science*

It should be noted at the outset that the finding of a sizeable genetic component to cognitive ability is quite distinct from claims of genetic differences between racial and ethnic groups. While Herrnstein and Murray (1994), for instance, convincingly document evidence of genetic influences on IQ as well as differences in average IQ scores for different races, it is a logical fallacy to combine these two findings and conclude that racial differences in IQ are in large part genetically driven.²⁰ The fallacy is easily illustrated: it is well known among biologists, for instance, that the height of wheat is due in large part (but not exclusively) to genetics. Suppose then that two identical crops of wheat are planted in two different plots of land, one full of nutrients conducive to the growth and development of wheat and the other lacking in such nutrients. Since the resulting crops grown on these different plots spring from genetically identical wheat, any observed difference in the average height of the crops is entirely environmental (i.e., due to the different quality of the land the two crops are growing on). Thus, despite the fact that genes determine a large part of height, group differences in wheat height in our thought experiment are 100 percent due to environmental differences. Similarly, the joint finding of a 50 percent genetic component to IQ and of racial differences in average IQs has nothing whatsoever to say about genetic differences between races (i.e., 100% of the difference in average IQs between groups can be due to environmental factors even if 50% of IQ is genetically transmitted). Neither Herrnstein and Murray nor anyone else, as far as we have been able to determine, have provided consistent and credible evidence that this is not the case. Furthermore, others have illustrated that many racial differences in cognitive ability both in the U.S. and New Zealand disappear when environmental factors are taken into account, just as none of the differences in wheat height in our example can be explained through genetics when the environmental factors are properly taken into account.²¹

As suggested by Plomin (1997), social scientists have often shied away from the topic of heredity in large part because of the controversy surrounding ethnicity and race. When raised with a recognition that a finding of hereditary effects has no implications for differences between races, however, the topic is typically less controversial. A good example of this is the wide publicity surrounding the recent work by Harris (1998). While much controversy surrounded her work, little attention was given to her claim that 50 percent of the variance in personality traits is genetically driven. Similarly, Jencks (1992) feels free to acknowledge an important role for heredity in general while arguing that it is unlikely to have much to do with racial differences. The lack of controversy surrounding such writings stands in stark contrast to the furor surrounding the publication of *The Bell Curve* that claimed, without much solid evidence to support it, that racial differences in IQ were probably genetically driven.

²⁰ Murray and Herrnstein themselves are quite aware of this fallacy but nevertheless speculate that racial differences might be genetically based.

²¹ Many criticisms of Herrnstein and Murray's conclusions on race exist (Goldberger and Manski 1995; Heckman 1995; Devlin, Fienberg et al. 1997), and a thorough review of all these is beyond the scope of this review. In the case of African-American children, economic and social differences seem to account for virtually all of the differences in IQ scores (relative to Caucasian children) (Brooks-Gunn, Klebanov et al. 1996). With respect to New Zealand, increasing evidence from the Christchurch Health and Development Study also indicates that observed group differences in cognitive ability, school achievement and youth crime between Maori and Pakeha children can be explained through environmental factors (Ferguson, Lloyd et al. 1991; Ferguson, Horwood et al. 1993; Barker and Maloney 1999), and that at least some measures of intelligence do not differ between Maori and Pakeha children even without such controls.

In principle, it is of course possible that part of inter-group differences between races is genetically driven. Our main point here, however, is that this is not a conclusion supported by the behavioral genetic studies we review, studies that generally do not have sufficient data to address such questions. In fact, some of the more convincing studies of the kind reviewed in this section suggest an almost entirely environmental channel for observed racial differences in average IQ.²² Furthermore, the literature is full of examples of stark group differences in average IQs that vanish within one or two generations, especially in the case of quite different immigrant groups to the US such as Irish and Italian immigrants in the early part of this century and West Indies immigrants later on (Sowell 1994). In many of these cases, observers at the time blamed average inter-group IQ differences on genetics only to be refuted by the quick transformation of these groups within short periods of time. There is no particular reason to believe that current observed group differences, whether these involve American blacks or New Zealand Maori, are particularly different, although we argue later on in this review that environmental factors may prevent convergence of mean IQ levels for some of these groups.

4.2 *Relevance of Heredity for Social Scientists and Policy Makers*²³

In addition to being clear that we make no claims regarding genetic causes of racial differences, we also think it important to point out that ignoring heredity entirely for the sake of being uncontroversial may lead to a serious misunderstanding of the empirical literature with which this review is concerned. In particular (and as we point out frequently below), the methodological challenges of identification such as omitted variables bias are likely to be important, at least for some of the factors we study. As we will point out again and again, the hesitancy of social scientists to explicitly include heredity in their analysis or in their interpretations of findings has therefore probably led to many biases in statistical analysis, and such biases have entered the policy debate. If heredity is indeed important, no one is served in the long run by excluding it from analyses.

For example, suppose that income is partially produced by cognitive ability which in turn is partially heritable. Suppose further that scores on student achievement tests are also causally linked to cognitive ability. Researchers investigating correlations between family income and student achievement will then find strong correlations of the two, but these correlations are not necessarily indicative of a causal link because income and student achievement are both in part caused by cognitive ability which is genetically transmitted from parents to children. In other words, the observed correlation between income and student achievement may simply mask the underlying fact that parents with higher cognitive abilities have both higher incomes and children with higher cognitive abilities, children who then score well on achievement tests. Research bias of this kind can appear in virtually all the studies that do not take seriously the potential for genetic links, and we will repeatedly emphasize this possibility as we review some of the literature below. Only for some issues, in particular issues surrounding the returns to schooling, have we found serious recent attempts by social scientists to see whether such biases exist.²⁴ While the consensus in the returns-to-schooling literature seems

²² In a study of illegitimate children fathered by US soldiers during World War II in Europe (and raised by their mothers in Europe), for instance, children of black servicemen were found to demonstrate average IQs no different from children of white servicemen (Flynn 1980). For a more comprehensive account of this issue, see Nisbett (1997). For references particularly relevant to New Zealand, see those cited in the previous footnote.

²³ The concept of heritability and its usefulness have been debated in the past (Goldberger 1979; Taubman 1981), and this debate continues today (Devlin, Fienberg et al. 1997). The issues raised in this debate are the ones we attempt to address in this section.

²⁴ Early analysis using twins data had suggested substantial bias from leaving out innate ability from the analysis (Taubman 1976; Taubman 1976; Behrman and Taubman 1989), while more recent studies have found smaller biases. These studies have used twins data from various countries including the U.S. and Australia (Ashenfelter and Krueger 1994; Miller, Mulvey et al. 1995; Miller, Mulvey et al.

to increasingly be that commonly used variables for family background are sufficient to adjust for genetic background, we will argue in Section 5 that the evidence in many other areas is suggestive of more significant biases from ignoring heredity.

For similar reasons, policy makers should also be concerned about the implications of heredity. In particular, if policy is to be guided by social science research, then research bias of the kind discussed above will lead to ill-advised policy proposals informed by incorrect conclusions regarding causal links in the data. In our example above, for instance, researchers who ignore genetic effects may infer a causal link between family income and student achievement despite the fact that the underlying cause of the correlation between the two is that cognitive ability within families partially causes both high incomes and good student achievement. Policy makers may then use such research to justify income subsidies for families whose children are not achieving in school, all with the intention of raising student achievement. While redistribution may be desirable for all kinds of other reasons, the social science research that ignores heredity (under the assumptions of the example above) incorrectly gives policy makers the impression that it will also benefit student achievement when in fact it might not. Research bias from excluding heredity in analysis of data is therefore easily translated into policy bias.

At the same time, the existence of strong hereditary forces does not have to introduce complacency into the policy arena. While policy makers clearly have little control over hereditary factors themselves, it is quite conceivable and even likely that hereditary disadvantages can at least in part be overcome with environmental changes (unless all outcomes are 100 percent genetically predetermined.) As Plomin and Petrill (1997) put it, “evidence of genetic influence does not imply that differences are immutable or irremediable.” In some sense, it may well be irrelevant for policy makers whether particular disadvantages stem from genetic or environmental causes so long as the same environmental improvements can successfully ameliorate both types of disadvantage (Jencks 1992). The finding of hereditary factors thus re-enforces the need for a better understanding of environmental factors that may positively impact child development. At the same time, a better understanding of hereditary factors may also unveil the limits of social policy aimed at equalizing outcomes. One has to be extremely careful, however, in jumping to quick conclusions regarding the potential for policy to have real impacts even if heritability estimates are high. The mere presence of such heritability has, by itself, very little to say regarding such limits (Goldberger and Manski 1995).

Finally, some might argue that a finding of genetic explanations for inter-personal differences in the absence of genetic explanations for inter-group differences renders the discussion of genetics less important for policy makers who may be concerned precisely about group differences (e.g., Maori and Pacific Island outcomes in New Zealand, black and latino test scores in the U.S., etc.). We disagree with this view on two counts: First, if reducing or eliminating group differences in child outcomes is an important policy goal, it remains vital for policy makers to understand the environmental factors underlying inter-personal differences. After all, it is a proper understanding of those factors that must form the basis for any policy that aims to reduce both individual and group differences through environmental channels. And if research bias of the kind discussed above introduces policy bias, policies intended to reduce group differences may not have the desired effect. Second, while bordering on philosophy, we find it doubtful that policy makers genuinely care only about group differences to the exclusion of aiding under-performing children in groups that do well on average. Policies aimed at raising the performance of under-performing children in general

1997; Ashenfelter and Rouse 1998). Debate regarding the issue continues, however, as evidenced in an entire section of a recent journal (Behrman and Rosenzweig 1999; Bound and Solon 1999; Neumark 1999; Rouse 1999).

will naturally have the impact of reducing group differences, given that groups whose children have disproportionately worse outcomes would benefit disproportionately.²⁵

4.3 *Heritability versus Environment: Evidence from Behavioral Genetics*

Because of the complex interplay between environmental and genetic components of child development, much of what we think we know regarding the role of genetic influences relative to environmental ones comes from studies of twins and adopted children. Such studies are almost like controlled experiments in that they can disentangle genetic and environmental sources of family resemblance in a strikingly straightforward manner. If heredity plays a large role in contributing to family resemblance, for instance, then genetically related individuals reared apart should be similar to one another. Furthermore, genetically related family members living together should resemble one another more than adoptive members reared in the same environment, and identical twins should be more similar than fraternal twins. The data from studies of twins and adopted children, as well as data from wider studies including cousins, are remarkably consistent with all of these implications. In fact, not only are the data qualitatively suggestive, they can also be used to quantify what fraction of the observed variance of particular traits in the population can on average be attributed to heredity. We will focus in this section on hereditary components of intelligence and scholastic achievement, but we note in passing that similar results are emerging in relation to personality traits (Rowe 1994; Harris 1998). Furthermore, we should note the evidence of the kinds of correlations we find in siblings and twins studies is almost certainly to a large extent causal as such evidence comes from unusually controlled natural experiments. Some potential caveats to this are discussed in Section 4.3.2 after we review the evidence itself in Section 4.3.1.

4.3.1 *Quantitative Estimates of the Role of Heredity*

The field of behavioral genetics has made significant advances in the past decade, just as the field of DNA research seems poised to do in the near future. Much of the research has concerned itself with the genetic determinants of cognitive ability (as typically measured by IQ scores or some similar test), but some has expanded beyond this focus to find measures of heritability of specific cognitive skills, of school achievement, of juvenile crime and of personality traits. Furthermore, insights regarding non-genetic determinants of cognitive ability in particular have emerged. One feature of this literature that is remarkable to the outsider is how relatively little controversy it raises among behavioral geneticists and how relatively unnoticed it is by social scientists who are concerned about many of the same issues. Much of the research on cognitive ability is reviewed thoroughly and accessibly by Plomin and Petrill (1997), and we refer the reader to their review for more details than we are able to provide here.²⁶

Plomin and Petrill (1997) use and summarize data from twins and adopted child studies to arrive at quantitative estimates of the role of heredity in determining differences in IQs. Their survey is based on a large number of studies in behavioral genetics, and a full review of each of these studies is beyond the scope of our review. However, because of the special nature of the data used in these studies, in particular because of the fact that much of the data comes close to resembling data arising from relatively closely controlled experiments, statistical analysis is quite straightforward, generally accepted, and difficult to dispute. Furthermore,

²⁵ Issues related to heredity also arise in policy discussions of intergenerational mobility as well as economic growth. For a recent theoretical analysis of some of these issues, see Zak (1999).

²⁶ A review of heritability of outcomes like juvenile crime is beyond the scope of this paper. For recent work indicating rather high levels of heritability, see Wilson (1985), and for treatments of personality see references in Harris (1995; 1998).

while the general methodology is often similar across studies, it is applied to a variety of different types of data: identical twins separated early in childhood, identical twins not separated, non-identical twins, adopted children, related siblings and more recently wider kinship relations like cousins. As we discuss more fully in Section 4.3.2, the relative consistency of the results across all these data gives us reasonable confidence that the features that keep each particular data set from having been truly experimental are not biasing the results appreciably. Discussion of other caveats is saved for Section 4.3.2.

4.3.1.1 Heritability of IQ

The most basic question addressed in these studies regards the extent to which cognitive ability among children and adolescents is heritable or attributable to genetic factors.²⁷ Here, the degree of quantitative consistency found across studies is indeed striking and has been accepted for some time (Snyderman and Rothman 1988). For example, approximately 48 percent of IQ variance among adopted children is heritable,²⁸ whereas the same number arising from data comparing fraternal and identical twins raised in the same home is 52 percent.²⁹ A somewhat different methodology involving a comparison of adoptive and genetic relatives yields a slightly more modest estimate of 40 percent heritability, while heritability estimates from small samples of identical twins reared apart reach as high as 72 percent.³⁰ Statistical analysis that merges family, adoption and twin data, however, yields estimates of heritability closer to the 50 percent mark (Loehlin, Horn et al. 1989; Chipuer and Plomin 1990). In addition, newer regression approaches are expanding the quantitative search for genetic links in wider kinship data (such as cousins) and again arrive at almost precisely the same 50 percent estimates (Rodgers, Rowe et al. 1994). Although the scientific “consensus” on the question of heritability of cognitive ability has swung between one extreme and the other throughout much of this century (Plomin and Petrill 1997), the accumulated evidence now seems to be converging to a more stable consensus among those who actually conduct research in this area: roughly 50 percent of observed differences in IQ and cognitive ability when assessed in childhood or adolescence seems on average due to heredity.³¹ Estimates of higher heritability reported by Herrenstein and Murray (1994) are probably overstated for various reasons and have been criticized on a variety of grounds. One excellent critical analysis (Daniels, Devlin et al. 1997) explores the variety of technical issues surrounding this debate, and concludes that heritability of IQ is likely to lie between 43 and 54 percent.³²

²⁷ Heritability itself is a statistic that gives the proportion of observed differences between individuals that can be attributed to genetic differences among them.

²⁸ Biological relatives adopted into separate homes are said to “correlate” approximately 24% in their measured IQ, which means that 24% of the IQ variance in the population covaries for relatives who do not share the same environment but are 50% genetically similar. Heritability is then twice this correlation, or 48%, because it represents the degree of variation attributable to heredity with 100% rather than 50% genetic similarity. This method is statistically sound so long as environments are not consistently similar for adopted children (Plomin and Petrill 1997).

²⁹ Differences between fraternal and identical twins reared in the same home can be used to arrive at this number. Identical twins (who are 100 percent genetically identical) correlate approximately 86% while fraternal twins (who are 50% genetically identical) correlate 60%. Multiplying the difference in these correlations by two gives the estimate of heritability: 52%.

³⁰ The substantially larger estimate using identical twins is often cited but tends to be attributed in part to prenatal environmental factors (Bouchard 1990; Pedersen 1992), not strictly genetic ones.

³¹ Further references on similar studies confirming these estimates can be found in Plomin and Petrill (1997) and Harris (1995; 1998).

³² This analysis is part of a thorough response presented by a large number of scientists from various disciplines to *The Bell Curve*. Since the topic of this review is only tangentially related to *The Bell Curve*, we forego a thorough summary of all the possible criticisms of Herrenstein and Murray’s analysis and refer the interested reader to this volume (Devlin, Fienberg et al. 1997). We note in passing that this volume agrees with our assessment above that Herrenstein and Murray’s suggestion that racial differences in IQ are due to genetics is “mere speculation.” We also note that even within these studies, studies that are explicitly aimed at criticizing Herrenstein and Murray’s emphasis on

One common element of most of these studies is, as just stated, that they assess cognitive ability in childhood or adolescence. It seems natural to suppose that environmental factors increase in importance during the life span relative to genetic factors, and it has thus sometimes been assumed that these estimates of heritability measured early in life represent an upper bound of the heritability of cognitive ability when assessed in adulthood. Evidence brought to light in the past decade, however, increasingly suggests the opposite. Plomin and Petrill (1997) report heritability (from studies similar to those described above but conducted at different ages) increasing from 40% in early childhood to 60% in early adulthood and to as much as 80% later in life. One possible interpretation offered by these authors is that there may be genetic pre-dispositions directing individuals toward environments that accentuate genetic propensities, a hypothesis we return to later. Another interpretation is that heritability manifests itself in the form of genetic triggers that “kick in” over the course of a lifetime. Heritability estimates of cognitive ability when measured in adulthood are therefore generally higher than 50 percent, but we remain uncertain of how to interpret this insight of a growing consensus that genes and environments probably interact in quite complex ways.

4.3.1.2 Is Heritability of IQ Informative

Since the bulk of the evidence on heritability has involved research on tests of cognitive ability such as IQ, we might ask to what extent heritability of IQ is representative of heritability of other types of outcomes. Research by behavioral geneticists has only recently begun to focus on such questions, and three preliminary findings seem to be emerging: First, not all aspects of what is commonly called “intelligence” or “cognitive ability” are equally heritable. This implies that some dimensions of cognitive ability can more easily be influenced by policy induced environmental changes. Second, to the extent different cognitive abilities are heritable, the same genetic factors seem to influence these various abilities. Third, and perhaps particularly interesting for readers of this review, genetic effects on scholastic achievement scores overlap substantially with genetic effects on intelligence.

Plomin et al. (1997) provide one of the most recent and extensive studies on the heritability of various aspects of cognitive abilities using a relatively large data set including adoptive children, their biological parents and their adoptive parents. They replicate a heritability estimate for general cognitive ability (using a standard IQ test) close to that reported above (56 percent), but then proceed to estimate heritability for four sub-categories of cognitive ability. Somewhat surprisingly, the most heritable of these was verbal ability (54 percent), followed by spatial ability (39 percent), speed (26 percent) and memory (26 percent). While these results have not been replicated to the degree that those on heritability of general cognitive ability have, they are indicative of differences across different types of cognitive functions. This research may lead to solid policy implications as less heritable dimensions of cognitive ability are identified.

An immediate question raised by these results is, of course, to what extent genetic pre-dispositions for one aspect of cognitive ability are correlated with genetic pre-dispositions in other dimensions. The methodology used to investigate this question is known as multivariate genetic analysis. Rather than focusing on the variance of a trait as was done in the studies reviewed above, such analysis focuses on the covariance between traits. In the previous set of studies, twin data, for instance, was used to assess the heritability of IQ by comparing how one twin’s IQ correlates with another’s. In multivariate genetic analysis, the researcher would instead compare a particular cognitive ability, say spatial ability, of the first twin with a different cognitive ability, say verbal ability, of the other twin. As before, if such correlations are greater for identical twins than for non-identical twins, genetic influences can be inferred.

genetics and IQ, the consensus is still that IQ is heritable on the order of somewhere between 43% and 54%.

Similarly, the other types of data (adoptive relationships, non-twin siblings, etc.) can also be used in this fashion. One interpretation of a finding of genetic overlap is that, if a specific gene were found that is associated with one of the cognitive abilities investigated, the same gene would also be expected to be associated with the overlapping ability (Plomin and Petrill, 1997). The general conclusion from such multivariate genetic analysis has been that different cognitive abilities do seem to be influenced by similar genetic factors.

Finally, the same multivariate technique has been employed to investigate the extent to which scholastic achievement scores are influenced by the same genetic factors as cognitive ability or IQ scores. In light of the fact that so much of the evidence on the heritability of child outcomes is based on IQ tests, this becomes a particularly important question. Plomin and Petrill (1997) report on several studies (Thompson 1991; Wadsworth 1994) that have replicated the same general result several times: genetic effects on scholastic achievement overlap almost completely with genetic effects on intelligence. This, combined with a finding of heritability of scholastic achievement of magnitudes roughly similar to that of intelligence (Plomin and Petrill 1997), implies that the roughly 50 percent of the variation in scholastic achievement that is directly attributable to genetic influences is in fact influenced by essentially the same genetic factors that influence intelligence. Of course, since about fifty percent of both intelligence and scholastic achievement are still caused by non-genetic factors, plenty of room remains for the classic under- and over-achiever (highly intelligent children that perform poorly on scholastic achievement tests, and low IQ children that outperform their high IQ peers).³³ However, the near complete overlap of genetic effects on intelligence and scholastic achievement indicates that “under-achievement” and “over-achievement” are phenomena almost entirely due to non-genetic factors. Nevertheless, cognitive ability remains among the most consistent predictors of scholastic achievement, and its inclusion in regression analysis tends to diminish or eliminate the impact of socioeconomic family and parental education variables (Teachman 1996).³⁴

4.3.1.3 Behavioral Genetics and Home Environments

Perhaps the most surprising insights emerging from the research on twins, siblings and adopted children, however, involves the role of shared environments in determining cognitive abilities (and also personalities). It has long been assumed that growing up in the same home environment would cause siblings to become more similar over time, but precisely the opposite seems to be the case. In the transition from childhood to adolescence, shared home environments, in particular those relevant for cognitive development but perhaps also for other traits like personality, seem to contribute to making family members more different. More precisely, while shared home environments seem to have relatively large power in explaining similarities among siblings in early childhood, they lose this power as siblings grow up. As Plomin and Petrill (1997) put it: “Genetically related children growing up in the same family are similar for both genetic and environmental reasons, but by adolescence heredity accounts for this familial resemblance” (p. 64).

Earlier studies of adoptive children provided strong evidence for a large influence of shared home environment on IQ, an effect of approximately 25 to 30 percent. While many studies still cite this work as evidence in favor of important shared home environment impacts on cognitive developments, it now appears that these estimates were due entirely to the fact that measurements of cognitive abilities were done when the subjects were still relatively young children. In a follow-up study of 181 adoptive siblings, the correlations in IQ used to arrive at the 25 to 30 percent estimates when children were approximately 8 years old had vanished entirely by the time those same children reached an average age of 18 (Scarr and Weinberg

³³ In fact, the two measures are correlated by only 50 percent (Plomin and Petrill, 1997).

³⁴ In the New Zealand literature, for instance, Baker and Maloney (1999) find strong correlations between cognitive ability and reading ability in the Christchurch Child Development Study.

1978; Loehlin, Horn et al. 1989). In another striking study, Plomin et al. (1997) replicate these results in an ambitious 20 year study. Their data, however, are not only on adopted children but also on both their adoptive and their natural parents. As the adoptive children grow into adolescence and adulthood, they gradually resemble their adoptive parents less. Furthermore, and even more surprisingly, they become more like their biological parents (in cognitive ability) and at a pace similar to that found in a control group of non-adoptive children. Not only does this study suggest that shared home environmental factors are of less significance in contributing to cognitive ability on average than previously thought, they also suggest that shared home environment factors that do contribute to cognitive development are not correlated with parental cognitive ability. This is an important finding in that it should shift attention away from aspects of home environments that are correlated with parental cognitive abilities and towards other home environment factors (such as parental warmth and support) that may be less correlated with parental cognitive abilities. However, given that other studies find relatively little influence of shared home environment in general, even those factors may not be of great importance in causing cognitive development in most homes. These insights are beginning to be confirmed in broader statistical analysis that demonstrates a correlation between shared environment in siblings with cognitive ability in childhood but not in later life, while measures of non-shared environment show little correlation with cognitive ability early in life but more later on (Chipuer and Plomin 1992). One tentative conclusion arising from this work is that some genetic factors that determine long run cognitive abilities may not come into play until middle to late childhood (whether through a genetic “trigger” or through genetically influenced choices of environments).³⁵ Finally, the fact that correlations between fraternal twins are higher than those for non-twin siblings (0.60 versus 0.47) indicates that shared environmental factors (perhaps as early as the womb) contribute more to the resemblance of twins than to non-twin siblings (Plomin and Petrill 1997).

Additional evidence regarding the impact of shared home environments has been provided by an emerging literature using more traditional social science methods applied to data sets of twins. Taubman (1976) uses data on twins born between 1917 and 1927 (collected by the U.S. National Academy of Sciences and the National Resource Council) and finds only a small role for shared environments in explaining earnings and schooling at the approximate age of 50.³⁶ Similar findings arise in more recent studies on Australian twins (Miller, Mulvey et al. 1996).

³⁵ One important side effect of this research is an increasing awareness of how previously used variables of home environments actually seem to include genetic factors and may be leading social scientists in general to over-interpreting regression coefficients associated with those variables. Many past studies by psychologists, for instance, use self-reported data on home environment to illustrate environmental impacts on cognitive abilities. Recent studies of such reports (Hur and Bouchard 1995), however, compare sets of identical and fraternal twins’ perceptions of home environments to find that a large portion (around 45%) of the variance in such perceptions is in fact attributable to heredity. Thus, unless parents consistently treat children dramatically differently, heredity in part shapes perceptions of home environments, and statistical tests finding a positive correlation between such variables and cognitive outcomes cannot be used to infer an environmental effect. Other measures of home environment seem to be similarly affected by heredity, most particularly the widely used HOME (Home Observation for Measurement of the Environment) measure. Plomin and Petrill (1997) report that approximately half of HOME’s prediction of individual differences in cognitive development can in fact be accounted for by genetic factors. They conclude that the “more profound meaning of finding genetic influence on measures of the environment” is that “genes contribute to experience itself” (p. 66).

³⁶ Estimates of genetic contributions range as high as 50% for income and 46% for schooling, and nonshared environments account for around 46% for income and 24% for schooling. In a related study, he also finds that typical family background variables do not sufficiently proxy for genetic factors (Taubman 1976), a finding that has been disputed by the more recent literature using more recent data sets on twins (Ashenfelter and Krueger 1994; Miller, Mulvey et al. 1995; Miller, Mulvey et al. 1997; Ashenfelter and Rouse 1998).

Behavioral geneticists (and others) have therefore presented persuasive evidence that environmental characteristics shared by children within a home account on average at best for a small share of the total environmental component causing variances in intelligence and achievement. Estimates of this average influence of shared home environments on variances in cognitive outcomes range between 0 and 10 percent (Loehlin 1992; Bouchard 1994), with increasing evidence that it may actually lie closer to the lower of these. What remains unclear, however, is precisely which non-shared environmental factors account for the remaining bulk of the unexplained variance in cognitive abilities and under what circumstances these average findings for home environment do not hold. Even so, while we postpone much of our discussion of this until later, we need to emphasize, as has been done most recently by Harris (1998), just how strong a statement regarding environmental influences has been delivered by the research thus far reviewed in this section. The behavioral genetic evidence on twins, siblings and adopted children suggests that only those environmental factors within average homes that are not shared by children can be of significant causal importance for cognitive development within the samples of families investigated, or alternatively that children experience the same home environment quite differently. Note that the behavioral genetic research does not offer evidence that such non-shared home environmental factors are in fact important, but their evidence simply leaves room for the possibility. However, it excludes shared environmental components such as parental personalities, broadly applied parenting styles, and the presence of public goods within the home—all those factors psychologists, sociologists and economists have reported as important in fundamentally shaping child outcomes, at least as long as these shared environments are perceived similarly by children of different genetic backgrounds. We return in Section 8 to the possibility that such perceptions might differ, and how this might carry implications for the social sciences.

4.3.2 Caveats and Cautions: What Do These Studies Really Tell Us?

While we have argued that the studies conducted by behavioral geneticists are quite unique in their design and thus not subject to many of the criticisms we levy against the more mainstream social science literature reviewed in the rest of this paper, the data sets on twins, siblings and adopted children are not entirely equivalent to data derived from randomized experiments. As a result, a few caveats are in order. Furthermore, it is easy to caricature this literature with statements like “parents don’t matter” or “home environments don’t matter”, caricatures which we think are not warranted and highly misleading. We therefore conclude our discussion of heredity with a brief discussion of the statistical caveats in Section 4.3.2.1 and some cautions in over-interpreting these studies in Section 4.3.2.2.

4.3.2.1 Caveats Regarding the Validity of the “Experiments” in Behavioral Genetic Studies

None of the many studies that underlie the basic conclusions from behavioral genetic studies reported above is “perfect” in the sense that no possible alternative explanations can be offered for the observed results. For example, twins, even those separated at birth, share the environment of the womb, and we know from medical evidence that this environment is crucial to the healthy development of a child. This probably accounts at least partially for the unusually high estimate of heritability of IQ for twins separated at birth. Similarly, twins who volunteer to participate in research studies may be more alike than those that do not, and adoption agencies may be particularly sensitive in placing twins into similar home environments. These factors, too, would contribute to an overestimate of heritability. Furthermore, even in the largest study on identical twins separated at birth, the sample size was a meager 95. Were the results from behavioral geneticists entirely based on one kind of study, identical twins separated at birth for instance, we would therefore naturally be quite skeptical.

What is remarkable, however, is that results, especially those for heritability estimates of cognitive ability, are, by and large, quite consistent across very different kinds of studies using very different kinds of data. Criticisms involving small sample sizes can be levied against some but not all of them. Charges of bias from unaccounted pre-natal environments are relevant for twins but not as much for other siblings and certainly not for adopted children. Procedures employed by adoption agencies may create bias in some studies but are largely irrelevant in many others. Some studies required the voluntary participation of subjects and may involve some responder bias, but others were based on data collected without such cooperation. Even the possibility that parents are genetically programmed to treat adoptive children differently than natural children is relevant only for certain types of comparisons and plays no role in others. Still, the results on heritability are usually very similar and consistent. It is therefore difficult to believe that the thrust of the results reported above has anything much to do with these potential biases, especially when one focuses on the more conservative estimates of 50 percent heritability.

A more subtle caveat regards the question of what precisely these heritability estimates mean. In one sense, Jencks (1992) points out that studies such as some of those we report technically find an upper bound of how much variation among individuals is due to genetics. This is because certain genetic characteristics may cause children to be consistently treated differently which may then affect their development. If, for example, two twins are both physically beautiful, and if adults consistently treat beautiful children better than other children, then some of the similarities the studies attribute to gene similarities are hereditary only indirectly: the genes that produced beauty caused environments to be consistently more pleasant. For practical purposes, however, this distinction seems unimportant as policy-makers are unlikely to find means of altering how such characteristics as beauty affect adult behavior. Furthermore, non-twin studies simply are not subject to this observation, yet such studies still find similar estimates for heritability.³⁷ In another sense, however, even more than what is found in these studies could be due to heredity if one takes the view of some recent evolutionary psychologists. Tooby (1992), for instance, argues that the distinction between heredity and environment is somewhat artificial because what is inherited is likely to be substantially more complex than previously imagined. In particular, the human mind may be a collection of many types of “triggers” that are set off by environmental conditions, which implies that the environment may be important in setting off a trigger that is genetically programmed. Thus, the distinction between environment and heredity becomes blurred. Again, however, the dimensions added to the debate by this perspective seem largely irrelevant from a policy perspective: whether a particular environmental condition “sets off a genetically determined trigger” or whether it causes learned behavior to emerge is largely academic if we are simply trying to determine what environments are most likely to cause good outcomes to emerge. At the same time, a recognition of such more complex interactions is helpful in avoiding too quick a judgement regarding the potential for policy to make a difference.

4.3.2.2 *What does the Evidence NOT Imply?*

While the cautions of potential bias of heritability estimates of the kind we have reported are, we believe, largely unwarranted, some cautions regarding interpretation of the findings are extremely important. First, while the evidence does suggest that roughly 50 percent of cognitive ability and scholastic achievement is genetically influenced, it does not imply that intelligence is predetermined or that genetic disadvantages cannot be overcome by environmental factors, nor does it imply anything regarding the genetic basis for inter-race differences (as demonstrated in Section 4.1). Furthermore, the finding does not imply that policy cannot be effective in improving outcomes for disadvantaged children, whether the

³⁷ This criticism has also been successfully addressed more directly in studies that compare the impact of parentally perceived identical twin status and actual identical twin status.

disadvantage is genetic or environmental. More importantly, neither the finding of roughly 50 percent heritability nor the finding of a small impact of shared home environments in the samples that are studied implies that these are necessarily applicable to all homes and all contexts. Rather, given the data that are analyzed in these studies, these represent findings broadly applicable to the “normal range of environments and of intelligence” (Plomin and Petrill, 1997). They may not apply to extreme cases such as extremely poor households, abusive homes, and children of extremely low intelligence on the one hand, or extremely rich home environments or extremely intelligent children on the other. Current samples in behavioral genetic studies simply do not include sufficient numbers in those extremes to provide any conclusions.³⁸

More generally, we do not think the evidence we have reported can support statements such as “parents do not matter,” statements which are often attached to reports of this evidence.³⁹ First, as suggested already, the evidence speaks only to non-extreme cases and does not imply that shared home environments cannot make an extraordinary difference in certain cases. Second, even for the sample of households in the large middle of the distribution, the evidence is equally consistent with shared home environments being quite important. In particular, the evidence only states that differences in shared home environments for the vast majority of households cannot explain rather large differences in outcomes. One possible interpretation of this is that most parents by and large create home environments that are broadly speaking “healthy,” even if they do so in very different ways, using different parenting styles under different family structures with different incomes, etc. Under this interpretation, what parents do may be of tremendous importance for cognitive development in an absolute sense even if the differences in the home environments they create do not consistently contribute to differences in cognitive ability. Even under this interpretation, of course, while parents themselves remain important much of the attention given to specific parenting styles and choices in modern parenting handbooks remains unwarranted. Furthermore, when most people express a belief that “parents matter,” they are probably not merely referring to their impact on outcomes but also on the general ability of parents to provide pleasant environments for their children even if certain features of those environments do not raise cognitive ability. Finally, it may be that the “home environment” experienced by two children in the same household is importantly dissimilar, at least as observed by children, a hypothesis that seems to be confirmed by some of the evidence of child perceptions of home environments. Given the evidence on birth order presented in the next section, Harris (1995) has argued that the last caveat may be of minor significance compared to the others. However, as we argue in Section 8, we do not think this evidence is sufficient to rule out potentially important child specific effects arising from different interpretations of similar environments. Under all these interpretations, two lessons that can be drawn by social scientists are that it may be important (1) to view “environment” more broadly than some have done traditionally and to include community and school characteristics in the analysis, and (2) to allow for the possibility that what appear to be identical shared home environments are in fact not all that identical.

Given these caveats, we now turn to a review of the evidence from more traditional social science literatures which explore the influence of shared and non-shared home environment impacts through samples other than twins, adopted children and other siblings. We should point out that this literature is less concerned with cognitive ability per se and more with outcomes like student achievement and attainment, teenage fertility, juvenile crime, etc. Although behavioral geneticists have begun to look at such issues and are finding similar trends, their main emphasis thus far has been on cognitive ability. Still, much of the statistical evidence from the social sciences appears at first glance to be in stark contrast to the evidence

³⁸ In fact, some of the larger samples of twins from Europe are indeed quite homogeneous.

³⁹ Much of the popular press, for instance, greeted Harris’ (1998) presentation of the behavioral genetic evidence in her book with precisely this caricature, which she did not altogether discourage.

arising from behavioral genetic studies – especially the conclusion that shared home environments have little impact on cognitive outcomes. However, because of substantial methodological difficulties, both those involving the neglect of potential genetic factors as well as others related to topics raised in Section 3, much of this literature is less at odds with results reported thus far than may be apparent at first. Finally, as we point out again in Section 8, it is important to realize that, while behavioral geneticists attempt to find broad evidence on the importance of the total shared home environment, many of the studies we are about to analyze search for evidence on particular aspects of home environments. Even without the behavioral genetic evidence indicating a small role for home environments in explaining differences across individuals, we should therefore expect that no one specific factor of home environments will yield huge effects.

5 Causal Links of Child Outcomes to Within-Household Choices

This section explores how child outcomes are linked to several features of non-shared and shared family environments. First, we examine how non-shared home environments such as birth-order may affect outcomes. Despite conventional wisdom, our review of the evidence suggests little role for birth-order in the determination of outcomes, although this may differ according to specific cultural contexts. The rest of the section is devoted to an examination of the evidence on shared home environments. Genetic evidence from Section 4 already indicated a rather small role for shared home environments. Beyond this evidence, there is an enormous collection of statistical research in other social science disciplines on specific aspects of home environments. These include family structure (sibship size and spacing, divorce, single parenthood, and parental age) and parental investments of time and money in raising their children. Often the results of these studies accord with our expectations (i.e., that divorce is negatively associated with child outcomes). However, a common theme is that methodological challenges prevent us from ascribing a strong causal interpretation to the correlational results.

One group of studies focuses on family structure. For example, larger numbers of siblings tend to be associated with less cognitive development and lower educational attainment. But there is increasing statistical evidence that this relationship is spurious, and actually reflects the presence of other unobserved environmental or genetic factors in the home. Similarly, we find only limited evidence that the configuration of siblings (the number of girls relative to boys) affects outcomes. Another vast literature explores the impact of parental divorce on child outcomes. While there is a consistent negative association, the roots of this are not entirely clear. Three theories might be forwarded. First, the effects of divorce might actually reflect genetic influences, although evidence is inconclusive. Second, home environments might be quite stressful prior to the dissolution of the marriage, which could be the true cause of negative outcomes, rather than the trauma of divorce; some evidence points to this interpretation, but it is difficult to separate from hereditary explanations. Third, divorce may negatively affect outcomes by reducing the family income. Evidence suggests that direct effects of income are not large, although it does appear that home relocations following divorce could account for perhaps a quarter of the negative effects of divorce. Single parenthood is also negatively correlated with outcomes, particularly among lower-income families. But the causal meaning of these results is unclear. Several explanations exist, ranging from the lower-quality home environments of single-parent families, to hereditary factors. A quite robust finding is that children have lower outcomes when their mothers choose to live with, but not marry, another partner. Again, the causal channels underlying this observed correlation may include factors such as heredity or neighborhood effects; if such variables are omitted from the analysis, then their effects are confounded with family structure. Finally, neither parental age nor non-traditional family structures such as gay parents have been found to have strong impacts on outcomes.

Another group of studies examines how the quantity and quality of parental time investments affect childhood outcomes. The most recent evidence indicates that maternal work and child care during infancy is associated with negative behavioral and cognitive outcomes in early childhood but that these negative impacts are likely to be temporary and dissipate by middle childhood. For at-risk children, childcare is often found to improve outcomes, although these results also tend to “fade-out” as children get older. Furthermore, recent and more sophisticated research casts some doubt on whether simple indicators of child-care quality are truly associated with child outcomes. A second class of studies examines the effects of parental time investments, or lack thereof, on school-age children. Evidence from studies involving adolescent after-school care indicates that arrangements chosen by parents seem to play no large part in explaining differences in outcomes. Exceptions to this finding, however, are notable. While there is little evidence that time spent by parents monitoring adolescents within the home is important, lack of supervision after school outside the home may matter,

with children in unsupervised non-home environments associating with more deviant peers and demonstrating worse outcomes in school. Furthermore, while structured after-school programs seem to have little (or possibly negative) impact on middle class children, there is evidence suggesting they may improve outcomes for at-risk children, especially for boys. A common limitation of many of these studies is their inattention to problems of selection bias.

Apart from the literature on child care and maternal employment, a number of studies purport to show that parenting “quality” and parenting styles are related to child outcomes. While much evidence is suggestive of these influences, it is hampered by an over-reliance on Western conceptions of “good” parenting, data drawn from Western families, and the methodological challenges of selection bias and omitted variables bias.

Finally, parental income is often correlated with child outcomes. However, the best empirical research increasingly suggests that income is generally not a strong causal channel through which child outcomes improve significantly and that policies aimed at raising family incomes are likely to have only small effects. The observed income effect probably operates through other causal channels. An important caveat is that almost all research is focused on the U.S., where lower-income households already have higher incomes than poor households of many countries. Thus, income may still play a role in conditions of extreme poverty. Furthermore, income may have more appreciable effects in some special circumstances, most notably in the very early ages of at risk-children.

While we could fairly easily attribute a causal interpretation to correlations found in the twins and adoptive child studies reviewed in Section 4, this becomes more difficult as we venture into statistical studies of broader populations. In the previous section we followed Harris (1995) by arguing that much of the evidence derived from behavioral genetic studies suggests that, to the extent that home environments provide an important causal channel through which child outcomes are affected, such environments must typically be child-specific. Anything that has a “public good” nature within the home and is experienced similarly by siblings would be excluded by these studies as a significant causal channel (in average homes) for explaining differences in outcomes.

In this section we will approach the question surrounding the potential impact of home environments on child outcomes from a more traditional social science angle and will search for evidence of both shared and non-shared environmental factors. Unfortunately, many of the studies we have encountered run into the methodological challenges outlined in Section 3. Furthermore, most studies (other than those of behavioral geneticists) exclude considerations of genetic influences from the start. This makes it difficult to reconcile results on the impact of home environments from studies reviewed in the previous section with many of those in the mainstream psychology, sociology and economics literatures. We begin our review of such studies with those testing for non-shared environmental factors and then proceed to research dealing with the more commonly studied shared within-family influences.

5.1 Non-Shared Home Environments: Birth Order

Children living in the same home may experience different environments for a variety of reasons, and such child-specific environmental factors are not ruled out as possible causal channels for impacting personality and cognitive development by the behavioral genetic studies. These types of child-specific effects have been classified by McCartney et al. (1991), and Harris (1995) reviews the evidence in favor of each. She finds little consistent support. A detailed discussion of each effect is beyond the scope of this review, and we therefore limit ourselves to discussing the evidence of the most studied child-specific within-home environmental factor: birth order.

Birth order offers a particularly good avenue for studying child-specific environmental differences that are naturally distinct from genetic factors. First-borns clearly experience different home environments, at least early in life, but their genes do not consistently differ from their biological siblings'. Thus, any consistent difference in outcomes due to birth order must be environmental. Harris' (1995) reading of this literature, however, suggests there is little evidence in favor of the hypothesis that birth order affects personality and that, if it impacts cognitive development, it does so only slightly. While this is in stark contrast with much of the folklore surrounding birth order, most of the earlier studies claiming to find strong causal effects of birth order turned out to be fundamentally flawed. Since these studies were largely reports of rough statistical correlations, they lacked proper controls for demographic and socioeconomic variables (Ernst 1983). When proper controls were used, no effects of birth order on personality were found. Furthermore, more recent work indicates birth order plays little role in the determination of the level or growth of wages (Kessler 1991), and while some effect on schooling appears in one study, its effect on earnings in that study is not robust to model specification (Behrman and Taubman 1986).⁴⁰

In the past few years, however, this conclusion has again been challenged by provocative work by Sulloway (1995; 1997) who claims, among other findings, to uncover powerful historical evidence in favor of birth order determinants of social attitudes. But again, when tested on a large contemporary data set, Freese et al. (1999) find a complete vacuum of evidence supporting these claims and conclude that "birth-order theories may be better conceptualized in terms of modest effects in limited domains for specific societies." The latest evidence on the impact of birth order on cognitive ability is similarly weak (Retherford 1991; Kuo and Hauser 1997), although this evidence cannot rule out larger effects in cultures that may place more weight on birth order.

If our common presumption that home environments fundamentally shape child characteristics is correct, we are now left with a puzzle. The robust findings from a persuasive set of behavioral genetic studies suggest that, while there is a large (50%) unexplained role for non-genetic contributions to cognitive development and scholastic achievement, those features of home environments that are shared by siblings explain at best a minor portion of the non-genetic variance. Furthermore, psychologists and sociologists have found relatively little consistent evidence that the bulk of the remaining variance can be causally attributed to non-shared or child-specific home environments. This provides persuasive evidence in favor of the importance of broader environmental influences as well as impacts of micro-environments sought out by individuals as they grow up (Harris 1995, 1998). It also suggests an important role for future research to investigate to what extent shared home environments are perceived similarly by siblings within a home; i.e. to what extent "shared" home environments are actually shared. The presumption in both the behavioral genetics and the broader social science literature is that such perceptions largely coincide.

Before conceding that home environments are not as important as we might have thought, we now turn to research offered by scholars other than behavioral geneticists to see whether other types of evidence in favor of a large role for particular aspects of shared home environments can be found. As mentioned above, our search for such causal links is hampered by a general unwillingness on the part of most social scientists to consider such genetic influences carefully. In many cases, this leads to possible reinterpretations of their findings with quite different implications. In an attempt to provide us with the best possible chance of reconciling such studies with findings from behavioral geneticists, we will search particularly for evidence that suggests an important home environment influence under particular circumstances (such as extreme poverty, single motherhood, etc.). Given that much of the behavioral genetic evidence is about averages, the findings suggesting a causal link between

⁴⁰ The same study also provides a variety of theoretical reasons as to why birth order may matter (Behrman and Taubman 1986).

shared home environment and outcomes under special circumstances would not immediately contradict results from twins and adopted child studies. We begin this investigation in Section 5.2 with a review of the literature on the impact of family structure and continue in Section 5.3 with evidence on parental home investments.

5.2 *Family Structure*

Parental choices regarding family structure have long been seen as important environmental influences on child development. Such choices include the number and spacing of children, as well as choices regarding divorce and remarriage and the age at which to have the first child. Furthermore, they include the extent to which “family” refers to more than merely parents and siblings to include grandparents, aunts and uncles, and other extended family members. In the context of the United States, much of the research has dealt primarily with the narrower conception of family structure—parents and siblings. Correlations in the data are consistent with conventional wisdom—a negative relationship between family size (or sibship)⁴¹ and intellectual development, between divorce/single parenthood and child outcomes, and between young parental age and achievement. In each of these cases, however, there remains doubt as to whether the relationships in the data are causal or reflect other underlying variables. Nevertheless, to the extent that such negative relationships withstand the inclusion of various control variables, they seem to be strongest for at-risk segments of the population.

5.2.1 *Sibship Size, Spacing and Configuration*

The general result that sibship size is negatively correlated with cognitive development and educational attainment is well established⁴² and has again been confirmed in recent studies. Kuo and Hauser (1997) demonstrate that such correlation can persist even when family income and parental education/occupation are explicitly controlled for.⁴³ Other studies find that inclusion of family background and other variables tends to reduce (though not eliminate) the sibship effect.⁴⁴ In New Zealand, evidence from the Christchurch Health and Development Study confirm such international findings in the case of reading performance when various controls are included (Baker and Maloney 1999). Furthermore, some evidence has suggested that sibship size may have differential effects on boys and girls.⁴⁵ Because of data limitations, however, little has been done in this literature to investigate whether this sibship size effect is not masking a different unobserved factor such as the intellectual environment within the home or genetic inheritance.⁴⁶ If such unobserved variables are

⁴¹ Sibship differs from family size in that it does not count parents or other individuals that are in fact not siblings (Kuo and Hauser 1997).

⁴² See, for example, Hanushek (1992).

⁴³ This contradicts an earlier finding by Downey (1995) who finds that much of the sibship size effect is eliminated by an inclusion of family resources. His study differs from others, however, in that he explicitly measures parental resources devoted to educationally productive in-home expenditures, and he finds that these decline on a per child basis as sibship grows. We return to a more thorough discussion of this study in our review of parental resource investments below.

⁴⁴ One study (Anh, Knodel et al. 1998) reports that an initially strong negative correlation between family size and educational attainment in Vietnam largely disappears when urban residence, region, parental education and household wealth are taken into account, and the remaining effect is due mainly to the largest families in the sample.

⁴⁵ In particular, one study shows that women from small families work less early in life and more later as compared to women from large families, but no such effect was found for boys (Kessler 1991).

⁴⁶ Guo and VanWey (1999), for instance, refer to an older study which included adopted children and which found the usual sibship size effect for natural children but not for adoptive children, evidence for an unobserved genetic effect. Similar criticisms apply to studies of only children (as, for example, Falbo and Poston (1993)). Menaghan (1997) provides evidence that, holding maternal cognitive skill

important, and if they are correlated with sibship size, standard regression analysis would attribute them to sibship size even if there is no actual causal influence of sibship size on child outcomes. On a more macro level, even if sibship effects do exist, other evidence indicates that they cannot be large. Cohort differences in sibship size and configuration due to post-war changes in fertility, for instance, cannot explain any aspect of declining test scores in the U.S. (Alwin 1991).

In one of the most recent studies of sibship size on intellectual development, Guo and VanWey (1999) employ previously unused statistical techniques to isolate unobserved family-specific effects from effects due to sibship size. In particular, using a large data set from the National Longitudinal Survey of Youth, they employ a fixed-effects approach that differences out permanent, family-specific effects unrelated to sibship size. In the absence of reliable data on family-specific genetic factors and within-home intellectual climates, such a fixed-effects approach is indeed the best alternative.⁴⁷ While Guo and VanWey (1999) are able to replicate the usual negative sibship size effect on cognitive development using conventional regression techniques, this effect disappears entirely in the fixed-effects model, and the results seem to hold regardless of time spacing of siblings (Guo and VanWey 1999).⁴⁸ Furthermore, Phillips (1999) confirms this result using a different statistical technique proposed by Mayer (1997). Although neither of these techniques have been applied to data sets other than the NLSY, the results offer at least tentative evidence that the significant negative correlation between sibship size and cognitive development is likely not to be causal or, if it is, it is much smaller than previously thought.⁴⁹ We caution, however, that these recent results are from U.S. data, and that results from developing countries and from other contexts may be different. Particularly, resource dilution may be a bigger factor in developing countries, especially in light of the decreased probability of attending school that is associated with family size in such settings (Knodel and Wongsith 1991). Furthermore, these results do not speak to the issue of whether sibship size impacts the ability of children to attend college due to declines in per capita family resources.⁵⁰

A second argument often put forth is that the configuration of sibship—how many girls relative to boys there are in the family—may have important environmental impacts. Particularly, various theories suggest that girls who grow up with primarily male siblings will attain higher levels of education than girls who grow up with only sisters. While Butcher and Case (1994) provide some evidence of such effects, the more recent and more comprehensive evidence that utilizes significantly more data cast doubt on such findings (Hauser 1997). At best there is a small and declining impact of sibship configuration on the educational attainment of women. Our overall conclusion is that, while there is considerable suggestive evidence in earlier studies, we find little evidence for a major causal link between sibship

fixed, an increase in sibship within single parent homes is associated with declining cognitive stimulation.

⁴⁷ It does, however, require substantially more data than is contained in the typical cross-section, because it requires at least two separate and comparable measurements of intellectual ability at time intervals that permit the observation of sibship size changes.

⁴⁸ Powell and Steelman (1995) find evidence that spacing of children affect educational and economic outcomes, but their study is subject to similar criticisms as the earlier birth order literature.

⁴⁹ Incidentally, as Downey et al. (1999) point out in a comment to Guo and VanWey (1999), not only do fixed effects eliminate the impact of sibship size, but they also reduce in significance the impact of parental education, family income and marital status, all consistent with an interpretation that shared home environments have little impact once genetic factors are taken into account.

⁵⁰ Evidence comparing U.S. veterans who had free access to higher education through the GI Bill to their children who had no such free access indicates an inverse relationship between sibship size and schooling (as well as earnings) for the children but not the veteran parents (Behrman, Pollack et al. 1989). At the same time, recent evidence on college attendance suggests that long term factors associated with parental background rather than such credit constraints underlie differences in college attendance among ethnic groups in the U.S. today (Cameron and Heckman 1999).

size, spacing or configuration and cognitive development or educational outcomes in the most recent literature involving populations outside developing countries.⁵¹

5.2.2 *Divorce*

Much speculation concerning the potential effects of rising divorce rates and increasing incidences of single parenthood has driven discussions in a variety of social policy arenas, and corresponding levels of research attention have been lavished on these topics. In regard to the impact of divorce, such research involves a comparison of groups of children of divorced parents with groups who have not experienced divorce. With respect to single parenthood, additional attention has been given to out-of-wedlock births, especially those to teenage mothers. We touch on both these lines of research below. Correlations in the data are consistent and irrefutable even if they are not always large in magnitude (Amato and Keith 1991; McLanahan 1994). Children of divorce and children of single parents display greater behavioral and emotional problems, have higher incidences of marital divorce later in life, are more likely to become teen-parents, and have worse educational outcomes.⁵² These adverse impacts of divorce appear in the data whether family disruptions occur at birth, in early childhood or in adolescence and do not seem to be dramatically ameliorated by remarriage (McLanahan 1994), although there appear to be some differences for certain groups.⁵³ Furthermore, such correlations appear across international data sets, including those in New Zealand (Wylie, Thompson et al. 1999).⁵⁴

A variety of theoretical explanations for the correlation of divorce with adverse outcomes have been offered (Amato 1993).⁵⁵ The case for an empirically documented causal link between divorce and child outcomes is, however, somewhat more tenuous. Few studies have persuasively argued that children of divorce are at risk because of the traumatic nature of the divorce event. Much of the literature has therefore focused on some other factor connected to divorce. Particularly, we offer three possible explanations for the high correlation between

⁵¹ We are aware of a relatively large literature investigating such issues for developing countries. Since this literature is likely to be relatively uninformative for developed countries such as New Zealand, we exclude these from our review.

⁵² McLanahan and Sandefur (1994), for instance, report that 31 percent of children from families of divorce failed to graduate from high school while the same is true of only 13% of children from intact families. For data collected during childhood, however, there is some debate as to whether correlations between divorce and outcomes persist when observations outside the home (i.e. by teachers or other school officials) rather than observations within the home are used (Harris 1995; Harris 1998). For correlational evidence of the effect of family disruption on behavioral problems in New Zealand, see Pavuluri (1995). Furthermore, there is evidence in the demography literature on strong correlation of attitudes between parents and children toward divorce, and the correlations we cite extend well beyond the U.S. context (Cherlin, Kiernan et al. 1995; Axinn and Thornton 1996; Kiernan and Hobcraft 1997).

⁵³ McLanahan and Sandefur (1994), for instance, report increased risks for boys born to unwed mothers relative to those who live with divorced or separated mothers, and black children in the U.S. seem to perform better in school if their parents were married early in life. Also, they report some evidence indicating that a stepfather reduced the risks for young black men and women.

⁵⁴ While the broad correlations are also present in the Christchurch Health and Development Study, Baker and Maloney (1999) find little statistically significant impacts of single parenthood on reading ability in that data once other factors are taken into account. Similarly, Wylie et. al. (1999) report that, in the Competent Children longitudinal study, the correlation between family structure and child outcomes disappears when family income is taken into account.

⁵⁵ Some additional literature investigates the hypothesis that effects of divorce differ by gender. Phelps (1998), for instance, uses psychology data to assess the differential impact of divorce on the achievement motivation of girls relative to boys. He finds that divorce is actually correlated with higher achievement aspirations for girls while it has no impact on boys. Girls that have experienced divorce, however, tend to have lower earnings (while boys do not), although they tend to attain the same level of family income as their peers from intact families through marriage.

divorce and child outcomes that suggest this correlation is not due to the trauma of divorce itself: heredity, pre-divorce environmental factors, and post-divorce environmental factors.

First, we note that heredity may create an upward bias in our estimates of the effect of divorce due to an omitted variable effect. Unfortunately, the role of genetic factors that influence the likelihood of divorce and are also contributors to negative outcomes in children that inherit those traits is almost entirely absent from the analysis. Yet if there exist heritable traits that are contributors to divorce, and if such heritable traits are also contributors to child outcomes, then the omission of genetic variables in standard regressions will bias the estimates of the impact of divorce upward. This would lead researchers to believe that divorce itself is a causal link to child outcomes when in fact it is not. McGue and Lykken (1992), using methods similar to those reviewed in Section 4, estimate the heritability of divorce to be 50%. This may imply that such heritable traits as impulsiveness and tendency toward alcoholism (Loehlin 1992; McGue 1993) which tend to contribute to the dissolution of marriages may be a crucial causal link. Due to the general lack of further similar studies in these areas, however, we still know little about the extent to which inherited traits that help cause divorce may also be the cause of lower educational attainments or behavioral problems in children of divorce. McLanahan and Sandefur (1994) come closest to investigating the possibility that omitted variables, genetic or otherwise, cause their estimates of the impact of divorce to be upwardly biased, and they report mixed evidence.⁵⁶ It is therefore certainly not possible to use the available evidence to argue that the negative correlation between divorce and child outcomes is caused primarily or even in part by heredity. Neither, however, is it possible to reject such a hypothesis entirely.

A second concern is that divorce may be attributed with the impact of the marital stress and home environments that led to divorce. To the extent that pre-divorce stressors are important, divorce itself may not be the crucial causal factor behind the fact that children of divorce experience worse outcomes; these outcomes would have been worse even had the relatively dissatisfied marriages continued.⁵⁷ Feldman et al. (1990) present evidence that marital satisfaction is correlated with child outcomes, although much of the correlation disappears when the impact of family background and child rearing practices is accounted for. In a related study, Furstenberg and Teitler (1994) demonstrate that, while divorce is associated with some child outcomes, much of its effect vanishes when pre-divorce factors are taken into account. This is further confirmed by findings that child behavioral problems, particularly among boys (Cherlin 1991; Cherlin, Chase-Landsdale et al. 1998), tend to precede divorce, a finding that is consistent with both genetic and home environment explanations (Harris 1995) but inconsistent with a causal link between the divorce event and child outcomes. Similarly consistent with both the heredity and the pre-divorce environment explanation is the finding that children growing up with a widowed parent experience no detectable decline in outcomes (McLanahan 1994; McLanahan 1997). Finally, McLanahan and Sandefur (1994) report that, when pre-divorce income is taken into account, the negative correlation of divorce with such outcomes as high school graduation or teen pregnancy declines substantially. Pre-divorce factors, whether genetic or otherwise, therefore appear to be important for at least some outcomes.

The third avenue through which divorce may affect children involves the changes in family financial status following divorce. This may lead to fewer in-home resources, greater parental

⁵⁶ In particular, they use bivariate probit models to find that some of the estimates of the impact of divorce are moderated by the likely presence of omitted variables. In one of their data sets the impact of divorce on graduation rates, for instance, vanishes completely upon inclusion of additional variables, but in other data sets it does not. Furthermore, they find less evidence that omitted variables are problematic in their estimates of the impact of divorce on other outcome measures such as teen pregnancy and idleness.

⁵⁷ For example, studies consistently reveal a correlation between marital stress and levels of cognitive stimulation and maternal warmth as well as child outcomes (Menaghan, Kowaleski-Jones et al. 1997).

stressors, and re-locations of children to worse neighborhoods and worse schools.⁵⁸ McLanahan and Sandefur (1994), for instance, find that changes in parenting practices after divorce lead to higher rates of teen pregnancy (while contributing little to idleness),⁵⁹ and changes in household income can explain almost half of the divorce-related increase in high school drop out rates. More generally, however, while children in married mother-stepfather families are found to have environments no different from children growing up in biological two parent families (Menaghan, Kowaleski-Jones et al. 1997), mounting evidence suggests that children of divorce fare equally well (in terms of school or behavioral outcomes) in newly formed two parent households (with one step-parent) as they would had they remained in single parent settings (McLanahan 1994; McLanahan 1997).⁶⁰ Combined with evidence on the causal impact of income on child outcomes reviewed in Section 5.3, this provides strong evidence that, to the extent that post-divorce environments are responsible for the link between divorce and poor child outcomes, a change in within-home resources is not the crucial factor. At the same time, divorce does entail re-location for children of divorce regardless of whether their mother ultimately remarries or not. New residences tend to be in more disadvantaged neighborhoods with peers that display worse attitudes and schools of lower quality. Furthermore, even if (for step-parent families) there is no decline in neighborhood quality, the act of relocation nevertheless typically involves substantial peer disruptions. McLanahan and Sandefur (1994) attribute as much as 25 percent of the divorce-related high school drop-outs to such relocations, and they find that adjusting for differences in school quality and residential mobility eliminates a large part of differences in the risk of dropping out of high school between children in two-parent families and stepfamilies. Any post-divorce environmental impact of divorce is therefore likely to be due to relocation effects, not within home resource effects. We will have more to say on such effects in our discussion of household income as well as our later discussion of neighborhood effects below.

5.2.3 *Single Parenthood*

The question of whether divorce adversely impacts children is in some ways distinct from the question of whether single parenthood alone does the same. Again, correlations with adverse outcomes are well documented and seem to be additive (i.e., the more years spent in single parent homes, the worse are the outcomes).⁶¹ While such correlations are consistent across studies, they are sometimes significantly weaker in magnitude than other correlations. Garasky (1995), for instance, reports that a child is more likely to be observed graduating from high school if she grows up with a single mother who is herself a high school graduate than if she grows up in a two parent household in which the mother was not a high school graduate. Furthermore, some recent empirical findings suggest that the correlation between single parenthood and lowered educational attainment essentially disappears when economic

⁵⁸ Wood et al. (1993) find that, after controlling for demographic differences, children whose families have moved often are almost twice as likely to have behavioral problems relative to their peers who did not move. While we again caution that this ought not be interpreted as a causal link without further work, it is suggestive of one possible causal channel through which divorce might impact children.

⁵⁹ Data limitations did not permit investigation of whether changes in parenting styles contribute to other child outcomes.

⁶⁰ Although this finding appears across a large variety of studies (Duncan and Brooks-Gunn 1997), it is not entirely without controversy. Menaghan et al. (1997) suggest in a sample skewed toward low income and young mothers of adolescents that those growing up in mother-stepfather settings do not show cognitive deficits. In the same sample, however, no cognitive deficits are found for adolescents in single-parent households in the absence of unmarried partners. The importance of re-marriage in rising out of poverty is also discussed in the literature (Kniesner, McElroy et al. 1988).

⁶¹ See, for example, McLanahan (1989) for an overview of some such results, and Wojtkiewicz (1993) for a duration analysis. These studies also observe a generally negative correlation of living with step-parents and negative outcomes. Recent studies confirming such correlations include Thomas et al. (1996), Sandefur et al. (1992), Boggess (1998), and McLanahan and Sandefur (1994).

status is properly controlled for (Boggess 1998). In recent work on New Zealand, single parenthood is found to generally have little correlation with child competency at age 8 once the effect of income is accounted for, although some remaining correlations do appear for households with low incomes for long periods (Wylie, Thompson et al. 1999). This is mirrored in several studies that consistently find the impact of single motherhood on educational attainment varying with the mother's occupational complexity, with more complex (i.e. better) jobs being correlated with significantly reduced effects of single motherhood (Parcel and Menaghan 1994; Menaghan and Parcel 1995; Menaghan, Kowaleski-Jones et al. 1997). Mulkey et al. (1992) report that the significantly lower test scores of children from single-parent households is explained almost entirely by family background and variables such as school lateness, not doing homework, not having contact with parents and frequent dating. At the same time, however, McLanahan (1997) points out that a stronger role for single parenthood can be found in explaining behavioral and psychological difficulties than in accounting for educational attainment. In any case, it is quite clear from the data that adverse correlations between single parenthood and child outcomes are significantly stronger in low-income families, especially those in which mothers work in jobs involving little complexity.

Causal relationships are again more difficult to identify. One common explanation for the adverse correlation between single parenthood and child outcomes is that single parents are simply not able to provide the same kind of nurturing home environment as two parent families. This is certainly confirmed in the data. Menaghan et al. (1997) report that children living with single mothers experience lower cognitive stimulation and less active structuring. Echoing the findings on job complexity and educational attainment, however, this is less true for single mothers with jobs involving complexity than for those with no work, and the effect is further aggravated for mothers with jobs of low complexity. This is suggestive evidence that the negative impact of growing up in a single parent household is due to a lower quality of the home environment in such households. It is equally consistent, however, with genetic explanations. If heritable traits that influence job complexity and parenting are correlated with heritable traits influencing child outcomes, then the statistical analysis will yield upwardly biased coefficients on single parenthood and parenting techniques in single parent homes as long as unobserved genetic factors are not included in the analysis. We have little evidence to report in this regard. However, while the degree of parental encouragement is often cited as the causal channel through which single parenthood affects child outcomes, Astone and McLanahan (1991) find that the significant observed differences in parental encouragement between single and two parent households cannot account for significant differences in educational attainment of children that grow up in those households once other variables are controlled for.⁶² Further evidence against the hypothesis that lower quality of home environment is responsible for single parent effects is provided by the already cited studies on remarriage. Remarriage seems to have little positive impact on children of single mothers even though home environments in mother-stepfather families are not consistently different from home environments involving two biological parents.

There is, however, an important exception to this finding. Children that live in a home which does not have a biological father demonstrate substantially worse cognitive, behavioral and psychological outcomes when an unmarried partner is present in the home than when there is no partner or when there is a married partner other than the biological father (Murray 1984; Menaghan, Kowaleski-Jones et al. 1997). Children of divorce or out-of-wedlock children whose single mothers stay single or marry another man therefore do better than similar children whose mothers live with a partner in the absence of marriage. This result is robust to the inclusion of large numbers of control variables for maternal background (including

⁶² They find that they cannot explain more than 10% of increased drop out rates with such parental behavior variables, although McLanahan and Sandefur (1994) report some less sophisticated and tentative results indicating a somewhat higher potential for parental behavior as a causal factor.

cognitive ability and educational attainment). Furthermore, the data consistently suggest that single motherhood in the presence of unmarried partners is accompanied by less cognitive stimulation and less maternal warmth in the home (Menaghan, Kowaleski-Jones et al. 1997). Given the weak evidence of home environment effects in other settings, one is nevertheless cautious in interpreting the correlation between unmarried live-in partners and poor home environments as the causal channel through which poorer child outcomes arise. Rather, there are two significant selection issues that have simply not been adequately dealt with in the literature.

The first issue involves the unobserved personal characteristics of mothers who choose live-in partnerships over no partners or remarriage. It is unlikely that the control variables on maternal background adequately capture underlying personality traits that are in part heritable, traits that are also responsible for poor child performance and behavioral problems, nor that they capture all environmentally relevant variables. If we had data on such unobservables, we could isolate the impact of poor home environments in homes with unmarried partners from such genetic and environmental channels. Reading current results in light of Harris' (1995) theory of group socialization, for example, one could argue that such maternal behaviors are proxies for the ability of mothers to conform to social norms, an ability that is likely to be, at least in part, heritable and correlated with the likelihood that children will adapt to norms set for them in schools. This means that we should observe less of an impact of single parenthood and the presence of unmarried partners on child outcomes in cultures that do not stigmatize such family arrangements, a finding consistent with evidence from Western European countries such as Sweden. In the absence of more data, we are therefore left suspecting that estimates attributing causation of poor child outcomes to home environments with live-in partners are biased upward. No studies we have found can tell us by how much.

The second issue involves neighborhood impacts. Single mothers who choose to live with partners without marriage are likely to choose neighborhoods different from those chosen by single mothers that choose marriage. The studies we have cited thus far do not adequately address this. In particular, without the inclusion of neighborhood effects, estimates of the role of unmarried partners in the home will reflect such effects and thus be upwardly biased.

While both divorce and single-parenthood are therefore associated with poorer child outcomes, the causal channels for this correlation have not always been adequately identified. Given the lack of evidence on important shared home environment effects from behavioral genetic studies, one would suspect that at least some focus ought to be given to a more thorough investigation of whether this correlation does not mask some underlying genetic factors or important neighborhood effects. As in prior sections of this review, we were surprised to find so little consensus regarding many of the plausible causal channels and the weak evidence for some of them. When added to the increasing evidence of little impact of sibship size and structure, and mounting evidence on the neutral impact of growing up in non-traditional households, our tentative impression from the evidence is that family structure has less causal impact on child outcomes than commonly thought and represents only a small fraction of the impact of environmental forces on child outcomes. Further plausible channels include heredity, neighborhoods and school environments. However, the fraction of effects due to home environments are consistently higher for low-income families, especially those with poor jobs, and it remains to be seen whether such households represent one category to which insights regarding home environment from behavioral genetic studies do not apply.

5.2.4 Parental Age and Non-Traditional Households

Geronimus et al. (1994) provide an insightful review of the literature on the impact of maternal age and child outcomes, as well as new evidence on the topic. In particular, while they recognize a large body of literature suggesting adverse impacts of parental age on child development, they raise methodological issues similar to those we have raised in this review. Their conclusion is that the impact of maternal age on child outcomes has likely been substantially overestimated. While we do not undertake a full review of the methodological issues in this area (which are similar to those already mentioned for other studies), we note in particular their finding with first cousin data that “observed differences in developmental indicators between children with teen mothers and those with non-teen mothers are confounded by maternal family background” (p.601). This result is confirmed for a range of child development indicators. In other words, Geronimus et al. find that children of teenage mothers fare no worse than the children of those mothers’ cousins who chose to postpone fertility until later.

One final issue tangentially related to family structure involves recent findings on the impact of growing up in non-traditional families involving gay parents. Evidence from this literature finds virtually no effects different from growing up in traditional households (Flaks, Fichter et al. 1995).

5.3 *Parental Resource Investments in Children*

Virtually all economic models of household production (Becker and Tomes 1976; Becker, Murphy et al. 1990; Rosenzweig 1990; Becker 1991) assume a large role for parental resources in the creation of human capital in children. Empirical support for such models has been reported in various outlets. Hanushek (1992), for instance, interprets his finding of declining ability with increased family size as supportive of such models because larger family size is associated with declining home resources per child.⁶³ He admits, however, that in the absence of observations on time allocations within households, his results are subject to alternative interpretations. Similar interpretations of the previously reviewed sibship size literature abound across disciplines. These interpretations are bolstered by evidence that the availability of parental resources does in fact decline in family size (Downey 1995), and that this decline in parental resources can be directly linked to declines in child outcomes. The theory of “resource dilution” is therefore part of the assumed explanation for the consistent negative correlation between sibship size and child outcomes in virtually every study that claims to document such correlations, and moreover, it is almost always assumed that resources within home environments are the key. In this section, we will therefore search for more general findings regarding the impact of parental time and financial resources on child outcomes, whether in traditional two-parent families or not.

5.3.1 *Parental Time Investments: Quantity and Quality*

Household investment theories of child development often emphasize parental time investments in their children. Such investments can be delineated along both quantity and quality dimensions. While many studies have documented great differences in such investments, not all uniformly report finding strong causal links between them and observed child outcomes. Golombok et al. (1995), for example, found that parents of children conceived through the use of modern reproductive technologies score significantly higher than parents in control groups on a variety of dimensions including stress associated with parenting, marital satisfaction and quality of interactions with the child. However, no differences were found between the groups in a variety of emotional and behavioral measures of child outcomes. The study concludes that, in non-dysfunctional families, increased levels

⁶³ Hanushek similarly finds that virtually all of the positive effect of being born first is due to an increased probability that first borns are only borns.

of warmth and parental involvement do not result in greater well-being for the child. Similarly, Benasich and Brooks-Gunn (1996) find that measures of maternal knowledge of child development and child rearing have a strong correlation with the quality of the home environment and the number of child behavioral problems but only a small impact on cognitive development (IQ) at age 36 months. We begin by concentrating on different implications from a variety of such studies.

5.3.1.1 Early Childcare, Maternal Work and Child Care Quality

By far the largest set of studies investigating the impact of the quantity of time invested by parents (particularly mothers) on child outcomes involve those studies comparing outcomes for families with mothers that work in the labor force as opposed to mothers that stay at home to raise their children.⁶⁴ Although there are exceptions,⁶⁵ early (infant) childcare in these studies has been consistently related to generally poorer behavioral and cognitive outcomes for most children in early childhood, except in some cases for those coming from particularly disadvantaged backgrounds.⁶⁶ However, these effects of early childcare seem to be muted or absent when the same children are observed in middle childhood or adolescence. Burchinal et al. (1995), for instance, report no consistent relationship between non-maternal care in infancy and child cognitive and behavioral outcomes in children between the ages of 6 and 12. Harvey (1999), while confirming earlier results of negative impacts for younger children, finds only slight negative impacts of early maternal employment on 12 year-olds. In the Competent Children longitudinal study of New Zealand, effects of quality early child care remain detectable at age 8, although no adjustments for selection bias are made (Wylie, Thompson et al. 1999). Modest effects of maternal work on reading scores between age 8 and 13 are documented for the Christchurch Longitudinal data (Baker and Maloney 1999), while none are found in the correlational analysis of the Competent Children study (Wylie, Thompson et al. 1999).

Studies such as these, of course, run into problems similar to those previously identified in other sections. In particular, there may be consistent differences between mothers who choose to work and those who do not, and different forms of childcare arrangements may be associated with some underlying parental characteristics that we are not measuring. Burchinal et al. (1995), for instance, found selection of childcare arrangements to be strongly correlated with family and child characteristics already known to relate to developmental outcomes, and inclusion of such characteristics substantially muted at least some of the observed impact.⁶⁷

⁶⁴ Various arguments as to why early maternal work might matter are offered in the literature, with some of the more recent studies emphasizing the relationship between early work and breastfeeding (Roe 1999).

⁶⁵ Greenstein (1995) finds essentially no evidence of any adverse impact of early maternal employment among the “most advantaged” children in the NLSY.

⁶⁶ This is firmly established in several studies using data from the NLSY. Controlling for the usual socioeconomic family background, Baydar and Brooks-Gunn (1991), for instance, report substantial negative effects of maternal employment during infancy on the cognitive and behavioral outcomes of 3 and 4 year-olds, while such effects were more negligible if mothers postponed employment until the second and third years. Furthermore, for children, especially boys, in poverty, grandmother care was associated with significantly better outcomes than non-family child care, possibly due to the lack of high quality child care available to poor families. Similarly, Belsky and Eggebeen (1991) found children whose mothers returned to work in the first two years of life to be less compliant when they were 4 to 6 years old than children whose mothers did not work in the first two years of life. And Caughy et al. (1994) find that reading achievement in 5 to 6 year olds was negatively impacted by early child care for children from stimulating and responsive home environments and positively impacted from children with poor home environments. Finally, Blau (1992) reports negative impacts of maternal work in the first year of a child’s life on cognitive outcomes at age 3 and 4, but potentially offsetting positive effects if work continues through the next several years of the child’s life.

⁶⁷ These selection effects are confirmed in other studies (Dunn 1993; Peisner-Feinberg and Burchinal 1997).

None of these studies, however, attempt to correct for selection bias using the statistical techniques outlined in Section 3. Furthermore, many of the studies focus on type and quantity of child care without paying much attention to quality. When taking quality into account, Peisner-Feinberg and Burchinal (1997) and Dunn (1993), document a modest positive relation between childcare quality and cognitive/socioemotional outcomes with stronger (positive) effects on at-risk children but with negative effects of poor childcare no weaker for children from advantaged backgrounds. Kontos (1991), on the other hand, finds no such relationships. No attempt, however, was made in these studies to investigate whether these effects persist as children age, nor were any tests performed to detect the potential role of correlated unobservables.⁶⁸

The most recent evidence therefore indicates that maternal work and child care during infancy is associated with negative behavioral and cognitive outcomes in early childhood but that these negative impacts are likely to be temporary and dissipate by middle childhood. In addition, some studies suggest that early infant childcare has positive impacts for at-risk children, but the evidence is mixed.⁶⁹ Pre-school and center-based childcare for young children (past infancy), on the other hand, are generally associated with improvements in cognitive, behavioral and social adjustment dimensions during the time children are in such environments, with some evidence suggesting a stronger positive impact for at-risk children.⁷⁰ Furthermore, the effect of preschool childcare seems to be stronger for social adjustment variables than for cognitive outcomes (Kontos 1991). But the results on infant child care are mirrored in the sense that most such impacts seem to dissipate over time and are no longer detectable in middle childhood, at least in middle class children in the US (Clarke-Stewart 1991; Burchinal 1995). Similar dissipation has been documented for early childhood intervention programs (such as Head Start) targeted at raising cognitive skills among underprivileged at-risk children (Haskins 1989; Zigler and Muenchow 1992; Bryant 1994), although some studies indicate that somewhat muted positive effects in other dimensions such as school achievement and behavior persist into middle childhood (Barnett 1995). Burchinal et al. (1995) find the largest cognitive and behavioral outcome gains for African-American children attending center preschools and demonstrate gains that persist into middle childhood.⁷¹ However, other recent work using (household) fixed effects reports that dramatic scholastic gains for African American children do vanish quickly – much more so than for white at-risk children (Currie and Duncan 1995).⁷²

As with studies on maternal care, much of the work on child-care is plagued by the selection and endogeneity issues raised above as well as small sample sizes. Only recently has research begun to grapple with these issues. In particular, Blau (1997; forthcoming) has focused precisely on these issues in the most recent and most careful econometric work on child care done to date. First he observes that, while there is substantial evidence that the nature of interactions between child care providers and children properly constitute elements of child

⁶⁸ It is notable that earlier studies found significantly higher negative impacts of early maternal work, effects that are not found in more recent data using similar methodologies. It is likely that mothers going to work early in a child's life in older data were in fact consistently different than mothers who do so today when maternal work is more socially acceptable, and that those differences drove the earlier results.

⁶⁹ See a review of some of this evidence in Peisner-Feinberg and Burchinal (1997).

⁷⁰ Such positive impacts for at-risk children through age 8 seems to be confirmed in the New Zealand "Competent Children" data (Wylie, Thompson et al. 1999), although it is difficult to be certain unless more thorough statistical analysis is conducted. In the Christchurch Longitudinal data, Baker and Maloney (1999) find no evidence for a positive impact of preschool on reading scores at age 8 through 13.

⁷¹ For an accessible critical discussion of much of the research on Head Start, see Besharov (1992).

⁷² Head Start does, in this study, reduce the probability of grade repetition for white at-risk children. The study also notes other benefits from Head Start that are unrelated to cognitive or school outcomes – such as access to health care.

care quality, it is not clear that easily observed characteristics of child care centers provide valuable information regarding such quality. In particular, despite a plethora of largely correlational studies in the developmental psychology literature, studies that link child care quality with such easily observed inputs as staff-child ratios, group size, Blau (1997) finds that such linkages are not robust to changes in functional forms and to the inclusion of unobserved heterogeneity among child care providers. More importantly for purposes of this review, Blau (forthcoming) then proceeds to ask whether child care inputs are related to child development outcomes (rather than child care quality). He begins with a review of the developmental psychology literature and finds, for reasons similar to those enumerated throughout this review, that much of this literature is uninformative. It generally consists of correlational studies and ignores problems of small sample sizes, nonrandomly selected samples, a lack of inclusion of control variables such as family characteristics, and selection bias. He then proceeds to using a large sample of data from the U.S. National Longitudinal Survey of Youth, a data set that contains information on home inputs, household characteristics, child care characteristics and child outcomes for behavioral problems as well as mathematics, reading and vocabulary skills. Not only does the richness of this data set allow for the inclusion of many variables, it also makes possible a series of careful econometric tests (a review of which is beyond the scope of this review).⁷³ Blau's basic conclusion is that child care inputs such as staff-child ratios, group size and provider training generally have little impact on child outcomes.⁷⁴ This is not to say that differences in actual child care quality do not matter. Rather, the measures commonly used to proxy for child care quality do not seem adequate.⁷⁵

5.3.1.2 Parental Time and School-Age Children

A variety of other studies investigate the impact of after-school childcare arrangements for school-age children. These studies differentiate between various after-school experiences including self-care, distant adult supervision, in-home adult supervision, and structured after-school programs of various kinds. While some conflicting findings have been published, a few consistent findings emerge.

One fairly consistent finding is that surprisingly little impact from different levels of within-home parent supervision can be detected in child cognitive, behavioral and emotional outcomes once other background factors are accounted for (Rodman, Pratto et al. 1985; Vandell and Corasaniti 1988; Galambos and Maggs 1991). Vandell and Ramanan (1991), for instance, find self-care and care by single mothers of fifth graders in the NLSY to be correlated with worse vocabulary test scores and greater behavioral problems in school, but such effects disappear when family income and emotional support are controlled for. Similarly, Galambos and Maggs (1991) report no systematic difference between self-image and peer experience for sixth graders in self care at home and their peers in after-school adult supervision. However, children (especially girls) under more distant adult supervision (i.e. self-care outside the home) reported problem behaviors and greater deviant peer interactions, while firm parental control mediated such effects. These results mirror earlier studies that indicate physical distance from parents whose children are in self-care outside the home correlates with negative outcomes while no such results could be found regarding the level of

⁷³ These include the use of fixed effects controlling for unobserved maternal characteristics as well as a variety of interaction effects for race, poverty and the mode of child care. Furthermore, "value added" specifications yield similar results.

⁷⁴ One small exception is the finding that smaller group sizes in child care settings for children in the second three years of life may have small but statistically significant positive effects. Blau cautions, however, that the data did not permit sufficient robustness tests to control for unobserved heterogeneity.

⁷⁵ It should, however, provide some caution to an overly optimistic interpretation of the correlational evidence in the New Zealand "Competent Children" data between some aspects of child care and competency tests at age 8 (Wylie, Thompson et al. 1999).

supervision within the home (Steinberg 1986). These studies thus indicate that adolescent children are more susceptible to negative peer influences if they are unsupervised in, for example, a friend's house or even more if they are "hanging out" in the community, but not if they are relatively unsupervised within the home. Again, we suspect that parents who leave their children unsupervised outside the home are consistently different in other ways that are not captured in such studies, indicating that estimates may be upward biased.

Studies investigating the impact of structured after-school programs differ most dramatically depending on whether low income or middle income households are studied. In a sample of middle class third graders in both intact and divorced families, Vandell and Corassaniti (1988) find no detectable difference (in test scores or behavioral measures) between latchkey children and those returning to mother supervision after school. At the same time, however, they document significant negative impacts for children attending after-school day care. This differs sharply from results found for adolescents from low-income households. After accounting for family background, several studies (Posner 1994; Pierce, Hamm et al. 1999; Posner and Vandell 1999), for instance, find significant positive effects of structured after-school programs on academic achievement and social adjustment for at-risk children. Children in such programs participated more in cognitively stimulating activities and less in unstructured outside interactions or passive watching of TV, and such activities were correlated with better outcomes. These results are suggestive of a potential for structured after-school programs to improve outcomes for at-risk adolescents, especially boys (Pierce, Hamm et al. 1999) who tend to have worse behavior problems to begin with. Furthermore, these insights are confirmed in recent evidence from a controlled experiment in Milwaukee where parents of at-risk children received subsidies of various kinds. The parents in this group tended to chose higher levels of structured child care, after school care, involvement in clubs, etc. for their boys than they did for their girls, and achievement and behavioral gains were visible in boys but not in girls (Huston, Duncan et al. 1999). Unfortunately, however, we know little about whether the effects that were observed are transitory or last into adulthood.⁷⁶ If patterns for adolescent after-school programs are similar to patterns for early childhood programs, much of the good news regarding the potential for after-school programs to improve outcomes for low-income children may be overstated.

Evidence from studies involving adolescent after-school care therefore provide several useful insights. For most children, after-school child care arrangements chosen by parents seem to play no large part in explaining differences in outcomes. Exceptions to this finding, however, are notable. While there is little evidence that time spent by parents monitoring adolescents within the home is important, lack of supervision after school outside the home may matter, with children in unsupervised non-home environments associating with more deviant peers and demonstrating worse outcomes in school. Furthermore, while structured after-school programs seem to have little (or possibly negative) impact on middle class children, there is evidence suggesting they may improve outcomes for at-risk children. Similar suggestive evidence exists for preschool interventions targeted at low-income families. However, it is not always so straightforward to interpret these results as their authors do. As in the previous section on preschool care, no attempts are generally made to account for selection bias despite evidence that families who place their children in after school care, for example, differ systematically in terms of income, race and emotional support (Vandell 1991; Posner 1994). Furthermore, parents may place their children in self-care within the home only when they

⁷⁶ More on this experimental program (known as the New Hope Project) is reported in Section 5.3.2. A follow-up study is soon to get under way to ascertain whether gains made by boys in the study are fading out. It should also be noted that parents frequently stated that their differential use of supervised care for boys was due to their perception that boys were at greater risk. This is indicative of a relatively unexplored possible selection bias in many studies: at-risk children who end up in supervised settings may be those who are at greater risk, which may cause our estimates of the effectiveness of these programs to be biased downward.

judge them to be sufficiently independent while arranging for more structured care otherwise. Thus, children in self-care may be quite different than children in structured environment, and the selection process may cause us to underestimate the impact of such environments. Finally, too little emphasis has been placed on long run outcomes. For studies of early childcare, much of the evidence suggests that influences of child-care and maternal employment dissipate or completely disappear over time, and similar effects may be true for adolescent child care arrangements.

The available evidence on the importance of quantity of parental time spent with children therefore suggests that early studies had overstated its importance when long-run rather than short run outcomes are considered. Similarly, however, the impact of child-care arrangements outside the home, whether positive or negative, seems to fade over time for most children. While some evidence suggests that results may differ for at-risk children, more work is required to determine under what conditions these results hold in the long run. In general, the evidence of smaller long run impacts is consistent with the evidence of a home-environment fade-out from the behavioral genetic studies, at least as it concerns cognitive ability. The evidence in favor of more lasting long run impacts on other outcomes (such as school completion rates and behavior problems) is more favorable, especially for at-risk children.

5.3.1.3 Parenting Techniques and Parenting “Quality”

Much of the previous discussion has focused on the quantity of time invested by parents in their children. While we have found little evidence suggesting that such quantities are consistently related to better outcomes, we have only briefly touched on the question of the impact of the quality of time invested by parents. Much has been written, for instance, on the impact of different parenting techniques ranging from permissive to authoritarian, and a thorough review of all that has been written on the subject is beyond the scope of this review. However, a few observations are in order, especially in light of other findings we report. In particular, we found little evidence to support the hypothesis that higher levels of parental supervision of adolescents within the home causes detectable differences in child outcomes, and our previous review of evidence from behavioral geneticists finds little support for an important long run role of differing shared home environments (particularly for cognitive ability). This stands in stark contrast to the literature on parenting styles that suggests a large role for parents of adolescents. It is further contradicted by several recent studies.⁷⁷

Baumrind (1971) classified parenting techniques into three categories: authoritarian, authoritative and permissive. Broadly interpreted, authoritarian parents are those who take little account of their children’s opinions while authoritative parents are firm but warm and do take their children’s opinions into account. Permissive parents are those that allow children a wide range of freedom in choosing what they want to do. A consistent finding in the literature is that children growing up in authoritative households are most well adjusted and outperform others in academic achievement.⁷⁸ At the same time, however, this conclusion appears to be

⁷⁷ In recent work on the Christchurch Health and Development Study, evidence is found that emotional responsiveness of mothers when children are 3 years old is correlated with reading competence at ages 8 through 13, even when many other controls are included (Baker and Maloney 1999).

⁷⁸ See, for example, Lamborn et al. (1991). Dodge et al. (1994) find that half of the negative impact of low socioeconomic status early in childhood on child behavior problems in school can be explained by home factors such as harsh discipline, lack of maternal warmth and maternal lack of social support, peer group instability, family stressors, and lack of cognitive stimulation. There are similar findings for high school students. Brown et al. (1993), for example, find that parenting practices toward high school students (such as monitoring, encouragement of achievement and joint decision making) are correlated with positive adolescent behaviors such as academic achievement, lack of drug use and self reliance and that these were correlated with membership in common adolescent groups.

quite culturally biased.⁷⁹ In particular, it holds most consistently for Western middle class households, but does not hold within other cultures and settings. Asian Americans, for instance, display the most authoritarian parenting styles across cultural groups, and these parenting styles within the Asian-American community are correlated with positive cognitive and school outcomes (Steinberg 1992). Similar results regarding positive impacts of authoritative parenting styles have been found for disadvantaged minority households (Baldwin 1990). It is difficult to reconcile such results with the broad generalizations often found in the literature that authoritative parenting is the best form of parenting in all circumstances.⁸⁰ Perhaps, parenting styles are complementary to other cultural factors, which implies that some styles are better for some cultures than others. Or perhaps certain aspects of parenting are damaging unless complemented by others.⁸¹

One recent explanation, however, is offered by Harris (1995). Suppose that different cultural groups have different norms and customs in regard to child rearing. Then a parent's adoption of the parenting style that is considered most acceptable within her cultural group may reflect a broader competence on the part of the parent to conform to cultural norms. This competence is then passed on to children (whether genetically or not) and is further reflected in school outcomes. Parenting styles in psychology studies may therefore proxy for an ability by parents to rear their children in the style approved by their culture. Since different sub-cultures (such as Asian-Americans or other minorities) have different norms regarding child rearing, we would then expect parenting styles to show different correlations depending on the cultural setting without attributing any causal impact of parenting styles per se on child outcomes. We can offer little evidence in this regard. In a somewhat related study, Whitbeck et al. (1997) investigate the correlation of parental working conditions and economic hardship in dual-earner families and find, at least for fathers, that more adverse conditions are negatively correlated with accepted norms of "good" parenting, and further that such good parenting (inductive parenting techniques and avoidance of harsh parenting behaviors) was correlated with adolescent self efficacy. One reading of this study attributes causal influence to parenting styles (impacted by poor working conditions). Another reading, however, suggests that those displaying "poor parenting" lacked the ability to conform to social norms in general, which caused them to end up in poor working conditions. If such abilities are in part heritable, this would further explain the correlation with poor child outcomes.

Again, we do not mean to suggest that we have found firm evidence that parenting styles do not represent causal links to child outcomes. In the extreme, they most certainly do even if they do not within the range of family experiences generally studied in the types of papers we have reviewed. There is a long history of documented long-run effects from severe childhood abuse. However, both evidence from behavioral genetics as well as evidence from cross-cultural studies of parenting styles suggest that many studies are currently too quick to jump to conclusions regarding the potential for a causal link between variations in parenting style and child success among non-dysfunctional families.

5.3.2 *Parental Income and Money Investment*

For much of this century, the strong correlation between family income and child outcomes has been assumed by social scientists to be at least in great part causal. Such correlations in

⁷⁹ See, for example, Ambert (1994) for an illustration of how Western biases of parenting dominate the scholarly research.

⁸⁰ Furthermore, those studies that rely on recollections of parenting styles in adulthood are suspect as twins and adopted child studies have shown such perceptions to be heritable (i.e., identical twins' recollections are substantially more correlated than fraternal twin's perceptions).

⁸¹ Conservative Protestants in the U.S., for instance, combine strict discipline with unusually warm parent-child interactions (Wilcox 1998).

the international literature as well as in New Zealand continue to give rise to policy advice suggesting that income redistribution must be a central component to any government policy aimed at raising student outcomes.⁸² This is true despite the recognition that income is correlated with a variety of variables about families that are difficult or impossible to measure, and that those omitted variables may be the true causes of the observed correlation. We have touched on some of these issues in our discussion of single parenthood and divorce. The parental characteristics that employers value and pay for—skills, honesty, cognitive ability, reliability—are also the same characteristics that, whether through hereditary or environmental channels, may cause children to succeed. As Blau (1999) points out, very few studies adequately address the “policy relevant” question of how child outcomes would be affected by a direct increase in family incomes. The assumption that exogenous increases in family income will lead to improvements in child welfare can hold only if parents receiving the increased income devote significant portions of that income to investments in their children and if those investments have real effects. Recent research that has begun to address these issues increasingly suggests that income, in general, is not a strong causal channel through which child outcomes improve significantly and that policies aimed at raising family incomes are likely to have only small effects. At this point, we find evidence favorable to possible impacts of family income only for very poor households and only for families with very young children, and even this evidence is still subject to further replication.

Much of this has been brought to public attention in a recent book by Susan Mayer (1997). In it, she provides some of the strongest statistical evidence to date that, at least once basic material needs are taken care of, parental income plays a relatively minor role in determining child outcomes important to policy makers. Children whose parents have certain underlying attributes seem to succeed even if their parents have low incomes. This finding reinforces our previous tenuous conclusion that the negative impacts of divorce and single parenthood are not likely to be primarily due to lower resources in divorced and single-parent households. Further recent work using increasingly sophisticated econometric techniques and published in academic journals by and large reinforces these findings by Mayer (Duncan, Brooks-Gunn et al. 1994; Haveman and Wolfe 1994; Korenman, Miller et al. 1995; Blau 1999; Blau forthcoming).

Mayer’s study is unique in that it uses five distinct statistical techniques to specifically search for evidence on the degree to which current estimates of the parental income effect on child outcomes is biased upward by omitted variables. First, she investigates the impact of certain sources of income that are less correlated with omitted variables than wages or government transfers, and she finds that such income has significantly less impact on child outcomes. Second, she compares the effect of parental income prior to the observation of a child outcome to the “effect” of parental income after the child outcome. While the former might be causal, the latter (she argues) cannot be.⁸³ Thus, to the extent that this argument holds, the relative size of an outcome’s correlation with parental income before and after an outcome can inform us how important stable unobserved parental characteristics are relative to income itself. Mayer finds that for most outcomes, these unmeasured characteristics account for much

⁸² Wylie et. al. (1999), for instance, concludes from correlational analysis using the Competent Children longitudinal data from New Zealand: “Change to the socioeconomic differences inherited by children would have the most impact [in minimizing gaps between children of different circumstances].” In careful statistical analysis using data from the Christchurch Health and Development Study, Baker and Maloney (1999) find much less significance for family income, although this study was not particularly focused to answer this question. This gives us confidence that the findings we report in this section are likely to be applicable to the New Zealand context. Other studies demonstrating correlational linkages with income in New Zealand include Densem and Keeling (1988).

⁸³ Economists will be most skeptical of this notion because parents may be planning their income earning over a life-time, which implies income after an observation may in fact have been caused by decisions before the observation.

of the observed impact of parental income. Third, she attempts to find whether specific spending patterns or specific differences in parental practices translate into differences in child outcomes. She reports little influence of within home spending pattern differences between high and low income parents.⁸⁴ Furthermore, she finds that the correlation between parental income and parental practices, at least in her sample, is not strong; lower income parents do not differ strongly in how they discipline their children, how much they interact with their children or how much they read to them.⁸⁵ Fourth, she observes that to the extent that parental income causes differences in child outcomes, trends in parental income ought to predict trends in child outcomes. Despite well-documented increases in inequality between the richest and the poorest U.S. households in the past three decades, however, Mayer finds little evidence of increasing gaps in child outcomes along income lines. Neither changes in the level of financial well being of households nor changes in the distribution seem to influence child outcomes on average. Finally, Mayer tests whether policy-induced variations in income influence outcomes. Particularly, she investigates whether the large U.S. interstate differences in welfare benefits targeted at single parent families have appreciable effects on the difference in how well single parent families in those states fare in relation to two-parent families. She again finds no evidence to indicate that this is the case, which implies that, at least at the level of material well-being of poor families in the US, additional exogenous income makes little difference for child outcomes.

As Mayer points out, one would doubt the conclusion that parental income has less impact than previously thought on child well-being if any one of these tests had been used by itself to make the argument. However, the fact that all these tests yield the same result suggests that within-home resources play only a minor role, and that the strong observed correlation between family income and child outcomes is actually caused by unobservable characteristics of families that are correlated with family income.

In the most recent and most thorough study to date, Blau (1999) is among the first to use the latest econometric techniques to attempt to directly test for the possibility of improvements in child outcomes from exogenous changes in family income. His study follows a series of studies (Hanushek 1992; Duncan, Brooks-Gunn et al. 1994; Korenman, Miller et al. 1995) that used a variety of data sets to conclude that the effects of exogenous increases in current family income are small, and that, although the effects of permanent increases in income are larger, these effects decline in importance as household characteristics enter the analysis and are generally dwarfed by the impact of those characteristics.⁸⁶ Blau's (1999) study, however, is particularly noteworthy in its breadth and its application of statistical techniques aimed at overcoming the econometric problems inherent in this kind of an analysis. Using data from the U.S. National Longitudinal Survey of Youth, he tests for the impact of exogenous increases in income on six different possible outcome measures ranging from reading and vocabulary tests to behavioral, memory and motor skill and social development problems. Furthermore, included in the analysis is a summary measure of the quantity and quality of home inputs, a measure arrived at through questionnaires as well as home observations. Additional variables include various other maternal background characteristics, and both fixed and random effects models are estimated to account for the potential endogeneity of

⁸⁴ Those within household spending items that do seem to be important, books, museum visits, etc., are not highly correlated with income. However, Mayer does not test for the importance of parental income on neighborhood and school quality and thus leaves these as potentially important indirect effects of parental income.

⁸⁵ We point out shortly, however, that these results are inconsistent with some other findings in the literature that indicate that spending patterns and time allocations of poor parents do differ from those of wealthier parents. However, it nevertheless remains interesting that such differences are not the cause of observed income effects on child outcomes in Mayer's data.

⁸⁶ Furthermore, somewhat larger effects had been reported for the lower end of the income scale.

income.⁸⁷ The conclusion from this quite thorough analysis is that, although there is an impact of increases in permanent income on child outcomes, this impact is “not large enough to make income-transfer policies a feasible method for obtaining meaningful improvements in child outcomes.” An increase of almost \$1,000 per year in permanent income, for instance, would lead to a mere 1 to 1.5 percent of a standard deviation change in outcomes. The largest effects are reported for behavioral problems, indicating (as other studies reviewed in this report indicate) that behavior is somewhat more malleable than other outcomes.

In addition to providing these estimates, Blau also discusses potential reasons why the impact of income on child outcomes might be so low. Aside from Mayer’s (1997) speculation that existing anti-poverty programs have alleviated the worst effects of low income to the point where additional improvements in income will accomplish little, Blau suggests two further possibilities. First, it may be the case that parents do not care much about the kinds of outcome measures he investigates and therefore do little to invest in improving them when faced with exogenous increases in income. Second, investments that parents do make as a result of increases in income (such as parental time, child care, books for the home, etc.) may have little impact on child development. Blau then provides tentative evidence against the first of the explanations and in favor of the second, which is consistent with other evidence we report in this literature review. More precisely, he finds that the summary index of quantity and quality of home inputs is quite responsive to increases in permanent income, thus indicating that parents do invest in children when incomes increase. Nevertheless, these investments are accompanied by relatively small improvements in outcomes which implies that they are not effective at raising such outcomes. This is, of course, consistent with evidence from behavioral genetic studies.

While these studies thus increasingly point to family income as being a relatively minor causal channel through which home environments impact outcomes, they have not entirely addressed all the potential circumstances under which family income might matter. One recent study uses the PSID data and modern econometric techniques (to adjust for biases) to estimate the impact of family income at various stages of childhood and adolescence on school completion and non-marital fertility (Duncan, Yeung et al. 1998). Results from this study suggest that family income may be important for eventual school completion in the case of very young poor children (ages 0 to 5) while it is less important later in childhood.⁸⁸ More research of the potential impact of very low family income early in life is therefore needed before one can dismiss family income entirely as a causal link for all children.

Finally, evidence supporting this increasing body of literature was recently reported from a major 10-year project aimed at reducing poverty in an experimental Milwaukee program called Project New Hope. The project offered families in the experimental group a guaranteed job working 30 hours per week that would pay sufficiently to lift them out of poverty. Only 27 percent of the main target group, however, continued in their job long enough to

⁸⁷ Three fixed effects models are presented. In one the sample is restricted to a subsample of mothers that also have a sister in the sample. The fixed effect therefore controls for unobserved variables common to the mother and her sister. In the second set of fixed effects models, a subsample consisting of families with at least two children in the sample is used. In this case, the fixed effect includes unobserved genetic or home environment factors specific to that family. In the third set, a subsample of children who took the outcome tests more than once was used to isolate a fixed effect specific to the child.

⁸⁸ The study finds little impact, however, of income at any stage on non-marital fertility. One explanation offered by the authors is that low income may disadvantage children in both behavior and aptitude levels early in life, but that school systems classify (and track) children early on in terms of aptitude but allow for greater malleability of behavior later on. Further studies investigating other types of child outcomes can be found elsewhere (Duncan and Brooks-Gunn 1997).

accomplish the goal of escaping poverty.⁸⁹ The reasons cited for ending work were drug and alcohol abuse, dysfunctional family relationships, conflicts with employers, problems with baby-sitters and cars, and simply depression. These results are consistent with Mayer's findings in that they point to the kinds of unobservables that she argues are at the heart of child problems associated with poor households. They are even more discouraging if viewed in light of the fact that both the experimental and the control group consisted of families that had volunteered for the experiment, not of families randomly chosen from the population. The one encouraging result from the experiment was that boys (but not girls) of those in the experimental group registered large gains in academic achievement as well as large improvements in behavior. While no conclusive explanation for this has been identified, researchers working on the project suspect that this is due to an increase in formal, structured child and after-school care experienced by boys in the experimental group (Huston, Duncan et al. 1999). As suggested earlier, parents in the experimental group chose such subsidized care (as well as club and community activities) for their boys but not their girls, in many cases stating that they were more worried about their boys.⁹⁰ This would be consistent with our findings in the literature that document such positive effects for at-risk children, especially boys.⁹¹ However, it is not clear from these data whether these effects will be subject to "fade-out" similar to what we reported in studies of child and adolescent care. A follow-up study is planned and will shed more light on this issue.⁹²

Again, it does need to be emphasized that these results are obtained with U.S. data, which implies the results hold only for levels of family incomes that are substantially above what is observed in some parts of the world, even for the very poorest households. At the same time, recent evidence from New Zealand also points to a considerably smaller role for family income (in explaining reading scores among 8 to 13 year olds) once a variety of statistical controls have been made (Baker and Maloney 1999), although more can be done to see whether these results will further converge to those using U.S. data. The increasing consensus emerging from several carefully done analyses, therefore, is that family income seems to do little to help child outcomes, at least once children's basic material needs are met and they have reached school age. Plenty of evidence exists that, at least in extreme circumstances when such needs are not met, basic prenatal and early child nutrition, for instance, can make a large contribution to cognitive child outcomes.⁹³

What, then, might be the omitted variables that are behind the correlation between income and outcomes? We are, once again, left with several possibilities. Mayer herself assumes that parental income is correlated with parent characteristics such as intelligence, personality and habits, and that differences in these characteristics create fundamentally different home environments for children growing up in poor and rich households. Given our previous evidence as well as the evidence in studies like Blau's that suggest little impact from additional within home investments, there is substantial doubt that this is the whole story.

⁸⁹ Average income from all sources was higher in the first year of the program for the experimental group but statistically indistinguishable after the second year (Huston, Duncan et al. 1999).

⁹⁰ In fact, even after the gains made by boys, their performance was still slightly below that of girls (Huston, Duncan et al. 1999).

⁹¹ Researchers are particularly confident that structured care is the reason for higher achievement among boys because there was no evidence of a change in parenting practices in the experimental group (Huston, Duncan et al. 1999).

⁹² Effects from the Milwaukee project are reported in DeParle (1999) and summarized more formally in Huston et al. (1999). More details provided by the Manpower Demonstration Research Corporation (<http://www.mdr.org/>).

⁹³ Baker and Maloney's (1999) work on New Zealand further indicates potential interactions with ethnicity, although some are counterintuitive. In particular, they find that family income has a negative impact on reading ability for Maori and Pacific Island children and a positive impact on Pakeha children. More work is required before conclusions can be drawn from this finding, particularly given the under-representation of certain ethnicities in the Christchurch Longitudinal data set.

First, at least some of these parental characteristics are in part heritable. Second, these characteristics are likely to cause households to reside in different neighborhoods, implying that environmental factors outside the home environment may also be consistently different. Mayer has little to say on such issues, and it is left to other researchers to sort them out more thoroughly. Nevertheless, her main point that parental income seems less of a causal factor in generating positive child outcomes than commonly thought remains valid and is increasingly becoming conventional wisdom.

5.3.3 *A Brief Note on Maternal Education*

Family income is similar to another variable often included in regressions as a control variable: maternal education. Because maternal education is a variable less easily influenced through policy, we will not take the reader through the various studies that find large correlational effects for maternal education on child outcomes but acknowledge that they are at least as many in number as those studies finding an effect of family income. The same methodological criticisms leveled against interpreting correlational evidence on family income as causal applies, however, to correlational evidence on maternal education.

5.4 *Summary*

Much of the recent literature has become more sophisticated in its attempts to isolate methodological problems of prior studies. In the process, the message that seems to emerge is that, consistent with behavioral genetic evidence, shared home environments are less important to child outcomes than had previously been indicated by largely correlational studies. We have reviewed the evidence in regard to family structure, parental time resources, parenting styles, household incomes, etc., and each time we have either found emerging evidence indicating less of a causal impact than early studies had suggested, or that there were sufficient methodological difficulties to cast a shadow of doubt over current studies. In some cases, however, we found more suggestive evidence in favor of home environmental impacts for lower income children and children in some minority households. This would not contradict the general conclusion from Section 4 that shared home environments explain little of the variation in outcomes for the broad middle of the distribution. Similarly, the frequent finding of fade-out, especially as it regard cognitive ability, is consistent with behavioral genetic evidence of a larger role for shared home environments early in childhood than during adolescence. Finally, it should be emphasized that these findings are for large data sets that draw from many different types of households, and they do not imply that parental influences might not be extremely important in some restricted set of instances.⁹⁴

⁹⁴ Thornton (1995) offers one of the more dramatic such cases in which a New York ditchdigger's desire to have his five daughters become doctors caused him to create a highly structured home environment in which peer influences outside the family were not permitted. All of his daughters became successful professionals, including two doctors and one lawyer.

6 Causal Links Between School Resources and Child Outcomes

In this section, we review the evidence on the contribution of primary and secondary schools to child development. In particular, we ask to what degree parents who get involved in their children's schools improve outcomes, what effect peers within schools have on one another, and how other inputs such as class size and teacher attributes impact school quality (and thus child outcomes). These are difficult questions that call for statistical techniques which are often not utilized by researchers in this area. While some conclusions can nevertheless be made, they are more tentative and speculative than one might hope. However, unlike in Section 5 where increasingly sophisticated analysis was generally found to yield smaller estimates of home environment effects, we now find that more sophisticated techniques can also lead to larger and more significant results.

Parental involvement in schools represents one such example. While recent carefully controlled studies are more promising, the general enthusiasm for increased parental involvement in schools finds relatively little support in a sea of earlier studies that report rather mixed results and that rarely tackle the potential biases introduced by the research methods employed. For example, parents who become involved in their children's school, as well as parents who choose schools in which overall parental involvement is high, are likely to differ in various unobserved ways from other parents, a fact that would cause us to overestimate parental impacts on schools. At the same time, one factor that might motivate parents to become involved in schools is poor student performance, which might cause us to underestimate the impact of parental involvement. Of the few recent studies that take these potential biases into account, either through the inclusion of a rich set of control variables or through the use of more sophisticated statistical techniques, one finds stronger evidence for the proposition that parents who become involved in schools benefit their own children a lot while benefiting the overall school much less, while others find the opposite result. It is therefore too early to judge the precise nature of parental involvement effects in schools, even if the best evidence is suggestive of at least a moderate, positive impact of parents getting involved.

Studies of the potential of peers to affect each other within schools and classroom are even more challenging to interpret. One difficulty faced by researchers relates to the appropriate definition of the relevant peer group, a definition that is too often dictated by data constraints. Many studies therefore investigate whether the degree of mixing of different SES students within an entire school has significant impacts on student outcomes (separate from impacts of control variables), and we find little consistent evidence for such "school-wide" peer effects. A second set of studies investigates the extent to which different peer mixes induced by across-classroom tracking (or streaming) impact outcomes. These studies are particularly difficult to interpret because they confound the impact of different peer mixes with the impact of different teacher and classroom characteristics. Often, for instance, such tracking is specifically designed to teach different materials to different tracks (with different types of teachers), yet the same test is used to assess how well students in different tracks are doing. While these tracking studies therefore find evidence that is interpreted to favor the presence of peer effects, an equally plausible interpretation is that they are finding classroom and teacher effects. A distinction between these competing interpretations is important for policy makers, because under a "peer effects" interpretation mixing is predicted to lead to an improvement of low SES outcomes while under alternative interpretations it might not. Virtually no effort is made to draw such a distinction in tracking studies, which causes us to conclude that current evidence generally neither supports nor denies the existence of peer effects in secondary schools (where most across-classroom tracking occurs). A final set of studies investigates peer effects in within-classroom groupings in elementary schools. In part because many of these studies are based on relatively controlled experiments, these studies provide the strongest evidence for peer effects (at least in elementary classrooms), evidence indicating that mixing of children within elementary classrooms raises performance for low

SES children more than it lowers performance by high SES children. However, even in these studies it is not always clear whether the causal channel is a pure peer effect or whether the changes in teaching approach that accompany the formation of within-class groupings is primarily at work. Nevertheless, the policy implications from this line of research more strongly favor within class peer group mixing in elementary schools. Even if the positive effect comes primarily from different teaching methods triggered by the formation of within-class peer groups and not from the peer groups themselves, information on the aggregate effect is still valuable in guiding elementary school teachers. Relatively recent theories arguing that children themselves are more involved in peer selection than previously thought have not yet been subjected to rigorous empirical testing.

Finally, we briefly review the most recent evidence on school inputs other than peers and parents. A large literature attempting to link inputs such as class size and teacher attributes to student outcomes has been unable to consistently document large effects, but much of this literature is plagued by methodological problems that are likely to bias results. While some recent work has indicated that wages are more responsive to these school inputs than traditional measures of student outcomes, much of this evidence has been called into questions on both methodological and data grounds. What remains from these statistical studies is a great deal of uncertainty and a sense that the school inputs commonly measured do not contribute much to student outcomes. Given the many methodological issues, however, true experiments would be extremely valuable to the resolution of this uncertainty. Only recently has experimental evidence of this kind come to light with respect to class size. This evidence indeed suggests that the previous pessimism regarding any impact of class size on student outcomes was unwarranted. Class size, at least, does seem to contribute positively to such outcomes.

Given that we are fairly certain of a relatively large environmental component to child outcomes, and given the weakening evidence for the proposition that shared home environments are the major factor in this component, we now move to consider the impact of schools. Following Figure 1, we will do so in three steps. First we investigate in Section 6.1 the extent to which parental involvement in schools may benefit children. Second, we consider the impact that peers have on one another in the school environment, focusing on their influence on school outcomes. Third, we briefly touch on issues relating to other school inputs. Although the latter is not the main focus of this review, it will be useful to contextualize the results of other sections.

6.1 Parental Involvement in Schools

Although families and schools have independent effects on student outcomes, there is a widespread belief that interactions between the two may yield additional benefits for students. In analyzing the data for 8-year olds in the New Zealand “Competent Children” study, for instance, correlational evidence suggests that “voluntary parental involvement with schools ... made a difference [to the competence demonstrated by children at age 8]” (p. xix). Henceforth referred to as “parental school involvement”, these interactions assume a variety of forms, ranging from parental attendance at teacher conferences to participation in local school councils. By its nature, however, parental involvement is “voluntary”, implying that some parents choose to involve themselves in school activities, and others do not (hence, parental school involvement is represented as a double-arrow in the conceptual model of Figure 1).⁹⁵ The choices made by individual parents could have consequences for their child’s outcomes, as well as those of other parents’ children. A growing empirical literature in the social sciences attempts to assess the magnitude of these effects.

⁹⁵ It is also important to note that schools may play a role in actively encouraging or discouraging parental involvement.

The aim of this section is threefold. First, it delineates several ways in which parental involvement is manifested in schools (focusing on the U.S. experience). Second, it describes some perspectives from sociology and economics which explain why some parents choose to become involved and others do not. It will also examine the empirical evidence on which parents are most likely to be involved. Third, it will assess the growing body of empirical evidence which relates measures of parental involvement to child outcomes such as academic achievement.

6.1.1 The Form of Parental Involvement

Parental involvement has been conceptualized in several ways. Epstein (1987) suggests two general categories of parental involvement: communication and direct involvement.⁹⁶ The first includes communication between families and schools about school structure and events (e.g., schedules of events, school goals, programs and services, rules) and about children (e.g., academic progress, parent-teacher conferences). The second includes participation in the parent-teacher association, direct assistance of teachers and administrators in the school building, attendance at student assemblies or presentations, and attendance at special workshops or discussion groups for parents.

In contrast, economists are more inclined to view parents as consumers, and their involvement as a means of ensuring that schools act in accordance with parental preferences (e.g., McMillan 1999). According to this view, teachers and administrators may not always maximize school quality. In such a case, parents may choose to exert pressure on teachers and administrators—perhaps via a parent-teacher association or by spending time in the school—in the hopes of raising school quality. Note that some of these types of parental involvement (providing information, organizing events, applying pressure) are privately provided public goods within the school. That is, an individual parent's involvement benefits not just that parent's child but also the school at large. Others (such as attending personal meetings with the teacher) are largely private goods in the sense that the primary beneficiary of the parent's involvement is that parent's child.

However it is viewed, parental involvement is widespread, at least in the U.S. (Carey, Lewis et al. 1998). Most schools initiate contact with parents to inform them of school curricula and overall performance on standardized tests. Almost all schools hold a variety of activities open to parents, such as back-to-school nights, parent-teacher conferences, and academic exhibitions. However, only about half of schools indicated that parental attendance was substantial. Many schools reported having an advisory group or policy council that includes parents, but a fairly small percentage of schools indicated that parental input is actually considered to a great extent in making decisions. About half of schools make use of voluntary, written contracts that define the mutual responsibilities of parents and schools. Our impression is that parental involvement in New Zealand is also high, and raising it is part of the policy goal behind the "Tomorrow's Schools" reforms a decade ago.⁹⁷

⁹⁶ Epstein, as other authors (e.g. Sui-Chu and Willms 1996), uses a broader conceptualization of parent involvement that includes home-based learning activities. Much of this research is covered in other sections of the review. In this section, we restrict our analysis to parental involvement in school activities. See Muller and Kerbow (1993) for a further description of how parents may become involved in schools.

⁹⁷ Beyond the usual forms of parental involvement, there are numerous special reforms that aim to increase involvement beyond current levels, or alter its form. The intent of such reforms is often to increase parental responsibilities in school governance and decision-making. James Comer's School Development Program creates a School Planning and Management Team that includes a small number of parent representatives (Comer and Haynes 1991). The team is responsible for designing and implementing a comprehensive school plan. Other parents are involved in day-to-day activities in the classroom and parent organization. Henry Levin's Accelerated Schools Project creates a series of

6.1.2 Which Parents Choose to Become Involved?

Some parents choose to become involved in their child's school, and others do not. Many authors express concern that less-involved parents will typically be those of lower socioeconomic status (Lareau 1987; Stevenson and Baker 1987; Lareau 1989; Sui-Chu and Willms 1996). This may be important for two reasons. First, if the involvement of parents improves schools overall (as opposed to improving outcomes for just the involved parents' children), different average levels of parental involvement across schools may result in differences in school quality (linked to socioeconomic characteristics of the student body). Second, if the involvement of parents primarily benefits the involved parents' children but not the school overall, then children within the same school will receive different school qualities (linked to individual socioeconomic characteristics). This section explores several frameworks for understanding the different choices made by families, particularly those of lower socioeconomic status. It then examines the empirical evidence on which families tend to become most involved in schools.

The sociologist Annette Lareau (1987) describes several approaches to explaining the differential involvement of high and low socioeconomic families in schools.⁹⁸ First, there may exist a "culture of poverty", in which working-class families place less value on education and are, therefore, less inclined to participate in school activities. Second, institutional features of schools may discourage some parents from becoming involved.⁹⁹ For example, schools may be more welcoming to middle-class than to working-class parents. In some cases, schools have been found to discourage parent involvement in the school and classroom, because of a perception that parents are "busy or disinterested or 'ignorant'" (Epstein and Dauber 1991; Eccles and Harold 1993). Third, parents of lower social classes may lack the "cultural capital" needed to pursue active involvement. Schools are fundamentally organized as middle-class institutions with middle-class values. In this context, working-class parents may feel uncomfortable interacting with teachers and involving themselves in school activities (Sui-Chu and Willms 1996).¹⁰⁰

Though little work has been done by economists on this issue, they might assume a different perspective. Involvement would be higher where the expected costs to the parent who is considering becoming involved are adequately compensated by the expected benefits for that parent. In this case, parental costs of involvement include time requirements and monetary costs in the form of transportation or child-care expenses.¹⁰¹ From this perspective, parents are more likely to become involved if their involvement directly translates into benefits for their children (as opposed to the school overall). Benefits include higher student outcomes and, perhaps, the enjoyment derived from participating in school activities. The perceived

school committees, including parents, that are responsible for diagnosing school problems and designing solutions (Levin 1997). The 1988 Chicago School Reform Act created mechanisms for direct parental participation in school governance (Downes and Horowitz 1995). Each school is governed by a council composed of six parents, two community members, two teachers, and the principal. The council is empowered to hire and fire the principal, allocate the budget, and adopt a School Improvement Plan. Similar institutions were introduced in New Zealand as part of the "Tomorrow's Schools" reforms and continue to operate (Fiske and Ladd 1999). All this suggests that administrators, teachers, and academics suspect that parental involvement within schools may be of significance.

⁹⁸ Also see Sui-Chu and Willms (1996).

⁹⁹ For general discussions, see Eccles and Harold (1993), Hoover-Dempsey and Sandler (1997), or Lareau (1987).

¹⁰⁰ Analogous barriers may exist for families whose ethnically based culture, rather than class based culture, is different from the dominant culture that has shaped school institutions.

¹⁰¹ See Kerbow and Bernhardt (1993), who discuss the importance of resource and time constraints on parental decisions.

costs and benefits might be affected by a number of factors. Families of lower education may simply perceive fewer benefits from student outcomes, akin to the “culture of poverty” argument. Alternately, such families may be less adept at exerting pressure or navigating school bureaucracy, lowering the benefit derived from a given exertion of effort. The benefits of “pressuring” school officials through greater involvement might be large if a parent perceives that their child’s school is of exceptionally low quality (McMillan 1999). In weighing the costs and benefits of becoming involved, families of high and low socioeconomic status face different constraints. Low-income or single-parent families, for example, may be unable to obtain child-care in order to attend school functions.

Several authors correlate family background with measures of parental involvement (see Table 6.1.2 for a summary). They use a wide variety of measures to proxy for parental involvement, such as the percentage of parents that are active in the parent-teacher association and the frequency of contact between school officials and parents. In comparing the various studies, we have sought to transform their estimated effects into a common metric. When possible, we provide the effect on parental involvement (in standard deviation units) of a one standard deviation increase in the measure of family background.

Three studies use the large NELS data set of eighth-graders in the United States, all finding that parental education and income are positively correlated with parental involvement. Sui-Chu and Willms (1996) and Kerbow and Bernhardt (1993) find somewhat smaller effects than McMillan (1999), perhaps because the first authors also include measures of race and ethnicity in their regression. Only one study finds no relation between SES and involvement (Zellman and Waterman 1998). This may be attributable to its small sample size and collinearity between SES and other independent variables such as race and ethnicity.

Sui-Chu and Willms (1996) find that Asians are somewhat less likely to be involved than white parents, consistent with other research on Chinese-Americans parents in the United States (Hidalgo, Siu et al. 1995). Kerbow and Bernhardt (1993) replicate this finding with the same data. The differences between whites and other ethnic groups are rather muted. In some of the prior analyses, blacks appear slightly more likely than whites to engage in communication with their schools.¹⁰² All this is remarkably consistent with reports from New Zealand. In particular, cultural barriers arising from a generally deeper respect for authority seems to keep Pacific Islander parents in New Zealand from involving themselves more in schools, often citing respect for the authority and expertise of teachers. Pakeha and Maori parents, on the other hand, exhibit no such cultural barriers.

There is scarce evidence on how other indicators of family background relate to parental involvement. Two-parent families appear somewhat more likely to participate in school activities, but less likely to communicate with schools (Sui-Chu and Willms 1996). Linney and Vernberg (1983) cite some evidence that working mothers are slightly less likely to be volunteers in their children’s school activities, although none of the studies in Table 6.1.2 include this variable.

The most robust pattern from existing research is that parental education and income are positively correlated with parental involvement, after controlling for a wide variety of independent variables. A conclusion with limited support is that race and ethnicity are somewhat less important determinants of involvement—with the exception of Asians—once SES is properly controlled. Unfortunately, these simple correlations provide few clues as to the underlying mechanisms that promote or inhibit parental involvement. They are consistent with many alternative explanations. For example, they might indicate that higher income families are more involved because they are able to absorb the costs of participation in school

¹⁰² While Zellman and Waterman (1998) find the contrary, they also have a much smaller and less representative sample.

activities (such as child-care or transportation) or they might reflect a greater value placed on student outcomes by families of higher socioeconomic status.

6.1.3 *Effects of Parent Involvement on Outcomes*

Despite a fairly small body of empirical evidence, there is faith among many researchers that parental involvement leads to increased student outcomes. Purkey and Smith (1983), in their influential review of “effective schools” research, name parental involvement as one of the important organizational factor in schools: “our feeling is that parent involvement is not sufficient, but that obtaining parental support is likely to influence student achievement positively” (p. 444). But an early review by Linney and Vernberg (1983) notes that “despite the widespread belief that such involvement is beneficial to children, there is limited research systematically examining the context and form of parent participation and its role in enhancing school achievement and adjustment” (p. 78). Despite some additional work since the early 1980’s, Epstein’s (1992) review also laments the lack of systematic research that relates parental involvement to student outcomes.¹⁰³

Table 6.1.3 summarizes the results from 10 studies that we were able to locate. They were chosen because they use similar proxies of parental involvement, such as the participation in parent-teacher associations and the frequency of parent communication with the school. Most use regression analysis to relate these proxies to academic performance, holding constant a series of background variables such as family SES. We report the standard deviation increase in outcomes that is due to a one standard deviation increase in the measure of parental involvement. Several of these studies find positive effects of such involvement on academic performance (Iverson, Brownlee et al. 1981; Stevenson and Baker 1987; Paulson 1994; Griffith 1997). About the same number find a mix of weakly positive, statistically insignificant, or even negative effects (Epstein 1991; Reynolds 1992; Muller 1993; Sui-Chu and Willms 1996; Zellman and Waterman 1998).

Are the estimates from these studies unbiased? We have already suggested that the decisions made by parents to become involved are non-random. As such, the prior studies attempt to control for the background characteristics of families. But the families that typically choose to become involved may possess unobserved characteristics that also influence outcomes. If these characteristics are omitted from regressions, the estimated effects of parental involvement will be biased. It is difficult to assess whether the direction of bias is positive or negative. High SES families appear more likely to become involved, and their children have higher achievement. If key measures of SES are omitted as regressors, then it is likely that estimates of involvement are biased upward because they spuriously reflect family background.¹⁰⁴ On the other hand, families may tend to become more involved when their child experiences problems at school, such as poor behavior. But if the child’s background and behavior are not adequately controlled for, then greater involvement will appear to be negatively related to student performance (a downward bias).¹⁰⁵

¹⁰³ Some reviews of the literature do suggest that evidence is supportive of links between parental involvement and student outcomes (Henderson 1987; Eccles and Harold 1993; Muller 1993; Zellman and Waterman 1998). In one case (Henderson 1987), most of the studies reviewed are mainly related to parental involvement in the home, rather than the school. In all of these reviews, authors appear to uncritically accept the causal validity of correlational studies. As we will suggest, this is unwarranted.

¹⁰⁴ Two of the four studies that find positive effects (Iverson, Brownlee et al. 1981; Paulson 1994) use minimal controls for student background. In these cases, upward bias in the estimates of parental involvement appears likely.

¹⁰⁵ Both Sui-Chu and Willms (1996) and McMillan (1999) make similar points.

Given these concerns, we are more likely to have confidence in studies that control for a rich set of family and child characteristics. Using the nationally representative NELS data set, Muller (1993) makes extensive controls for family and student background and finds only mixed evidence that parental involvement raises achievement and grades. Using the same data and similar controls, Sui-Chu and Willms (1996) are unable to identify important effects of individual parent involvement on achievement. They do find a fairly small effect of school-wide participation on individual children's reading achievement, thus indicating that much of the effect of parental involvement is of a public good nature. Their results are suggestive that students benefit the most when many parents in the school participate simultaneously. Sui-Chu and Willms also estimate simpler models excluding indicators of parental involvement. In these models the estimated effects of SES change very little, suggesting that only a small amount of the "overall" SES effect is accounted for by the greater propensity of high-SES families to participate in schools.

Of the 10 studies, only McMillan (1999) uses more sophisticated statistical techniques to correct for potential bias. Also using the NELS data, he attempts to identify an instrumental variable that is correlated with parental involvement, but uncorrelated with unexplained variance in academic performance. He posits that the extent of parental participation in non-school organizations is a variable that fulfills these criteria. After making this correction, he finds larger effects of parental involvement than Sui-Chu and Willms, suggesting that standard regression estimates are biased downward. The effects of individual involvement on achievement are twice the size of the effects of school-wide parental involvement. Thus, in contrast to Sui-Chu and Willms, the results indicate a large private component to parental involvement in schools. Without further corroboration, these conclusions should be taken cautiously, for they may not hold if his instrumental variable is even slightly correlated with unexplained variation in student outcomes.

There are few consistent lessons to be drawn from the large set of existing empirical evidence. Correlational studies have produced a good number of negative and statistically insignificant effects, but many of these studies use small or non-randomly selected data sets. Of the studies using the large and nationally representative NELS data set, Sui-Chu and Willms and Muller find mixed results. Their evidence is also suggestive that parental school involvement accounts for a quite small portion of the overall family effect on student outcomes. The more positive results of McMillan (who employs more sophisticated statistical techniques) suggest that typical estimates of parental involvement are biased downward. This conclusion is promising but awaits further replication. We should note, however, that his results indicate a larger private, rather than public, component to parental involvement. That is, the effects of each parent's involvement mainly accrue to their own children, rather than those of others. Individual benefits for children might even come at a cost to others in the same classroom; for instance, one parent's pressure may encourage a teacher to devote additional time to one child and less to others. From a policy perspective, this provides good reason for caution. Policies that encourage involvement of some parents (but inevitably fall short of reaching every parent) might have unintended distributional consequences within the classroom or school. Clearly it provides good incentives for us to obtain a better empirical understanding of the mechanisms, whether public or private, which relate parental involvement to student outcomes.

6.2 *School Peer Groups*

A prominent theme in the sociology of education is that students are influenced by features of the school context, particularly the composition of their peer groups within the school. Perhaps a student's achievement rises when she attends class with relatively better-achieving students. Or perhaps a student is less inclined to drop out of the school when he attends class

with peers of a relatively higher socioeconomic status.¹⁰⁶ Beginning with the Coleman report in the 1960s (Coleman, Campbell et al. 1966), hypotheses about “peer effects” have spawned a voluminous literature in several disciplines that relates student outcomes (such as academic achievement) to a wide range of peer measures (such as race and ethnicity, socioeconomic status, and the outcomes of other students). This practice continues in studies of schools around the world, not just in the U.S. In New Zealand, for instance, one recent report by the New Zealand Council for Educational Research interprets correlational evidence to imply that “the socioeconomic composition of a school makes an additional contribution to children’s Literacy, Mathematics and Logical Problem-Solving competency levels, over and above their individual family income” (p. xix) (Wylie, Thompson et al. 1999).

Within the broad category of peer-effect studies, researchers in economics, education, psychology, and sociology vary in their research questions and methodologies. We can distinguish among three general approaches, in ascending order of specificity. In the first, typically used by economists and sociologists, school-wide measures of peer-characteristics—such as the percentage of minority students or average student achievement—are correlated with individual outcomes. Either implicitly or explicitly, these studies presume that the entire student body constitutes a relevant “peer group,” though this decision often seems motivated by data limitations rather than a well-defined theory of peer interaction. The second group of studies, mostly in education and sociology, examine the effects of peer group variation across classrooms. A common policy in many countries is the “tracking” or “streaming” of students into classes that vary by curricula (vocational or academic), ability-level (advanced, intermediate, or remedial), or another criterion. While tracking could alter many aspects of student classrooms—including the quality of teachers and the pace of instruction—we shall focus on how it alters a student’s immediate peer group, and how this affects student outcomes. The third group of studies, almost entirely conducted in education, assesses the effects of peer group variation within classrooms. A common instructional technique in many elementary school classrooms (both in New Zealand and elsewhere) is the use of small learning groups, organized by either homogeneous or heterogeneous ability levels. Varying group composition may affect student outcomes, independently of other classroom characteristics.

Before assessing the empirical evidence, the next section describes some conceptual approaches to explaining the causal connection between peer group composition and student outcomes. These will prove helpful in interpreting empirical studies. The next sections review the accumulated empirical evidence for the three categories of studies. In each section, we gauge whether some of the methodological concerns outlined in Section 3 have been adequately addressed and, therefore, the confidence we might have in attributing causal meaning to results. A final section attempts to resolve any contradictions among the three classes of studies, and draw out common lessons about the relation of school peer groups to student outcomes.

6.2.1 *Causal Mechanisms Underlying “Peer-Group” Effects*

Before evaluating the empirical evidence, it is useful to delineate more carefully what is meant by a “peer” effect. Many studies presume that correlations between peer characteristics and individual outcomes result directly from student interaction in the school or classroom (Summers and Wolfe 1977; Henderson, Mieszkowski et al. 1978). This, however, is a hypothesis rather than a finding. In fact, there are many conceivable routes through which

¹⁰⁶ Or perhaps a student who is on the margin of dropping out of school may choose to stay in school if the environment there is sufficiently pleasant, which is more likely if peers are of higher socioeconomic status with parents who demand greater amenities.

peer characteristics could affect outcomes.¹⁰⁷ These can be roughly categorized into student-centered versus school-centered explanations. Empirical frameworks rarely succeed in disentangling which mechanism is most important, yet this is fundamental for drawing policy implications.

There are several varieties of student-centered explanations. Lou et al. (1996) suggest that lower-ability students may benefit from having higher-ability students in their groups, because the latter provide “timely and elaborated assistance and guidance” (p. 449).¹⁰⁸ Likewise, higher-ability students may also benefit from such interactions because they can “clarify and organize their own learning better” (p. 449). This explanation conforms most closely to the typical intuitions of researchers and consumers of this research: that “peer effects” do indeed flow from direct peer interactions in schools and classrooms.

Another student-centered explanation suggests that groups composed of high-ability students provide a “social [setting] in which individual children evaluate their performance and internalize academic norms, thus forming expectations for their academic performance” (Pallas, Entwisle et al. 1994, p. 27). Exposure to high-ability peers tends to re-orient the expectations that students have for their own success in the classroom. More concretely, higher SES schools may provide more opportunities for students to develop peer networks of better “quality.” In these schools, “any individual student is more likely to establish friendships with high SES classmates” (Alexander, McDill et al. 1979, p. 223).

A second class of explanations centers on schools, namely on how the differing classroom experiences of high- and low-SES groups of students may alter outcomes. Pallas et al. (1994) suggest that ability group placement may influence the “quantity, quality, and pace of instruction and hence of learning” (p. 27). In other words, students of higher abilities will tend to benefit from their higher-ability peers, but in an indirect fashion, because teachers and schools provide different experiences for higher-ability students. In this vein, Dreeben and Barr (1988) show that classroom composition has important effects on moderating the quantity and quality of instruction provided to students in different reading groups.

In sum, there are many possible explanations for the observed patterns of correlations between characteristics of school peer groups and the outcomes of individual students.¹⁰⁹ Some may be true “peer” effects because they are rooted in student interaction. For example, students in small-groups may assist one another in learning, or good peers may influence the behavior of others by serving as role models. Other explanations suggest that different peer group compositions may simply influence how students are treated by teachers and schools. For example, higher SES students may receive more and better classroom instruction.

The literature on peer and neighborhood effects does not always distinguish among these explanations. As Jencks and Mayer (1990) observe, the true nature of peer effects often remains a “black box” because empirical studies are rarely capable of directly testing

¹⁰⁷ One review of the peer-effect literature that “it is not clear whether this effect comes from the influence of peers, school climate, teaching conditions, or differences in teacher expectations and curriculum...” (Levin 1998, p. 382). Also see Rumberger and Willms (1992) who observe that the overall effects of school “context” could “stem from several factors: peer effects, which may result from either face-to-face interactions among students or symbolic interactions, such as competition, emulation, or identification associated with students’ primary reference group; the academic, social, and disciplinary climate in the school; and teachers’ expectations” (p. 379).

¹⁰⁸ Also see Webb (1982).

¹⁰⁹ In their extensive review of the literature on neighborhood effects, Jencks and Mayer (1990) describe several avenues through which variation in peer or neighborhood characteristics may influence a student’s outcomes. See Section 7.1 of this review for a discussion.

hypotheses about the distinct causal mechanisms. In this review, we shall attempt to carefully state when effects are directly attributable to peers or some other influence, or whether we cannot adequately separate among the multitude of possible explanations.

6.2.2 *School-Wide Peer Influences*

Many researchers use non-experimental data and statistical methods such as regression analysis to relate peer measures to outcomes, while holding constant other variables such as family and student socioeconomic background. The choice of peer measures is often ad-hoc and lacking in strong theoretical or empirical rationale (Jencks and Mayer 1990). Typical choices include the average socioeconomic status and achievement of other students or the distribution of racial and ethnic characteristics in the student body.¹¹⁰ The intent of regression analyses is to isolate the unique contribution of peer characteristics to student outcomes. For several reasons discussed in Section 3, this is often fraught with methodological difficulties. Thus, following the discussion of the existing empirical evidence, we shall evaluate the evidence in light of methodological challenges.

6.2.2.1 *Reported Research on School Peer Effects*

One of the first, and still largest-scale efforts to assess the magnitude of peer effects was that of James Coleman and his colleagues (Coleman, Campbell et al. 1966). In carefully reviewing the Coleman results 25 years later, Jencks and Mayer (1990) conclude that school-level SES has limited effects on the academic achievement of white students, after measures of individual SES are controlled. However, they estimated effects are larger for black students. If the mean SES of students in a school rises by one standard deviation, then black student achievement can be expected to rise between 0.08 of a deviation (in the ninth grade) and 0.14 (in the twelfth-grade). This may be consistent with one of the themes of this review—that environmental effects are more likely to be found to “matter” in cases that are far from the typical. Black students tend to come from less advantaged families and neighborhoods, which may open more of a window for peer effects to make a difference. Nevertheless, the proxy of school SES—the proportion of families in the school that owns an encyclopedia—is rather imprecise and could be interpreted in a number of other ways.

Since the Coleman report, a wide variety of studies have used essentially the same methodology. The majority have sought to link peer attributes to measures of academic achievement (see Table 6.2.1 for a summary), though a few have analyzed other outcomes such as overall attainment (see Table 6.2.2). In comparing these studies, we have sought to transform their reported effects into a common metric. Thus, we report the change in outcomes associated with a one standard deviation increase in the peer measure. In the case of achievement, outcome changes are also expressed in standard deviation units. For attainment, effects are expressed as the change in the probability of dropping out. When it was not possible to transform the reported effects, we simply report whether the effect was positive, negative, or statistically indistinguishable from zero.

Of the 11 studies in Table 6.2.1, several consistently find that high levels of average SES or achievement are associated with higher individual achievement (Summers and Wolfe 1977; Henderson, Mieszkowski et al. 1978; Willms 1986; Link and Mulligan 1991; McEwan 1999). The estimated effect sizes range from around 0.1 of a standard deviation (Link and Mulligan 1991), to 0.4 (Willms 1986). Other studies yield some positive results that are, nonetheless, inconsistent enough to give pause. Caldas and Bankston (1997) find that mean SES increases achievement, but that the mean family incomes—proxied by the percentage eligible for free-

¹¹⁰ The school was, for example, the unit of analysis leading New Zealand researchers to find peer effects related to which decile school a child attends (Wylie, Thompson et al. 1999).

and-reduced lunch—are negatively associated with outcomes (multicollinearity might explain their result, though their sample size is sufficiently large to suggest that this is not a factor). Zimmer and Toma (1997) use several data sets and measures to test for peer effects in several countries. They find a mix of positive (New Zealand, Ontario, United States) and statistically insignificant peer effects (Belgium, France). Using a unique longitudinal data set, Robertson and Symons (1996) find some positive peer effects on achievement, but none on long-term earnings. Their study is, notably, the only one to assess the latter outcome. Bryk and Driscoll (1988) find a rather strong effect of mean SES on achievement (0.56) that is counter-balanced by a negative effect of increasing mean achievement. Gamoran's (1987) analysis, which uses the same data as the prior study and controls for a different set of independent variables, concludes essentially the opposite. Finally, Winkler (1975) finds that the percentage of low-SES students tends to lower white achievement, but not that of black students (a conclusion that contradicts results from the Coleman report).

There is only limited evidence on the effects of school-wide SES and achievement on other outcomes such as attainment. Both Mayer (1991) and Gaviria and Raphael (1997) find that advantaged peer groups tend to lower the probability of dropping out. Though interpretation of the relative magnitudes of these effects is subjective, Mayer's results seem to imply that these are somewhat small effects. In the case of Gaviria and Raphael, it was not possible to estimate the change in probability for a one standard deviation change in peer attributes because the standard deviation of the corresponding peer measure was not provided. Finally, Bryk and Driscoll (1988) do not find the predicted influences of mean SES and achievement on attainment.

In addition to mean SES and achievement, a common measure of peer attributes is the percentage of students that belong to various racial and ethnic categories. The presumption is usually that minority students will benefit from integration into school populations that are predominantly white or majority-race, holding constant other variables such as individual socioeconomic status and school characteristics. Examples of factors that might generate this result are (1) differences in the political power that parents of different ethnicities can bring to bear to improve school quality, (2) exposure to different cultural norms, or (3) the effect of integration on self-perception and feelings of social marginalization. Indeed, this was a major conclusion of the Coleman report in 1966 and spurred much of the effort in the United States to desegregate schools (Coleman, Campbell et al. 1966). In reviewing the Coleman results, Jencks and Mayer (1990) conclude that blacks in schools that are 90 percent white tend to score about 0.3 standard deviations higher than those in all-black schools, all else equal. However, the effect holds only for blacks in the northern United States. The results also suggest that whites benefit from being in schools that have a higher percentage of white students. Thus, if these results are to be believed, the process of school integration may create some winners and some losers.

Subsequent correlational studies have also explored how the concentration of minority students tends to affect outcomes (see Tables 6.2.1 and 6.2.2). One study in Chile finds that greater numbers of Native American students tends to lower achievement, all else equal, though the estimated effect sizes are quite small (McEwan 1999). With data from Louisiana, Caldas and Bankston (1997) find that the percentage of black students in a school tends to lower achievement. The results from other empirical studies are surprisingly inconsistent. An oft-cited analysis of peer-group effects in Philadelphia finds that student achievement is highest when the student population is racially balanced (40 to 60 percent black), rather than highly segregated in favor of blacks or whites (Summers and Wolfe 1977). Gamoran (1987) simply finds no consistent race effects, while Bryk and Driscoll (1988) and Winkler (1975) find that increasing numbers of black students tends to increase achievement, the opposite of

hypothesized effects. Finally, Link and Mulligan's (1991) results are suggestive that black students may benefit from being around other black students, rather than white students.¹¹¹

Further evidence on the effects of racial composition is provided by experimental and quasi-experimental studies of the effects of racial desegregation plans in the United States. This literature has been extensively reviewed (e.g., St. John 1975; Mahard and Crain 1983; Cook 1984) and the large number of reviews are synthesized in Schofield (1995) and Jencks and Mayer (1990). Schofield concludes that black reading achievement rises between 0.06 and 0.26 of a standard deviation after one year of desegregation. However, there are no consistent effects of desegregation on mathematics achievement among black students. There are only a limited number of studies that examine how desegregation affected the achievement of white students. From these, Schofield concludes that white achievement is relatively unaffected; while Jencks and Mayer note a similar pattern, they remain agnostic in their conclusions, given the relatively small sample sizes of most studies and the difficulty of finding statistically significant effect sizes. A common theme in many studies is that desegregation may be most effective when it is conducted in the early elementary school years. Furthermore, both Schofield and Jencks and Mayer observe that effect sizes are only pertinent to an evaluation of a single year of desegregation. Ultimately, we are interested in the net effects of desegregation on student outcomes over a 12 year education, but the empirical evidence is not suited to evaluate this.

6.2.2.2 *Making Causal Inferences from the Reported Evidence*

Given the accumulated empirical evidence, is it possible to infer causal relationships between peer attributes and student outcomes? Towards answering this question, let us consider two important methodological issues that were discussed in Section 3: omitted variables bias and endogenous group membership (or "selection bias"). First, statistical models should be completely specified to include all relevant independent variables as regressors. If omitted independent variables are correlated with peer group measures and outcomes, then our estimates of peer group effects are biased. Second, more attention must be paid to the process by which students are "selected into" different schools. For these reasons, we would, for example, resist interpreting all of the school level peer effects identified in the correlational New Zealand study of Wylie et. al. (1999) as true peer effects.

We can hypothesize several instances in which omitted variables could bias the previous estimates of peer effects. Recall that some researchers have suggested that "peer" effects actually reflect the different instruction received by higher-ability students, rather than student-level interactions. Such classes may be taught at a quicker pace, or their teachers could hold higher expectations for student success, or schools in wealthy neighborhoods with "good" peers could be favored with greater amounts of school resources. If correlational studies do not control for teacher and school characteristics, then we confuse the multiple influences of peers and schools. Table 6.2.1 shows that few empirical studies have extensively controlled for teacher and school variables; even these use relatively broad measures (such as class size or teacher experience) that are unlikely to capture the important variation in classroom practices that could be confounded with measures of peer attributes.

Another potential source of omitted variables bias occurs in studies that relate the proportion of black students to student outcomes. Imagine that one school has 10 percent black students and another has 90 percent. Regression analysis allows us to compare the achievement of students in each of those schools, holding constant a relatively simple set of individual variables, such as parental education. The presumption of many researchers is that students in the school with 10 percent black students will score better, even after individual background

¹¹¹ This has been forwarded as a hypothesis by some authors. For a review, see Moreland and Levine (1992).

is controlled. But it is possible that black families in the two schools are qualitatively different, in ways that not observed by the researcher. Perhaps black families that choose to live in white majority neighborhoods possess certain characteristics such as parental education or family wealth that are associated with higher achievement among their own children. If these characteristics are poorly measured or unmeasured, then the percentage of minority students will reflect the unmeasured influence of family background as well as any peer effects (Moffitt 1998). Bias of this sort probably exists in the Coleman report and in all the subsequent studies of Table 6.2.1.

A related endogenous group membership bias also holds for measures of the effect of group SES. The mere fact that some families consciously choose schools with higher SES peer groups suggests a great deal about those families, beyond that which is recorded in simple indicators of family income and education. It may suggest, for example, that such families are extremely motivated in their schooling decisions—perhaps more so than families of similar income and education in another school with lower SES peers. Thus, highly motivated families tend to be concentrated in high SES schools. As a result, the application of regression analysis, even controlling for basic measures of family background, could mistakenly attribute the effects of “motivation” to measured peer characteristics.

There are statistical methods of correcting for biases induced by endogenous group membership. Many researchers attempt to identify an “instrumental variable” that is correlated with the peer group measure, but uncorrelated with outcomes. The fundamental purpose of doing so is to identify exogenous variation in peer group characteristics across individuals, akin to a “natural experiment.” One can then compare outcomes across schools with higher and lower SES peers, with greater confidence that observed differences truly stem from “peer effects.” In a study of the determinants of teen pregnancy, Evans et al. (1992) correct for endogenous group membership and find estimates of peer effects are substantially altered. Uncorrected estimates appear to overestimate peer effects, perhaps by confounding them with the effects of individuals.

Only two studies in Tables 6.2.1 and 6.2.2 correct for endogenous group membership. As an instrument for school characteristics, Robertson and Symons (1996) use the characteristics of the region in the United Kingdom the student was born in. Unfortunately they do not provide a clear intuitive rationale explaining why the region of birth should be correlated with peer group characteristics, but uncorrelated with outcomes. Gaviria and Raphael (1997) estimate the effect of other students’ drop-out behavior on individual drop-out decisions. They instrument this peer measure with variables that reflect the aggregate characteristics of other students, such as the proportion of each student’s classmates that live in single-parent families. But it seems quite plausible that the chosen instrumental variables themselves independently influence achievement and should be included in the regression for their own sake. In neither case is it obvious that the authors’ empirical strategies have succeeded in eliminating the biases induced by endogenous group membership or that they provide insight into the stated question.

6.2.3 Ability-Grouping Within Schools

Most of the prior research assumes that the entire population of each student’s school constitutes the relevant peer group. In fact, it may be more useful to conceptualize and measure peer group characteristics at the level of classrooms. In the United States and other countries, it is common to “track” or “stream” students into different classrooms by ability-levels, particularly in secondary schools. Students may have substantially greater contact with certain schoolmates by virtue of belonging to a “high” or “low” track. Thus, school-wide

aggregation of each student's peer group characteristics could mask a great deal of within-school variation.

6.2.3.1 *Research Reported by Literature on Tracking*

Towards resolving these difficulties, a large body of research compares the achievement of students in classrooms that are designated as "high-ability," "average-ability," or "low-ability" tracks. A well-known review by Slavin (1987) summarized the results of numerous studies in elementary schools that compare student achievement in ability-grouped classrooms and those that are heterogeneously grouped. In a simple comparison of achievement in grouped classrooms to ungrouped classrooms, he finds a median effect size of 0.00, concluding that "it is surprising to see how unequivocally the research evidence refutes the assertion that ability-grouped class assignment can increase student achievement in elementary schools" (p. 307). A later review of similar research on secondary schools yielded a similar conclusion (Slavin 1990). Again, Slavin computes the median effect size in studies that compare achievement in ability-grouped and ungrouped classrooms. He finds that grouped classrooms have lower achievement, by an average of 0.02 standard deviations.

As Hallinan (1990) emphasizes, Slavin's analyses broadly compare ungrouped classrooms to ability-grouped classrooms, the latter including both high- and low-ability groups. However, if peer effects operate as some hypothesize, then high-ability students should benefit from being isolated with other high-ability students. In contrast, the achievement of low-ability students might decline because they are confined to classrooms with other low-ability students. Considering all grouped classrooms as the "treatment" could disguise an increase in the dispersion of outcomes created by ability-grouping. Slavin's (1990) review of secondary school research addressed this concern by computing, when possible, the median effect sizes of ability-grouping for students that are classified as high, average, and low-achieving. He finds effect sizes of 0.01, -0.08, and -0.02, respectively.¹¹² These results are only minimally suggestive that high-achieving students benefit from being placed in classrooms with high-ability students (relative to heterogeneous classrooms). Likewise, average- and low-achieving students have only slightly lower achievement when confined to classrooms with similar students, rather than a mix of high, average, and low students. Brewer et al. (1995), however, mention several caveats to the interpretation of Slavin's review of secondary research. First, the experimental studies he includes are often based on rather small samples, sometimes taken from a single school. Second, the studies were all conducted prior to 1978, with a single exception. Third, a good number of the secondary school studies (13) were unpublished dissertations, which were not subjected to extensive review.

A different tradition of tracking research has used non-experimental data, often drawn from comparatively recent and nationally representative surveys. Regression analysis is used to relate student outcomes to discrete measures of track placement, holding constant student characteristics such as socioeconomic status. The intent is to isolate the relative effect of being in a high track vs. a heterogeneous track, or a low track vs. a heterogeneous track, holding all else equal. These studies are similar to "peer-effect" studies summarized in the prior section, with two exceptions: (1) measures of peer-group composition are discrete (i.e., high, average, low, heterogeneous) rather than continuous (i.e., average SES) and (2) peer-group composition is always measured at the level of classrooms, rather than schools. Gamoran and Berends (1987) provide a fairly recent review of this literature. Since then, three high-quality studies have been conducted which provide some insights into the relative effects of tracking on student achievement (Gamoran and Mare 1989; Hoffer 1992; Argys, Rees et al. 1996).

¹¹² The effect sizes become slightly larger in magnitude when Slavin computes the effect sizes with only studies that use an experimental or matched experimental design.

Gamoran and Mare (1989) analyze the High School and Beyond survey of high school sophomores in 1980. Students report whether they belong to the college-preparatory or non-college track in their high school. The overall difference in mathematics achievement between the two groups is 0.84 of a standard deviation, favoring the college track.¹¹³ But upon controlling for student background, the achievement difference declines to 0.17. While still substantial, it suggests that most of the achievement difference between college and non-college tracks is due to pre-existing differences among individual students, rather than differences in peer composition.

Hoffer (1992) uses a nationally representative survey of seventh, eighth, and ninth graders in the late 1980s. Science and mathematics achievement of students in “high,” “middle,” “low,” and “non-grouped” tracks is compared. The track placement of students is derived from teacher-reported data. Hoffer finds that the mathematics achievement of eighth-graders in high tracks is 0.36 of a standard deviation higher than that of students in ungrouped classes, holding constant socioeconomic status and other variables.¹¹⁴ But students in middle and low tracks have achievement that is lower than that of ungrouped students, by 0.16 and 0.50 standard deviations, respectively. The results for science achievement are somewhat less conclusive. The only statistically significant difference is between ungrouped and low-track classes; the latter scored 0.41 standard deviations lower, all else equal.

Finally, Argys et al. (1996) compare mathematics achievement across tracks using the National Education Longitudinal Survey (NELS) which follows a nationally representative sample of the 1988 eighth-grade cohort. Using teacher-supplied data, tracks are defined as “above-average” in ability, “average,” “below-average,” or of heterogeneous ability levels. The authors conduct a parallel set of analyses using alternative track definitions: honors, academic, general, and vocational. Their empirical analyses suggest that students in above-average and average classes score somewhat better than students in heterogeneous classes, once background controls are made. Students in below-average tracks score somewhat lower. Considering the different track definitions, students in honors, academic and vocational tracks all score higher than general track students.

6.2.3.2 *Causal Inferences from Tracking Studies*

Taken at face value, Slavin’s reviews indicate few effects of classroom ability-grouping on academic achievement. In contrast, the best evidence from the non-experimental data just described suggests that ability-grouping may widen the gap between the achievement of low- and high-ability students, relative to classrooms where low- and high-ability students are mixed. How can these findings be reconciled, and are they suggestive that student outcomes are indeed altered by the composition of their classroom peer groups? As in the prior section, let us consider the evidence in light of two methodological issues: omitted variables and endogenous group membership.

We have already suggested that the apparent influences of peer groups could be explained by differences among in school or teacher resources that are typically available to classrooms with high and low ability peers. Slavin (1990) believes that many studies of tracking effectively compare apples and oranges, because students in high and low tracks receive vastly different kinds of instruction and curriculum.¹¹⁵ The high-track students—while perhaps of higher ability or higher socioeconomic status—may also be taking advanced algebra, while low-track students are taking remedial mathematics. The difficulty of

¹¹³ Effects sizes are computed by dividing the differences by the standard deviation of the mathematics post-test among all students.

¹¹⁴ Effect sizes are computed by dividing regression coefficients by the standard deviation of eighth-grade tests among ungrouped students.

¹¹⁵ Also see Gallagher (1995).

disentangling the effects of peer-group composition on mathematics achievement from those of curriculum and instruction are quite apparent, even when some effort is made to control for teacher and school characteristics. Towards addressing this problem, Slavin's (1990) meta-analysis explicitly excludes studies in which the type of course taken varies drastically between control and treatment groups. Thus, "conclusions...are limited...to the effects of between-class ability grouping within the same courses, and should not be read as indicating a lack of differential effects of tracking as it affects course selection and course requirements" (p. 487). Of course, Slavin then proceeds to find no substantial effects of ability-grouping, even when effects are divided by the relative ability of individual students. We should also recall that many of the reviewed studies have small samples and are quite old.

In contrast, the analyses of large-scale survey data make varying attempts to control for teacher and school characteristics. Gamoran and Mare (1989) control only for the number of advanced mathematics courses available in the school, while Hoffer (1992) controls for school size and the quantity of science instruction during the year. Argys et al. (1996) use a wider variety of teacher and school controls, including class size and teacher education, certification, and experience. Any of these studies could be omitting important teacher and school measures, thereby biasing the estimated effects of tracking. None, for example, controls explicitly for the types or pacing of courses that students are taking. If courses have more demanding curricula or cover more material, then it is possible that track effects reflect this instead of the effects of peer-group composition. Similarly, unobserved aspects of teachers could bias peer effects. The best teachers may be assigned to the higher tracks, and even controlling for experience or education may not adequately account for teacher quality differences. Thus, the positive effects of attending higher-track classes could be a reflection of teacher quality, rather than peer-group composition.

In addition to bias induced by omitted variables, there is endogenous group membership (or selection bias). Students are assigned to high and low tracks based on any number of characteristics, only some of which are observed by researchers. Schools may use objective criteria such as achievement tests. They could also assign students to tracks based on subjective evaluations of their potential or motivation. Other students may be assigned to tracks because their families exert pressure on school administrators to place the child in a higher track, regardless of ability. In any of these cases, it will prove difficult to compare student achievement across tracks because regression analysis can only control for a limited set of observed student characteristics.

There are some potential remedies to this selection bias. The most obvious is to conduct a randomized experiment, in which students are either ability-grouped or placed in heterogeneous classes based on chance instead of family or school preferences. Slavin's (1990) review considers several of these studies, but they are rather small-scale and mostly conducted in the 1960s. The other studies we discussed use more recent non-experimental data, but apply statistical corrections for selection bias. As with instrumental variables, the usefulness of these corrections depends on the success with which researchers can identify variables that are associated with the propensity of being in a high or low track, but unassociated with student outcomes. Variables include students' self-reported grade point averages (Hoffer), school-average characteristics (Argys et al. and Gamoran and Mare), and regional location (Argys et al.). Selection bias is ameliorated to the extent that variables identifying track assignment are not correlated with achievement. But in each of these cases, it is certainly plausible that the selection variables are directly associated with achievement, and that selection bias remains.

6.2.3.3 Policy Inferences from Tracking Literature

Given the current literature's lack of focus on disentangling the impact of peers in various tracks from the impact of teacher quality/expectations and curriculum differences among

tracks, a careful reader of this literature is necessarily skeptical that reported findings actually represent true peer effects (even if issues of selection bias have been properly dealt with). The question remains, however, whether policy makers can recover useful insights from such findings even if they cannot say whether the findings are due to peer effects, teacher quality or expectations or to curriculum differences. After all, if low SES children are shown to do better in mixed tracks and high SES children do not lose too much from being in such tracks, is that insight not sufficient to cause policy makers to view tracking unfavorably?

Unfortunately, we do not think so. If such measured effects of tracking are due to true peer effects, then the clear policy implication may favor less tracking if policy makers and “society” place more value on the gains achieved by low SES children than they do on the losses experienced by high SES children. However, if these effects result from a variety of possible unmeasured teacher or classroom factors, there is no such implication. More specifically, suppose that those teachers selected to teach in lower track classrooms have lower expectations of children or are otherwise different, and suppose that this difference in teacher quality underlies the poorer performance of low SES children in separate tracks. Then if tracking within a system is reduced, those same lower quality teachers remain in the more mixed classrooms causing de-tracking to have less impact than had the entire effect been a true peer effect.

Still less of an effect, perhaps even a perverse effect, would happen if the cause for the tracking results lies in intended curriculum differences. Suppose, for instance, that policy makers have instituted tracking in order to prepare students in different tracks for very different paths, teaching calculus and other college preparatory material to some and teaching technical or more applied trades to others. If the same test is used to evaluate how well students in different tracks perform, and if this test is primarily geared to subject matters treated more thoroughly in advanced tracks, then the results reported by many tracking studies are almost predetermined and quite uninformative. Furthermore, while it may be true that mixing students will cause lower track children to perform better on those tests, it may well cause them to perform worse on tests that assess skills focused on in lower tracks. The extent to which policy makers and society wish to actually use tracking to teach very different material involves broader philosophical judgements, but the fact that such differences most likely exist across tracks does bias current results as well as policy inferences.

6.2.4 Ability-Grouping Within Classes

Just as students may be assigned to classrooms by ability-level, it is a common practice to group students by ability within classrooms, especially in elementary schools and particularly in New Zealand. In many cases, ability-grouping is designed to create groups that are homogeneous in ability, but since the 1980s, it has become increasingly common for ability-groups to purposely include students of diverse abilities (Lou, Abrami et al. 1996).

6.2.4.1 Reported Research on Within Class Groupings

Slavin (1987) reviews eight studies of the effects of homogeneous ability grouping on mathematics achievement in elementary schools. He finds a median effect size of around 0.3, suggesting that students that were grouped for mathematics instruction tend to perform better than in ungrouped classrooms, regardless of whether groups were high- or low-ability. Slavin also breaks down the effect size by student achievement level, and finds that low-achieving students actually benefit more than others from homogeneous grouping. Because within-class grouping is much less common in secondary schools, Slavin’s (1990) review of secondary school research does not treat this issue in depth.

A more recent meta-analysis by Lou et al. (1996) compares the outcomes of grouped and ungrouped students, but uses a much larger number of studies (51) than Slavin. They calculate a mean effect size on achievement of 0.17, which is statistically different from zero. Many of the reviewed studies also gauge the effects of grouping on student self-concepts and student attitudes (e.g., towards the subject, their peers, and other features of school). In both cases, the mean effect sizes are positive: 0.26 and 0.18, respectively. Lou et al. extend their analysis by examining how the estimated effect sizes on achievement are “moderated” by features of the particular grouping program and participating students. They find that grouping is relatively more effective at increasing mathematics and science achievement (0.20) than reading or language arts (0.13). It is also more effective when applied in elementary grades four to six (0.29) instead of grades one to three (0.08). Finally, the relative effect size of grouping is much larger (0.29) in cases where special instruction is given to grouped students, but standard instruction is given to ungrouped students. In studies where instruction is purposely held constant across treatment and control groups, the mean effect size is only 0.02. Based on this finding, the authors conclude that

it appears that the positive effects of within-class grouping are maximized when the physical placement of students into groups for learning is accompanied by modifications to teaching methods and instructional materials. Merely placing students together is not sufficient for promoting substantive gains in achievement (p. 448).

Whereas Slavin’s analysis was restricted to homogeneous grouping, Lou et al. combine both homogeneous and heterogeneous ability-grouping in a single “treatment” category. This makes it difficult to assess whether the effects of grouping might stem from the ability composition of small group peers, because either type of ability-grouping could yield small groups with vastly different peer compositions. Thus, a further analysis of Lou et al. directly compares the outcomes of homogeneous to heterogeneous ability grouping, finding an overall effect size 0.12, favoring homogeneous grouping. The effects are moderated when calculated separately for low-, medium-, and high-ability students. Low-ability students have lower achievement when confined to homogeneous ability groups (-0.60). In contrast, both medium- and high-ability students have higher achievement (0.51 and 0.09, respectively). The authors consider several explanations for this finding. It could stem directly from different patterns of peer interaction among students of low- and high-abilities:

Low-ability students may gain most in heterogeneous groups from having other students provide them with timely and elaborated assistance and guidance. In contrast, when low-ability students are placed in homogenous groups there may be no student capable of providing those explanations. High-ability students may benefit from being placed in heterogeneous groups to the extent that they are often called upon to provide elaborated explanation by their less able peers. Medium-ability students, however, may act neither as tutor nor tutee and, therefore, neither give nor receive explanations (p. 449).

Alternatively, the achievement of low-ability students in homogeneous ability-groups might suffer if “demands for learning are set too low; if these students feel isolated, inadequate, or incompetent; or if the teacher has negative performance expectations” (p. 450).

6.2.4.2 *Causal Interpretation from Research on Within Classroom Groupings*

What are the causal interpretations of the prior results? Let us once again consider the methodological problems of endogenous group membership and omitted variables. The research summarized in the previous reviews often uses experimental designs, in which students or classrooms are randomly assigned to the “treatment” of ability-grouping. This provides a measure of confidence that endogenous group membership (or selection bias) does

not severely bias results. However, it is interesting to note that the mean effect sizes are slightly higher in the fewer studies that use quasi-experimental design with statistical controls, instead of an experimental design (Lou, Abrami et al. 1996). This is somewhat suggestive that positive selection bias in non-experimental studies tends to tilt results in favor of ability-grouping.

Despite the frequent use of experiments, the familiar problem of omitted variables bias could still affect results. Even though students and classrooms are randomly assigned to treatments and controls, it might be that the same teachers provide different “treatments” to individual students. That is, the experimental design is able to control for the student composition of peer groups and perhaps even the curriculum or method of instruction. But it is unable to ensure that teachers maintain the same level of expectations for both groups, or offer similar degrees of encouragement. It could be, as Lou et al. hypothesize, that lower-ability or heterogeneous groups have teachers with “negative performance expectation” (p. 450). Thus, the effects of peer composition and unobserved teacher behaviors are difficult to separate, and the caveats we discussed in Section 6.2.3.3 regarding policy interpretations continue to apply.

6.2.5 Summary of Research on Three Types of Peer Effects

This section summarizes the results from three categories of “peer-effect” studies, and evaluates their potential to establish causal links between the composition of school peer groups and student outcomes. Among the studies that link measures of school-wide SES to achievement, there are surprisingly inconsistent findings. Authors use an eclectic variety of measures of SES. While some find the expected effects, it is easy to find regression coefficients of the “wrong” sign or examples of a single study with inconsistent results across several measures of peer attributes. When race or ethnicity are included as peer attributes, the results are also inconsistent. Notwithstanding the Coleman report and its substantial influence on popular opinion, correlational studies seem to turn up a mix of positive and negative effects. The evidence from experimental and quasi-experimental studies of desegregation is somewhat more positive. It indicates small but consistent effects of desegregation on black reading achievement, but none on mathematics. White achievement is relatively unaffected.

The evidence on within-school tracking is somewhat more consistent, or at least inconsistencies can be attributed to variation in methodologies. While Slavin’s review of the evidence finds little evidence that tracking affects student outcomes, the accumulated research with non-experimental data consistently finds that students benefit from being in higher tracks, and are disadvantaged by being placed in lower tracks, holding constant their individual backgrounds. An apparent explanation for the conflicting findings is that Slavin removed studies which substantially varied the types of classes that students in high and low groups were taught. Other non-experimental studies did not control for this variable.

Slavin’s (1987) review also finds positive effects of homogeneous ability-grouping within classrooms, but does not find that effects vary substantially by the ability-level of students. This seems to contradict the assumption that students benefit from exposure to higher-ability peers and that, therefore, low achieving students should have even lower achievement when confined to low-ability groups. The more sophisticated and extensive review of Lou et al. finds that the achievement of lower-ability students does indeed decline when they are confined to homogeneous groups of other lower-ability students. In contrast, average and high-ability students have somewhat higher achievement in their homogeneously formed groups.

The causal interpretation of the prior results is far from straightforward, however. In every case there are problems of omitted variables. Few studies can perfectly control for teacher and

school characteristics that differ in relation to peer group composition. For example, it is likely that better-qualified teachers are assigned to schools (or classrooms or small groups) that are relatively better-qualified. Even similarly qualified teachers may have higher expectations for students of higher ability or SES. In either case, the causal interpretation of peer-group effect is clouded. While they could stem from direct student interaction (such as friendships among students, or inter-student assistance in learning groups), they may be an indirect reflection of school or teacher effects. The only way to distinguish between the two is by collecting better data on schools and teachers to use as statistical controls. The extension of this research seems quite important for policy purposes.

Endogenous group membership could also induce bias if the families that tend to choose high-SES or high-ability schools and classrooms also possess unobserved characteristics that improve achievement. We suggested that such bias is probably present in all of the correlational studies with non-experimental data. Several studies have applied statistical corrections for selection bias (Gamoran and Mare 1989; Hoffer 1992; Argys, Rees et al. 1996; Robertson and Symons 1996; Gaviria and Raphael 1997). But all of these face the difficult task of rationalizing their choices of instrumental variables. If the chosen instruments are not strongly correlated with the peer measure and are even slightly correlated with unexplained variation in student outcomes, then bias may still be present. While randomized experiments ameliorate this form of bias, they have mostly been mainly applied in studies of within-classroom ability-grouping. Overall, the evidence provides limited opportunities to make strong causal inferences for true peer effects.

6.2.6 *A New (Untested) Perspective on Peer Effects*

An alternative view of peer effects has arisen from the work of Harris (1995, 1998) and has been hinted at throughout this paper. Harris, similarly frustrated by the behavioral genetics insights on home environments and the resulting lack of an adequate explanation of the environmental component of child outcomes, has developed what she calls group socialization theory. According to this theory, parents have as much influence on their own child as they do on their children's friends. Likewise, their children's friends have as much influence on their child as they do. Culture and norms are passed from groups of parents to groups of children through peer interactions, and little that is taught at home and contradicted by peers lasts in the long run as children attempt to succeed in their peer environment. The relevant peer group, however, is not a census track or even an entire classroom, but rather a small number of children that each child selects as her relevant peer group. (Note that this is represented by the double-arrow in Figure 1 that indicates each child's ability to choose peer groups.)

Harris argues that children can switch personalities dramatically as soon as they are outside the supervision of their parents, and that they do so for peer group reasons. Evidence on small peer groups by psychologists suggests that such smaller peer groups are composed of peers similar to one another and are associated with at least some common outcomes (Cairns, Leung et al. 1995). A peer effect of this kind would explain why studies with data sets commonly used to test for peer effects have often been unsuccessful. Furthermore, it is consistent with many of the findings we report in this review. Finally, Cairns et. al. (1995) suggest that these peer groups are still somewhat fluid in adolescence, and Brown et. al. (1993) find evidence that parental involvement has some impact on child selection of peer groups. This may explain in part the "private" effect of parental involvement in school by parents on their own child's outcome as well as the consistent finding that supervision level in the home has little to do with child outcomes while supervision levels outside the home do (see Section 5.3.1).

Nevertheless, it is premature to argue that this theory is confirmed by empirical evidence. Broadly speaking, it seems to not be contradicted by much of the evidence, but it has not actually been tested (because it is relatively new) and has not yet been subjected to intense scrutiny in the published literature.

6.3 *School Inputs*

The majority of this review addresses the causal influence of families, peers, and neighborhoods on child outcomes. Nevertheless, we have neglected an important aspect of the conceptual model in Figure 1: the causal links between school inputs and student outcomes. A comprehensive review of the role of schools in affecting outcomes is obviously beyond the scope of this paper.¹¹⁶ However, we shall briefly summarize key contributions to the literature, in order to place our broader discussion of non-school resources in its proper context. We review two general categories of studies. The first, and most common, assesses the relationship between school inputs and contemporaneous student outcomes such as academic achievement (often referred to as “production function” or “input-output” studies). The second considers the effects of school inputs on the subsequent labor market outcomes such as wages.

6.3.1 *School Effects on Achievement*

As in studies of families and peers, the overwhelming majority of school effects research uses non-experimental data on schools, teachers, and students. With these data, researchers compare the outcomes of students who have greater and lesser quantities of selected school inputs, while making statistical controls for variables such as family and student background. The purpose of making such controls is to remove the confounding influence of family attributes, thereby allowing us to “isolate” the unique effect of various school inputs. School inputs usually, but not always comprise, a limited range of school and teacher descriptors such as teacher-pupil ratios, expenditures per-pupil, and teacher education and experience.

One of the first examples of such a study is the widely cited Coleman Report (1966) in the United States, already discussed for its contributions to our understanding of family and peer influences. Its controversial conclusion was that school inputs contributed little to academic outcomes. Since then, hundreds of similar analyses have been conducted. Many were secondary analyses of the original Coleman data, but many others collected new data and applied a similar statistical methodology. An important review of the accumulated literature was conducted by Hanushek (1986). He concluded, much as Coleman, that evidence did not point towards a strong or consistent relationship between several kinds of school inputs and student outcomes.

A more recent review by Hanushek (1997) updates and extends his previous work, considering additional studies published up until 1994.¹¹⁷ Its principal findings are summarized in Table 6.3.1. Hanushek focuses on the several common measures of school resources, such as the teacher-pupil ratio and per-pupil expenditures. Column 2 provides the number of estimated effects that exist for each input. Other columns give the percentage of effects that are statistically distinguishable from zero; these are further divided into positive and negative effects. For example, 15 percent of studies found, as we would expect, that the effect of the teacher-pupil ratio on outcomes was positive and statistically different from zero

¹¹⁶ We will not, for instance, give a comprehensive review of the literature on teachers unionization, competition and organizational incentives, etc. but only briefly address the issue of school resources holding these other factors constant. Of course we recognize that these issues may well interact; for example, school inputs may have less impact under unionization or lack of competition.

¹¹⁷ Also see Hanushek (1996).

(at a five percent significance level). In contrast, 13 percent of studies actually found a negative and significant association. The percentages of positive and significant effects are somewhat higher for teacher experience (29) and expenditure per-pupil (27). An infrequently used measure of teacher test-score also yielded more favorable results; of 41 studies, 37% find it to be positively and significantly related to outcomes. In summarizing the evidence, Hanushek's conclusions are much the same as his earlier work: "...there is no strong or consistent relationship between school resources and student performance. In other words, there is little reason to be confident that simply adding more resources to schools as currently constituted will yield performance gains among students." (Hanushek 1997, p. 148). This conclusion is mirrored in a recent careful analysis of New Zealand data using the Christchurch Health and Development Study. In this analysis, Baker and Maloney (1999) find no evidence of class size making a difference in reading abilities of 8 to 13 year olds.

These conclusions are not, however, without controversy. Greenwald, Hedges, and Laine (1996) reach starkly different conclusions in their review of much the same evidence.¹¹⁸ They employ different analytical methods, rejecting a "vote-counting" approach which tallies the number of positive and negative effects. Instead, they use more sophisticated methods of meta-analysis to formally test whether school resources have effects on outcomes. The intuition of their approach can be explained in the following way. Even if a given input has no relation to outcomes, the vagaries of statistical sampling will still lead to a small number of estimated effects that are statistically distinguishable from zero (a two-tailed significance test at 95% suggests that 2.5 percent will be positive and 2.5 percent will be negative). A cursory examination of Table 6.3.1 shows that this is rarely the case. In fact, there appear to be more positive results than chance would produce if the true resource effects were zero. This is borne out by the "combined significance tests" which the authors apply. These results are disputed by Hanushek in a number of ways, who argues that the sample of studies used to carry out the analysis is biased (Hanushek 1997). Because the data requirements of a meta-analysis are more demanding, the authors are forced to eliminate many studies, including those failing to report the sign of statistically insignificant effects. Hanushek argues that this tilts the results in favor of their optimistic findings.

Nevertheless, Hanushek does acknowledge that some schools may be effective at using resources, noting that "...productive results are possible, even if seldom achieved currently" (Hanushek 1997: 152). But where Hedges, Greenwald, and Laine seem to be encouraged by this evidence, Hanushek is more skeptical. He argues that the preponderance of the evidence suggests that simple resource policies will sometimes, though rarely, have the intended effect on school quality. Thus, we are better served by focusing on policies that may alter the incentives facing schools to make efficient use of existing school and classroom resources.

Similar production function studies have been carried out in many developing countries of Latin America, Africa, and Asia. The results of over 100 studies are summarized in a review by Fuller and Clarke (1994). In many respects, these results are more encouraging than studies from developed countries. For example, there are positive findings in 19 of 26 studies which examine the links between textbook availability and outcomes in primary schools. School libraries are found to increase early grade achievement in 16 out of 18 studies. The amount of instructional time was another area in which quite consistent positive effects were uncovered. In contrast, some resources such as class size were found to have few consistent effects, much as in Hanushek's review.

The encouraging findings from developing countries could be interpreted in several ways. First, school effects could be erroneously capturing unmeasured aspects of family background. Fuller and Clarke (1994) observe that many variables used to measure family socioeconomic status (e.g., parental education) are not culturally relevant for some developing

¹¹⁸ Also see Hedges and Greenwald (1996).

societies. Better measures might include ownership of modern possessions such as a television or automobile. They observe that “if imprecise SES indicators from the West are simply imported and error terms contain unmeasured elements of family background that are highly correlated with school quality, achievement effects will be mistakenly attributed to school factors” (Fuller and Clarke 1994: 136).¹¹⁹ Second, school effects may be stronger because of the relatively scarcer resource endowments of schools in developing countries (Fuller and Clarke 1994; Hanushek 1995). If basic inputs like textbooks are subject to diminishing marginal returns, then effects on achievement will be stronger in schools and countries where there are few resources. The marginal effectiveness of inputs will decline in the better-endowed classrooms of developed countries. This is similar to findings that initial increases in family resources matter much more to children than increments beyond a given baseline level.

Hanushek and others frequently use the preceding research to infer a causal relationship (or lack thereof) between school resources and outcomes. Nevertheless, the causal interpretation of non-experimental estimates of school effects is often doubtful. Consider the familiar issue of omitted variables bias. Estimates of school effects are unbiased if adequate controls have been made for intervening variables such as family and student background. If controls are imprecise, then we have little guarantee that school effects do not erroneously reflect unobserved family attributes. Boozer and Rouse (1995) present convincing evidence that the usual estimates of the effects of class size reduction are biased by omitted variables. They note that the allocation of students to large and small classes is non-random. At least in the United States, lower-ability students are typically assigned to smaller classes. But if student “ability” is imperfectly controlled for in production functions, then smaller class sizes will appear to be associated with lower student achievement. This dilemma is best resolved by an experiment in which students are randomly assigned to large and small classes, which provides assurance that differences in achievement are attributable to schools rather than preexisting student differences. Regrettably, few experiments of this sort are conducted, even though they provide a much higher degree of confidence than the hundreds of non-experimental studies reviewed by Hanushek. Krueger concludes that “one well designed experiment should trump a phalanx of poorly controlled, imprecise observational studies based on uncertain statistical specifications” (Krueger 1999, p. 528).

A rare example of just such an experiment is Project STAR in Tennessee (Mosteller 1995; Finn 1998; Krueger 1999). Commissioned by the state legislature, the experiment compared the achievement of students that were assigned to three kinds of classes: (1) students in regularly sized classes (22-25 students) without teacher aides; (2) students in regular classes with teacher aides; and (3) students in small classes (13-17 students). In participating schools, the cohort of students that entered kindergarten in the 1985-86 school year was randomly assigned to each of the three classroom types. Their progress on achievement tests was followed until the third grade. The results from one evaluation suggested that performance on standardized tests increases by about four percentile points during the first year in which students were assigned to smaller classes (Krueger 1999). Effects tended to be relatively larger among minority and lower-income students. Each subsequent year of participation in a smaller class further increased achievement by about one percentile. In contrast, teacher aides were found to have little effect on student achievement.

Other studies, though not true experiments, manage to provide more credible evidence than the typical non-experimental study. In Israel, class sizes are largely governed by a twelfth century rule proposed by Maimonides. The rule dictates a maximum class size of 40. If an additional student enrolls in a school with single class of 40 students, two smaller classes are formed. Angrist and Lavy (1999) argue that this provides a useful source of “natural” variation in class size that is independent of student background. As school enrollments

¹¹⁹ For empirical evidence, see Lockheed et al. (1989).

increase, the authors find a zig-zagging pattern of average class size in which increasing enrollments tend to increase class size to a certain point. Thereafter, new classes are formed and class sizes drop. This pattern is found to mirror movements in test scores. As class sizes rise, test scores tend to decline. Further investigation suggests that disadvantaged students appear to reap the largest benefits from class size reduction, much as in the Tennessee experiment. Finally, recent evidence from South Africa in which the end of apartheid has led to dramatic changes in the distribution of school resources across students indicates substantial impacts of class size on student achievement and test scores (Case and Deaton 1999).

6.3.2 *School Effects on Labor Market Success*

There are fewer studies that examine the links between schools and the eventual labor market success of individuals, usually measured by wages. In one of the most widely cited, Card and Krueger (1992) used U.S. census data from 1980 to estimate the rate of return to education for several birth cohorts of men in a number of states (the rate of return is essentially the added income that individuals receive upon pursuing an additional year of education). The authors then proceeded to estimate the relationship between the rate of return and several measures of schooling quality that varied across states and time, including class size and teacher salaries. Their hypothesis was that higher levels of quality would lead to higher rates of return; that is, an additional year of education would yield greater income gains where schools were of relatively higher quality. In several cases they found statistically significant links between school resources and the estimated return to education.

More recently, the same authors have reviewed the accumulated literature which attempts to link school resources with wages (Card and Krueger 1996; Card and Krueger 1996). In summarizing a number of these studies, the authors conclude that “a 10 percent increase in school spending is associated with a 1 to 2 percent increase in annual earnings for students later in their lives” (Card and Krueger 1996, p. 133). Nevertheless, their conclusions are subject to some controversy. In conducting his own review, Betts (1996) finds that several types of studies are more likely to encounter positive results. First, positive results are common where studies utilize data on school resources that are aggregated to the state level, as in Card and Krueger (1992). When finer data on school resources are used, the effects of school resources are diminished, suggesting that results could be biased upward by using aggregate data. Second, positive results are more common in studies that examine workers who were educated in the first half of the century. This might be due, for example, to the phenomenon of diminishing marginal returns. In earlier periods, schools possessed fewer resources and increasing these may have yielded important effects. As resource endowments improved throughout the century, their marginal effects might have gradually diminished. The implication is that money does indeed “matter,” but only in deprived circumstances.

Recent empirical studies have cast further uncertainty on the hopeful findings of Card and Krueger. For example, Betts (1995) was unable to find any effects of school resources on earnings, using a data set with school-level measures of school resources. Curiously, he does find effects when he substitutes his school-level measures with the state-average measures employed by Card and Krueger. He concludes that state-level variables are imperfect measures of school quality. In another study, James Heckman and his colleagues (1996) are initially able to replicate the findings of Card and Krueger with census data. However, after varying several major assumptions of the analysis, their positive results are reduced or eliminated.

6.3.3 *Summary*

The literature which explores the links between school resources and student achievement is voluminous. Much of it appears to indicate that increasing amounts of school inputs have inconsistent or weak links to achievement. However, the causal interpretation of these results is highly uncertain. The omission of important independent variables could easily bias results of school effects. These methodological shortcomings are best resolved by randomized experiments, but this kind of evidence is scarce. The best experimental evidence suggests that reducing class size does improve student achievement, contrary to the usual non-experimental findings (see Table 6.3.1). There is less evidence on the links between school resources and wages. Despite some encouraging findings, much remains to be learned.

Due to a combination of available data and intuitive appeal, most of the prior evidence focuses on a limited set of school inputs such as class size, teacher experience, and teacher education. Even if such resources are associated with student outcomes or wages, they represent a small subset of what might be considered the relevant school and teacher attributes. For example, the literature has consistently neglected aspects of classroom processes and pedagogy. Hanushek (1997) argues persuasively that greater attention should be focused on gaining an understanding of the role of school incentives and organization, rather than basic resources. Our empirical knowledge in these areas is quite limited (although a complete review is not given here).

From the perspective of this review, the available evidence sheds little light on how families and schools may interact in the production of student outcomes. Most research uses a simple empirical framework to test for the presence of family, peer, neighborhood, or school effects. In the typical additive specification, it is presumed that families affect outcomes, and that schools affect outcomes, but that no interactions occur among the two. However, there is little basis for the assumption that families and schools operate in relative isolation. It seems intuitively reasonable that some school policies or inputs will function better or worse for certain kinds of students, and in certain contexts. Indeed, the evidence cited above suggests that disadvantaged and minority students reap disproportionate benefits from class size reduction, although the roots of this interaction are poorly understood.

7 Causal Links between Community Resources and Child Outcomes

We now examine the final causal force represented in Figure 1: community resources. Community resources can include the financial resources that are spent in a particular neighborhood on schools or other activities for children, but may also include a broader notion of human resources. For example, a community with many positive role models and the expectation that its children will do well in school and life may directly improve children's outcomes and chances for success.

Many theories have been proposed to illuminate how various attributes of neighborhoods and their inhabitants may be related to childhood outcomes. These explanations are grouped into four broad categories by Jencks and Mayer (1990).¹²⁰ The first category of models, "contagion" or "epidemic" models presume that communities have a dominant set of social norms. High levels of poor outcomes may be passed from generation to generation as unemployment, teen-fertility or juvenile delinquency become accepted behavior. Collective socialization models suggest that all the adults in a neighborhood, not just parents, exert an influence on children. Community members may provide passive inspiration through their own success or actively "parent" any child in the community. A third category of models, "institutional" models focus on variations in the treatment and opportunities that are provided to children in different neighborhoods. An otherwise identical child may face different job opportunities or different treatment by authorities if he lives in one neighborhood versus another. Finally, relative deprivation models emphasize the possible effects of basing perceptions about success or failure through comparisons with nearby individuals. A community with many high-achieving children may encourage a child to do at least as well as those around her. Alternately, a child with early disadvantage may lose all motivation if she feels that she will always be unsuccessful when compared to her peers.

There is much anecdotal evidence to support the intuition contained in the above models that shared norms matter to children and that certain norms are more conducive to favorable outcomes than others. Sowell (1994; 1996; 1998) compares the cultural norms and group achievement levels of blacks from the West Indies with American blacks whose roots are in last century's slavery. Within one to two generations, black immigrants from the West Indies, have IQ scores similar or better than the national average. Similar stories can be told of other immigrants to the U.S. Such progress has unfortunately not been replicated for American-born blacks, and Sowell attributes much of this to differences in culture. Additional evidence suggests that children from disadvantaged backgrounds that build ties outside their traditional peer group tend to do better than those that don't, suggesting that these ties grant children access to more favorable norms.

Despite some suggestive findings, studies relying on statistical analysis of non-experimental data do not consistently suggest that higher SES or more affluent neighbors tend to increase the cognitive ability or academic achievement of children. There is slightly stronger evidence that neighborhood effects are important among older children and for attainment as opposed to ability measures. This is consistent with evidence from behavioral genetic studies cited in Section 4 where home environments were found to explain less of the non-genetic variance in outcomes as children reached adolescence and adulthood. Our confidence in these findings, however, is limited for several reasons. Many of the studies rely on a single longitudinal data set (PSID); we have yet to see if these relationships will also be found in data on different populations. Additionally, none of the studies includes measures of school or teacher resources. While these resources are also outputs of community choices and characteristics, their omission could bias estimates of the direct effect of community members on one another. Furthermore, none of the studies reviewed has sufficiently corrected for the selection bias created by families' ability to choose (subject to some constraints) their own community.

¹²⁰ See Gephart (1997) for a further elaboration of conceptual frameworks of neighborhood effects.

While most of the analysis of community effects relies on non-experimental data a few controlled experiments do exist. A 1966 lawsuit against the Chicago Housing Authority (CHA) alleging discrimination in the operation of public housing projects has resulted in a series of experiments in which some minority public housing families are moved to suburban settings. More recently, these experiments have been specifically designed (in “Moving to Opportunity” (MTO)) programs to ascertain effects of the kind we are concerned with. While there is still some selection on the part of parents who choose whether to apply for a housing move, these experiments do afford a comparison between outcomes of chosen families and those who expressed interest but were not chosen. Data from these policy experiments point to stronger effects than those found through earlier analysis of non-experimental data sets, although the strongest effects are for outcomes like juvenile crime and behavior, not student achievement. The fact that experiments yield such results that were not consistently found in non-experimental data is suggestive of a selection effect that the empirical literature had not focused on. This effect would appear if those low SES families that move to higher SES neighborhoods are moving because they feel their children are particularly at-risk (as opposed to those families that are staying behind.)

Given the above caveats, there is a feeling that certain communities are much better than others at encouraging positive childhood outcomes. Even in the absence of the types of effect hypothesized by Jencks and Mayer, communities will differ greatly in the financial resources they provide to schools, parks and other children’s services. An examination of residential choice will enable us to understand the way in which parents can affect the educational resources that are available to their children. In particular, limitations on the communities that a poor family or families of particular ethnic backgrounds can easily choose may create an avenue through which parental socioeconomic characteristics will affect child outcomes. There is theoretical and empirical support for the notion that poor, low SES families will tend to cluster together in neighborhoods with few community resources. Unfortunately little empirical work has been done to directly examine the effect of residential choice on the distribution of community resources and child outcomes.

We now move to the final influence represented in Figure 1: community resources. Communities are broadly conceived to include both geographical groupings such as neighborhoods, as well as other social structures such as extended kinship circles, ethnic communities, and religious groups. While much of the emphasis on community resources relies on evidence of residential community characteristics, the “community” relevant for many households may be more precisely defined by race, ethnicity and culture. To the extent that this is the case, an inquiry into neighborhood and peer effects using only crude measures of neighborhood characteristics will miss many of the true underlying neighborhood and peer externalities. Furthermore, there is strong empirical evidence that within schools, children associate more with peers who share their cultural/ethnic background even if substantial institutional efforts are devoted to attempts at integration.

We distinguish between two avenues through which community characteristics may impact children’s educational outcomes: (1) a direct effect of the community environment (similar to the peer effects discussed in Section 6.2) and (2) an indirect effect operating through community determined resources, such as school inputs and other children’s services. Because the political forces that generate a relationship between community characteristics and school resources are more well-understood and less controversial, much of the theoretical interest and attention focuses on the first avenue. This direct effect of communities is often called a “neighborhood effect.” We will also employ this term, remembering that the effects of culture, norms and peers may operate through many social groupings including, but not limited to, residential neighborhoods.

In this section we begin by describing several causal mechanisms through which neighborhood characteristics might be expected to affect student outcomes. We will follow with anecdotal evidence on the importance of cultural norms. We then present the accumulated empirical evidence—mainly from sociology—which purports to test conjectures about how neighborhood attributes affect outcomes. In particular, we will focus on school outcomes, such as achievement and attainment, even though other outcomes are explored in the literature, such as the probability of early pregnancy or participation in criminal behavior. Following the review of the evidence, we shall evaluate whether observed correlations provide any guidance as to the causal links between neighborhood characteristics and student outcomes. Once we have evaluated the evidence on communities' importance to children and their life outcomes we will shift focus slightly to include a discussion of the impact of parental characteristics on community selection. The selection of communities by individual families provides another, indirect route through which parental characteristics influence children's cognitive ability and educational attainment.

7.1 Mechanisms Linking Neighborhoods to Outcomes

Varied attributes of neighborhoods and their inhabitants may be related to student outcomes, even after controlling for a wide range of individual characteristics. But what are the exact mechanisms through which neighborhood attributes end up affecting outcomes? Jencks and Mayer (1990) cite four categories of explanations (also see Section 6 on school peer groups for a related discussion of this topic).¹²¹

First, “contagion” or “epidemic” models of neighborhood effects presume that neighborhoods have a dominant set of social norms (regarding employment or success in school, for example). Neighborhoods with large percentages of impoverished or unemployed or blue-collar adults may establish norms which “infect” the behavior of children or adolescents who are exposed to them on a daily basis. Some formal models of such phenomena find that only certain norms may be self-sustaining. Nechyba (1999), for instance, demonstrates that self-sustaining norms for acceptance or rejection of out-of-wedlock fertility can typically take two equilibrium values—one with high rejection, the other with tacit approval. Other such models regarding social norms come to similar conclusions, often generating important threshold effects and multiple equilibria that are difficult to estimate empirically. To illustrate the implications of this model, imagine that a community in which teen-fertility is highly rejected faces a small change in the social costs associated with early child-bearing. Even if this leads to a quite small immediate effect on teen-pregnancy rates, the new increased occurrence of teen-pregnancy will lower the stigma and further increase the frequency of teen child-bearing. This feedback effect may continue until a new stable norm is reached of tacit approval and high rates of teen pregnancy.

This force also generates important implications for the effect of increased residential segregation on community norms. There may be many reasons why children from low-income would be slightly less likely to complete high school; family budget constraints may necessitate their early entry into the work force, for example. Nevertheless, in communities composed of families from all income levels, all children are exposed to the same graduation rates and similar norms regarding the desirability of higher education. If income segregation were to increase, however, children from low-income communities would be exposed to an environment with slightly lower rates of high school completion and might begin to change their estimation of the importance of a high school education. As with the teen-fertility example, a small initial difference in outcomes in some communities could then be exacerbated through changes in norms. This suggests that residential segregation in itself can effect lead to a change in norms and outcomes, exacerbating initially small differences across income groups. Policy makers might consider the impact of particular policies (such as those

¹²¹ See Gephart (1997) for a further elaboration of conceptual frameworks of neighborhood effects.

involving schooling) on the residential location choices of families and the subsequent influence on community norms.¹²²

A second class of models, collective socialization models, suggest that neighborhood adults exert influence on children insofar as they are role models for positive behaviors. Jencks and Mayer note that

those who believe in this model...see affluent adults as role models whose existence proves that success is possible if you work hard and behave prudently. They also see affluent adults as potential 'enforcers,' who keep children from running wild on the streets, call the police when trouble occurs, and generally help maintain public order (p. 115).

Thus, it is not so much the presence of "bad" influences that determines outcomes, but the presence of "good" role models who play an important role in socializing and constraining the behavior of children.

A third group, institutional models, shift attention from positive adult role models to the possible effects of other key adults in the neighborhood. Often from outside the community, these adults might include the police or teachers. For example, the police may treat children and adolescents differently in poor neighborhoods, influencing their chances of acquiring a criminal record. Or teachers in schools with a poor student body may treat their classes differently than they would if teaching in a school with a more affluent student body; this may be the effect of negative performance expectations or other stigmas.

The fourth set, relative deprivation models, emphasize the possible negative effects of being exposed to more privileged neighbors. Children and adults may judge their success or failure by comparing themselves with others in their immediate vicinity. For example, a student may judge his success by comparing it with those of neighboring families. When a poor student is around exceptional neighbors, he may judge his situation rather harshly. While some students could respond by competing more vigorously, others may respond negatively, perhaps by dropping out of school.¹²³

These broad categorizations may be of use in conceptualizing the potential forces that are at work in communities and suggesting policy responses to dangerous or "dysfunctional" neighborhoods. Little empirical analysis has been done to actually distinguish between the different theories. They do, however, provide justification and motivation for broad research into neighborhood and cultural peer effects.

7.2 *Evidence on the Effects of Culture and Community Resources*

We now proceed to an evaluation of the empirical evidence on the kinds of neighborhood effects specified above. We do this in two steps: first, we address the evidence on the potential "neighborhood effect" of culture, race and ethnicity, and second, we review the studies searching for more conventional residential neighborhood effects. While there exist many statistically based studies in the latter category, we know of few that focus on the specific issues of race, ethnicity and culture beyond including race or ethnicity in standard

¹²² In settings in which school districting is residentially based, Nechyba (1999; forthcoming) investigates the ways in which school policy might increase or decrease such segregation. His results indicate that reforms eliminating residential zoning for schools (such as those in New Zealand earlier this decade) should result in less residential segregation even if they raise school segregation along income lines. These predictions have not, however, been empirically tested in New Zealand.

¹²³ This would be an example of similar environments having different impacts on different children, a possibility alluded to earlier and covered in more detail in Section 8.

regressions. As we explain below, we think that the “neighborhood effect” aspects of race, ethnicity and culture are not likely to be captured by simple dummy variables, and we therefore do not explore the literature in that direction. We note, however, that we do not generally object to including race or ethnicity into regression specifications, and we think this exercise can be meaningful in exploring other issues. For instance, recent evidence in New Zealand indicates that cognitive outcomes as well as juvenile crime, while initially correlated with ethnicity, are no longer so correlated when other regressors on family background are included (Ferguson, Lloyd et al. 1991; Ferguson, Horwood et al. 1993; Barker and Maloney 1999). While this may not mean that culture as we discuss it below (Section 7.2.1) is causally irrelevant, it does indicate that “being Maori” in New Zealand should not in and of itself be associated with the status of cognitive or other disadvantage. We return to some of this evidence in our discussion of New Zealand in Section 9.

7.2.1 Effects of Non-Neighborhood Community Resources: Culture, Race and Ethnicity

While studies that directly search for evidence that cultural differences lead to differences in population outcomes is rare, some provocative work does exist. There is a large body of work that attempts, through a direct examination of different cultures and their different norms, to distinguish between norms that are conducive to positive outcomes and those that are detrimental. Most of these studies stray far from the methodologies we are reviewing in this paper, and we feel ill equipped to evaluate them. We therefore focus our discussion on the few studies that examine statistical relationships between differences in outcomes and differences in cultural backgrounds across time and populations, even if these studies do not use complex statistical tools to do so. This is one area where we find that simple statistical analysis, carefully thought through, can yield potential insights that can then lead to more general testing in the future.

In an impressive series of works on the impact of culture on race and migrations, Sowell (1994; 1996; 1998) documents the importance of culture through different historical times and into the current decade. Much of his evidence favors the proposition that culture, especially when households sharing a cultural background are geographically concentrated and relatively immobile, may indeed lead to different kinds of equilibria for different sets of households. In other words, isolated populations from common or similar origins may through history or chance develop radically different but persistent sets of norms and attitudes. Some cultural norms are identified as conducive to cognitive development and achievement, while others are connected to economic stagnation and lowered outcomes. Furthermore, cultures are not necessarily tied to race and ethnicity, although these are certainly important contributors to cultural identity. In the U.S., for example, Sowell traces dramatically different paths for blacks from the West Indies whose cognitive development and educational achievements on average mirror those of white Americans, and black Americans who trace their roots to the days of slavery. Not only do these groups differ in their outcomes, but the evidence indicates that, despite sharing race, children of the two different heritages tend not to interact. Norms in favor of achievement are found among children of West Indian background while norms against achievement are documented for many groups of blacks of different background—even within the same schools and neighborhoods.¹²⁴ Similar stories are true of immigrants

¹²⁴ Harris (1998) tells of a school in New York in which the principal reports large tensions between Haitian-born teenagers and American-born adolescents. Haitian-born blacks associate with one another while being shunned by American-born blacks for trying to “act white” when they “are nice and respectful of teachers.” Similarly, she tells of a different school in New York in which children and grandchildren of Jamaican immigrants identify themselves with groups that contrast themselves with other black children in the same school. Both the Jamaican and Haitian children develop norms of achievement and succeed as much as white middle class children, while American-born black children do, on average, considerably worse.

from a variety of different backgrounds. Irish and Italian immigrants in the early part of the century, for instance, scored substantially below white Americans on IQ tests (even when language problems are taken into account) but came with strong norms in support of achievement. Within one to two generations, these groups, just as black immigrants from the West Indies, had IQ scores similar or better than the average. Similar stories can be told of Asian and Jewish immigrants to the U.S.¹²⁵ Such progress has unfortunately not been replicated for American-born blacks, and Sowell attributes much of this to differences in the culture.

While many observers argue that the cause of poor outcomes among American-born African Americans lies in lack of self-esteem, this does not seem to be the case. In the most recent literature review of motivation among African American school children, Graham (1994), for example, finds virtually no evidence in differences in self-esteem or locus of control that could account for differences in achievement. These findings are echoed by Harris (1998) who argues that norms toward scholastic achievement among black child peer groups (of American-born origins) tend to be replaced by norms in favor of other types of achievement (such as athletics), and that this has little to do with self-esteem among children. Such norms, Harris argues, are culturally transmitted from groups of parents to groups of children. Additional evidence suggests that, whenever children from disadvantaged backgrounds build ties outside their traditional peer group, they tend to do better. This is true of the Gautreaux children reviewed in Section 7.2.3, and, to the extent that it has been documented in Section 6, it may be somewhat true for black students under school integration. Similarly, Stanton-Salazar and Dornbusch (1995) find that bilingualism among Mexican students in the US is the best measure of social capital in terms of predicting school success and social mobility. Bilingualism, of course, affords such students access to wider peer groups.

Combined with evidence from such works as Moreland and Levine (1992) that demonstrate the relatively small size of peer groups for children, the work in this section suggests that statistical analysis from large data sets with coarse variables for race, ethnicity and culture will tend to substantially underestimate the role of the peer groups relevant for individual child development. While such studies can, as we have suggested above, be useful in demonstrating that ethnic groups in general are not all that different once background variables are accounted for, they yield little insight into deeper issues of ethnicity and culture more finely defined. The historical work combining simple statistics with broad historical trends suggests to us a possibly strong role for community resources that are not merely related to residential neighborhoods but are further related to peers selected by families and children within neighborhoods and schools. Much of this, however, remains at the level of informed speculation until such hypotheses can be tested more rigorously. Much work remains to be done in this area.

7.2.2 Statistical Evidence on Neighborhood Effects

Despite the theoretical and empirical evidence suggesting that the relevant communities affecting children's outcomes are not easily captured in rough measures of neighborhood ethnic composition and income or SES, most of the current work on "neighborhood effects" relies on just such data. While this may limit our ability to identify the direct effects of communities on children, there may be important policy benefits to studying the way in which differences in community characteristics are broadly related to children's outcomes. Unless we are examining policies that will break the link from community characteristics to school inputs, the indirect effects of community characteristics at work through resource determination are still useful for any measure of the "quality" of a community and how

¹²⁵ This is further evidence against Herrnstein and Murray's (1994) argument that racial differences in IQs are genetically based.

conducive it is to positive outcomes. Nevertheless, much of the interest in the following discussion is in finding the peer effect component of community influence on outcomes. An overview of the general approach taken by these models and some likely limitations precedes our summary of research findings.

As Jencks and Mayer observe, the best method of establishing the influence of neighborhood characteristics on student outcomes would be to conduct a randomized experiment. Two groups of students could be randomly assigned to either a “good” or a “bad” neighborhood. After some specified amount of time, the relative outcomes of the two groups could be compared. Randomization ensures that any differences we observe in outcomes can be causally linked to varying neighborhood environments. For moral reasons, among others, we rarely observe such experiments, and we reserve the discussion of those that exist to Section 7.2.3. As an alternative, regression analysis is used to relate outcomes to neighborhood characteristics. Statistical controls are made for a variety of student and family characteristics, such as parental education and income.

Researchers have employed an eclectic variety of measures of neighborhood characteristics, including general indicators of socioeconomic status (education and income of neighborhood adults), unemployment and joblessness, the percentage of professional workers, and the distribution of race, among others. Several authors attempt to make the case that one measure or another provides a faithful test of a particular neighborhood effects model from the first section. For example, some have suggested that characteristics such as the percentage of high SES adults are a test of collective socialization models (and, conversely, that percentage low SES is a test of “epidemic” models).¹²⁶ We shall remain more agnostic about using the available empirical evidence to distinguish between the competing hypotheses on neighborhood effects. First, the literature has employed such a wide variety of measures that it will prove difficult to extract any consistent patterns of effects for a particular neighborhood measure. Second, most neighborhood variables are highly correlated, which suggests that one measure could well be reflecting the influences of multiple aspects of neighborhoods.¹²⁷ This is especially the case where only one or two measures have been included in regressions, or where several measures have been included in a single index measure.

Although different measures are employed, almost every study includes these measures at similar levels of aggregation. Most studies define neighborhoods geographically, as census tract or similar units, averaging from four to five thousand individuals. At least one study uses zip code areas, which contain roughly 10 to 20 census tracts (Datcher 1982). Another study of urban Boston (which does not focus on educational outcomes) uses smaller groups of city blocks (Case and Katz 1991). There is some evidence indicating that the appropriate unit for measuring neighborhoods is probably somewhat smaller than the census tract, because people do not perceive their “neighborhoods” as encompassing such a wide area.¹²⁸ Nonetheless, the measured differences between census tract and smaller block-groups are not substantial, suggesting that typical empirical strategies will not introduce much error into estimates. Despite the apparent importance of providing a theoretical and empirical rationale for census tract measures, most studies rarely devote much attention to doing so. The motivation for using this level of aggregation often seems to be the constraints of data that are available to researchers.

¹²⁶ See Brooks-Gunn et al. (1993).

¹²⁷ Gephart (1997) cites evidence that poverty and other forms of disadvantage have become increasingly clustered in geographic areas over time, suggesting that multicollinearity among neighborhood measures should also have been increasing over time. For further discussion on the measurement issues in studies of neighborhood effects, see section 3 of this paper, or Jencks and Mayer (1990).

¹²⁸ See Gephart (1997) and the citations therein.

Two prior reviews have been published on the neighborhood effects literature. Jencks and Mayer (1990) reviewed the few studies available at the time that related neighborhood measures to educational outcomes.¹²⁹ Based on these, the authors concluded that “growing up in a high-SES neighborhood raises a teenager’s expected educational attainment, even when the teenager’s own family characteristics are the same” (p. 137). A subsequent review by Gephart (1997) included a much wider variety of studies, given the substantial growth of the neighborhood effects literature during the 1990s. Despite the growth, she is sharply critical of existing studies for their reliance on overly broad and unsubstantiated measures of neighborhood characteristics. She believes that the “results of most existing studies should be viewed as exploratory analyses to see whether even weak indicators of neighborhood and community characteristics can help to clarify whether they influence child and adolescent outcomes” (p. 41). Nonetheless, she does note that several studies have successfully identified correlations between educational outcomes and the presence of middle-class, affluent, and professional-managerial neighbors, even after controlling for individual and family characteristics.

Tables 7.2.1 and 7.2.2 summarize the accumulated research that relates neighborhood characteristics to educational achievement and attainment, respectively. They include the studies reviewed by Jencks and Mayer and Gephart, as well as newer papers from Chase-Lansdale et al. (1997) and Halpern-Felsher et al. (1997), drawn from the two-volume compilation in which Gephart’s review was published. Before discussing the evidence, several points should be mentioned. First, we are not considering all outcomes which may potentially be affected by neighborhoods, such as criminal behavior or teenage pregnancy (see either of the previously mentioned reviews for further details). Second, we shall discuss not only the statistical significance of correlations, but their practical significance. While the former can tell us whether correlations are reliably positive or negative, the latter can tell us whether non-zero correlations are large enough to be of any social or educational importance. To do so, we have sought to identify the effect on student outcomes (expressed in standard deviation units) produced by a one standard deviation increase in a given neighborhood characteristic. Many studies do not provide sufficient information to do so, in which case we merely report positive or negative effects.

We first consider the studies in Table 7.2.1 which assess the determinants of cognitive ability or academic achievement. Chase-Lansdale et al. (1997) use two separate data sets to analyze how the cognitive abilities and achievements of young children are related to a common set of neighborhood measures. They use five indicators: (1) low SES; (2) high SES; (3) male joblessness; (4) ethnic diversity; and (5) family concentration. All are indexes constructed from a variety of underlying variables.¹³⁰ Of these, only high-SES (or the concentration of affluent individuals in a census tract) seems to have consistent links with outcomes (the IQs of 3-6 year-olds and the reading and mathematics scores of 5-6 year-olds). An earlier study with the one of the same data sets appears to uncover a similar link between relatively higher-income neighbors and child IQ (Brooks-Gunn, Duncan et al. 1993).

Another study by Halpern-Felsher et al. (1997), published in the same volume, uses a wider variety of data sets which include children and adolescents at different stages of development. Here the results are quite consistent: in the elementary grades there is almost no evidence of statistically significant neighborhood effects. Approaching middle and high school there is somewhat stronger evidence. For example, in a sample from upstate New York, a general index of “neighborhood risk” is positively associated with indicators of students’ “educational risk.” However, none of the effect sizes is over 0.13 of a standard deviation, and both samples are drawn from a single school district, which should temper impulses towards generalization.

¹²⁹ These included Datcher (1982) and Crane (1991), both of which we summarize.

¹³⁰ See Duncan and Abers (1997) for further details on the construction and data sources of these measures.

Using a fairly small sample of early-elementary students in Baltimore, Entwisle et al. (1994) do not find a statistically significant effect of median neighborhood income on mean mathematics achievement (although they do find that income moderates the school-level gender gap in achievement). Finally, both Dornbusch et al. (1991) and Garner and Raudenbush (1991) find that individual outcomes tend to increase when neighborhood characteristics are more favorable. The first authors use a fairly unique sample of Northern California high school students and self-reported grades. The second authors analyze secondary school achievement among a sample of Scottish students.

The studies in Table 7.2.1 yield a mixed bag of results. There are many statistically insignificant effects, and some significant ones. The only pattern that might be evident is a tendency towards more significant neighborhood effects among higher-grade students. But this conclusion is preliminary, given the generally small-scale and non-representative samples that have been used in most of the research up until now. It is, however, consistent with evidence from earlier sections of this review suggesting that family influence declines in importance relative to other environmental impacts as a child grows into adulthood.

Table 7.2.2 includes studies that analyze the determinants of educational attainment (i.e., the years of completed schooling and the propensity of individual students to drop out of high school). Invariably this research is focused on adolescents and young adults. Of the included studies, four use a single dataset: the Panel Study of Income Dynamics, an ongoing longitudinal survey in the state of Michigan. Using the PSID, Halpern-Felsher (1997) find that a general index of “neighborhood risk” is negatively associated with the years of completed schooling, a finding which is consistent across gender and race. On average, a one standard deviation increase in the risk index tends to lower the years of completed schooling between 0.08 and 0.16 of a standard deviation.

The tenor of their results is largely consistent with those of other analyses of PSID, although comparisons across studies are complicated by the use of different sub-samples, neighborhood measures, and analytical methods. Duncan (1994) finds that the percentage of families with higher incomes in a census tract is positively associated with years of schooling (with somewhat stronger effect sizes than Halpern-Felsher). But curiously, the percentage of families with lower incomes is also positively associated in the sub-samples of white males and black females. Using a limited dependent variable model, results in Brooks-Gunn et al. (1993) suggest that the probability of dropping out of high school declines by around five percentage points if the percentage of high income families rises by one standard deviation. Finally, Datcher’s results, also using PSID, find that average neighborhood income is positively associated with years of schooling among whites. The relationship among blacks is not statistically significant. Rather than census tract measures, Datcher uses larger zip code areas. An oft-cited study by Crane (1991) uses a different and perhaps more representative sample to show that the percentage of professional or managerial workers in a neighborhood is negatively associated with the probability of dropping out across different racial and ethnic groups. Ensminger (1996) finds that a similar neighborhood measure (the percentage of white-collar workers) is positively associated with educational attainment in a smaller and less representative sample of Chicago teenagers.

Cutler and Glaeser (1997) use a novel technique to study the effects of segregation on black and white student outcomes. To circumvent the potential endogeneity of residence in highly segregated neighborhood, the authors compare the effect on individual outcomes of living in a MSA with more or less segregated neighborhoods. While residential choice within MSAs may still generate endogeneity, households are likely to be less mobile across MSAs than they are across neighborhoods within an MSA. Across several specifications Cutler and Glaeser generally find strong negative effects on black student outcomes of growing up in a highly segregated MSA. A one standard deviation increase in segregation would reduce earnings of

25-30 year-old blacks by 7 percent. Averaging across the different outcomes, a one standard deviation increase in segregation leads to an increase of approximately 10-15 percentage points in the probability of an African American child having an adverse outcome: dropping out of high school, idleness or single motherhood.

A study by Borjas (1992) uses a quite different approach. It defines “neighborhoods” as ethnic groupings rather than geographic groupings of residences. Borjas hypothesizes that individuals’ educational attainments are influenced by family characteristics as well as by the average attainment of their respective ethnic groups. He finds positive associations between own-educational attainment and the overall ethnic group attainment, suggesting important effects of ethnic “neighborhoods.” Given the extensive levels of ethnic residential segregation, it is reasonable to presume that measuring educational attainment by ethnic group or by geographic census tract might yield fairly similar neighborhood measures.

In general, the evidence on attainment is much more consistent than the evidence on achievement and cognitive ability. Among secondary students and teenagers, more “affluent” neighbors tend to increase an individual’s average number of years of completed schooling and reduce his probability of dropping out of high school, even after controlling for individual background. One caveat to such an interpretation is that many studies rely on a single data set (PSID); thus, it would be desirable to replicate findings with recent and larger data sets that are nationally representative.

Given the evidence on correlations between neighborhood characteristics and student outcomes, is it possible to infer causal links between these two sets of variables? Towards assessing this, we must consider two methodological issues, already given a general discussion in Section 3: omitted variables and endogenous group membership. In any empirical analysis with non-experimental data, there is the omnipresent risk of omitting relevant control variables. If these omitted variables are correlated with the dependent variables (student outcomes, in this case) and other independent variables (such as neighborhood measures), then the estimated effects of neighborhood measures on outcomes are biased.

Let us consider an example in the context of the present analysis. There is a great deal of empirical evidence, beginning with the Coleman report (1966), that U.S. schools in the poorest neighborhoods also tend to have poorer endowments of school resources, such as quality teachers or sufficient textbooks. The roots of this are not germane to the present discussion, but it is important to consider that there is probably extensive collinearity between neighborhood measures and school resource measures. It is also likely that school resources are associated with the outcomes of children and adolescents. This is especially the case for outcomes that occupy an important position in school mission statements, such as mathematics achievement (but also, perhaps, for other outcomes, such as the propensity of committing a crime). Almost without exception, the studies of neighborhood effects that we have discussed neglect to include school and teacher variables in their analyses. Ostensibly this is because they wish to focus the analysis on neighborhoods rather than schools. But the omission of school variables could bias estimates of other effects. If better neighborhoods tend to have better schools, then does a positive coefficient on “neighborhood income” indicate that students are exposed to affluent adults, or that they attend good schools? The answer is not obvious from the available empirical evidence. As we discuss at the beginning of this section, the failure to include school resources is legitimate if we are looking for the whole effect of living in a more affluent community under given current political and financing institutions, but it will complicate our understanding of direct neighborhood peer effects.

A second methodological issue is related to endogenous group membership (or selection bias). As empirical researchers, we are not privy to the calculus of each family’s decision to

choose a given neighborhood (although there is some evidence, presented in section 7.3, on the factors that influence the residential choices of families.) It is possible that families which do elect to live in affluent neighborhoods are relatively more concerned about educational resources than other families—even those that are seem equivalent to the researcher. Perhaps some families have a greater preference for educational outcomes and thus choose “good” neighborhoods and schools that match their preferences. Of course, we are unable to perfectly control for unobserved characteristics of families such as motivation. Therefore, we run the risk of confounding the effects of “good” neighborhoods with unobserved family influences on educational outcomes. While this selection effect will tend to bias our estimates of neighborhood effects upward, there is a second selection effect that might also be at work and that has not been investigated as much. In particular, suppose low SES parents who observe their children to be in particular danger in low SES neighborhoods are the ones who move to higher SES communities. Then the children of such parents would, all else equal, do worse than expected by the researcher, and the neighborhood effect on that child would be underestimated.

Evans et al. (1992) highlight the perils of using simple regressions that control for a limited vector of observed family characteristics. When they make additional statistical corrections for endogenous group membership, using the technique of instrumental variables, they find that estimated neighborhood effects shrink substantially. In fact, there is widespread recognition among researchers that endogenous group membership could pose serious threats to the estimation of neighborhood effects (Tienda 1991; Duncan, Connell et al. 1997). Unfortunately, only one of the studies listed in Tables 7.2.1 and 7.2.2 attempts to correct for biases (Cutler and Glaeser 1997).¹³¹ There is, however, a good reason for this. One method of correcting for endogenous group membership—instrumental variables—requires the identification of variables that are strongly correlated with the endogenous neighborhood measure, but uncorrelated with student outcomes. For obvious reasons, this is a difficult task.

7.2.3 *Controlled Experiments*

While much of the analysis of peer and neighborhood effects is still in its infancy due to the lack of sufficiently controlled experiments, a few such experiments have (sometimes unintentionally) been conducted in the US. In a well-known lawsuit against the Chicago Housing Authority (CHA) in 1966, a black public housing tenant by the name of Dorothy Gautreaux alleged discrimination by the CHA in both the selection of (primarily minority) communities in which public housing was built and in its disproportionate placement of minority tenants within minority projects. The courts agreed. As a result, a series of “Gautreaux experiments” in which some minority public housing families are moved to suburban settings have taken shape, first in Chicago and then in other cities.¹³² In at least some of these experiments, interested families in large housing projects apply to participate, and some are selected (sometimes) randomly from the list of applicants. While this is not a completely controlled experiment in that families are chosen from a self-selected pool of applicants, it does afford researchers at least the opportunity to study how outcomes of those who were chosen compare to those who had expressed interest but were not chosen.

¹³¹ Besides Evans et al. (1992), the paper by Case and Katz (1991) also uses instrumental variables to make corrections. However, they do not directly consider school-related outcomes as we are doing here.

¹³² Additional court cases have led to such programs in Cincinnati, Dallas, Hartford, Memphis, Omaha, Parma, and Yonkers.

Since most of these programs were not launched until this decade,¹³³ much of the serious evaluation of their long run impacts is only now beginning. In particular, the most recent such experiments resulted from 1992 legislation and became known as “Moving to Opportunity” (MTO) experiments.¹³⁴ Results are encouraging and point to stronger effects than those picked up in the large data sets reviewed in Section 7.2.2. Those in the original program (i.e. prior to MTO) who moved to suburbs were more likely to find work, and their children were more likely to complete high school, attend college, and postpone fertility (Popkin, Rosenbaum et al. 1993; Rosenbaum, Fishman et al. 1993; Ladd and Ludwig 1997). In one of the most recent studies of MTO to date, Ludwig et al. (1998) investigate the impact on juvenile crime of one controlled experiment specifically designed to test for neighborhood effects. Their data allows them to not only test for neighborhood effects on juvenile crime using experimental data, but it also permits an investigation of how biased estimates of neighborhood effects would be using non-experimental (IV) difference-in-difference estimates (controlling for fixed effects). Their results are extremely provocative: in every case that they investigate, the non-experimental estimates are biased in the direction of understating neighborhood effects. For instance, “when comparing the prevalence of arrest for violent crime by teens in low versus high-poverty neighborhoods, the experimental estimates imply a reduction of 23 percentage points while the non-experimental estimates imply an increase of around five points.” The effects were largest for teenagers between the ages of 13 and 17, and they persisted when neighborhood institutions were controlled for.¹³⁵

One implication of this study is that the selection bias that researchers using non-experimental data worry about is in the opposite direction of what had originally been suspected. Recall that one obvious selection bias when investigating the impact of high SES neighbors on low SES children arises if the fact that a low SES parent chooses to reside in a high SES neighborhood tells us something positive about that parent. It may be, for instance, that the parent is particularly motivated to find good public schools for her child, and the neighborhood effect in part includes the parental impact of such motivation. A second possible selection bias, however, would point in the opposite direction. If low SES parents move to higher SES neighborhoods only when they see particular problems emerging in their children, then the children observed in high SES neighborhoods are particularly at risk. The existence of both selection biases may account for the inconsistent results of the empirical literature, especially since the second bias has rarely been considered. The results from the MTO experiments suggest that this second selection bias might be larger than the first, although this remains speculation until further evidence can be gathered.

While studies of this kind are rare, they do support the priors of many researchers that something large may be lost in non-experimental empirical studies of the kinds previously reviewed. Given the fact that these experiments are just now coming into the stage of being studied by researchers, more such studies are about to be conducted. Preliminary indications from some of these are that the results cited above are consistent across MTO experiments. A recent analysis of the MTO program in Boston, for instance, confirms significant reductions in behavior problems for experimental group children there (Katz, Kling et al. 1999). At the same time, even if these results continue to be replicated, their implications for policy are not entirely clear. The Clinton Administration, for instance, has pursued housing policies aimed at housing poor families increasingly in suburbs. The results from the Gautreaux /MTO experiments, however, are not directly applicable since the participants in the experiments

¹³³ Even the first Gautreaux program resulting from the Chicago law suit of 1966 did not start until 10 years later when appeals had been exhausted.

¹³⁴ These MTO experiments were specifically designed to study the Gautreaux model. Legislation was first proposed by the Bush Administration and passed by Congress in 1992, and the program was then launched in Baltimore, Boston, Chicago, Los Angeles and New York in 1994.

¹³⁵ The authors also noted conclusions from ongoing research suggesting that similar dramatic effects for education may not be as easy to come by despite the fact that experimental group children are attending better schools.

were a self-selected group of applicants. The tentative conclusion arising from these experiments is that, among households interested in such programs, neighborhood effects are large—those who were interested but left behind did worse than those who were interested and moved. Whether such effects are similar for households living in dysfunctional neighborhoods but uninterested in leaving remains to be seen. Furthermore, it is not clear whether there is a critical mass beyond which these effects diminish. If, as we suggest below, neighborhood effects and peer effects involve small groups, then moving entire apartment buildings of public housing into suburbs may not have the same effect as moving individual households by themselves.

7.2.4 *Summary*

Despite some suggestive findings, the correlational evidence from large statistical analysis (Table 7.2.1) does not consistently suggest that higher SES or more affluent neighbors tend to increase the cognitive ability or academic achievement of children. There is slightly stronger evidence that neighborhood effects are important among secondary students and adolescents, although one would hope that such a conclusion could be substantiated with larger-scale and more representative data. On attainment, the evidence is somewhat clearer. In general, neighborhood measures have the predicted association with years of completed schooling or drop-out probabilities. That is, students in better neighborhoods are more likely to stay in school, all else equal.

Despite these findings, there are good reasons to be skeptical about their causal significance. First, none of the studies includes measures of school or teacher resources, which could bias the estimates of neighborhood effects (perhaps upward). Second, none of the studies has managed to apply corrections for endogenous group membership, which is widely acknowledged to be a potentially large source of bias. As we have suggested, however, this latter bias could be either upward or downward. When evidence from randomized experiments is analyzed, significantly larger effects have been found at least in some studies and at least for juvenile crime and behavior. Furthermore, evidence from this study indicates that statistical techniques would not have documented such effects, indicating that a bias previously thought to be upward in non-experimental studies may in fact be downward. At the same time, it remains unclear how much such evidence can be generalized for policy purposes.

7.3 *Residential Choice: How do Families Choose Communities?*

Even with the reservations we may have with respect to findings on neighborhood effects, if certain communities are better places for children to grow up than others (a statement that becomes far less controversial once we consider the correlation between school resources and community characteristics) community choice may be another important causal link between parental characteristics and children's outcomes. An understanding of the impact of parental characteristics on residential choice will enable us to understand more fully and to distinguish between the different avenues through which parents affect children's attainment and cognitive ability. Additionally, the constraints on residential choice and mobility may illuminate the limitations currently faced by children from different family backgrounds. In the New Zealand context, we want to know if the fact that people of different ethnic backgrounds tend to live in different geographic communities (Ongley, Carson et al. 1998), affects in any way the resources that are available to their children.

Through the selection of neighborhoods, parents can select both the quality of their children's school and, to some extent, the quality of their peer group. Community choice can be seen as

an indirect family investment in children's educational outcomes.¹³⁶ This investment has real costs, such as higher housing prices associated with "higher quality" residences,¹³⁷ or the sacrifice of other community characteristics that are also valued by the household (such as proximity to parents' employment or a bustling urban environment). The ultimate effect of this investment on outcomes will come from the strength of neighborhood effects, the importance of school peer groups and of community-determined schooling inputs. In this section we will discuss what little is known about the effect of demographic factors on residential choice, primarily in the U.S. context and as is relevant to educational outcomes.

Ideally we would like to determine the effect of family characteristics on the level of community educational resources chosen through the location decision. In this way we would be able to separate direct family inputs into child achievement (say through help with homework or by teaching a child to read) from the indirect effect of family characteristics on the community inputs available to a child. We would also be better equipped to identify neighborhood effects by controlling for the selection of communities by households. Unfortunately little empirical work has been done in this field. There are, however, recent theoretical pieces that model residential choice in combination with endogenously determined community characteristics. These models generally imply the stratification of communities by income and schooling or peer group quality. They also imply that, if residential zoning for school districts is abandoned (as has been done in New Zealand), residential stratification will decline but school stratification will increase. Some preliminary analysis of New Zealand reforms suggests that at least the latter of these effects may have taken place in New Zealand. We further discuss the potential that racial segregation across communities may create differences in households' probability of living in communities with "high quality" peer groups. Finally we present the few recent studies of household community choice specifically related to school quality.

7.3.1 Residential Segregation by Income

Several theoretical models of residential choice and educational input determination predict income stratification with higher quality schools/peer groups in higher income communities. In one group of models this stratification is based on the determination of local voting for community school financing. Epple and Sieg (1999)¹³⁸ develop and test equilibrium implications of a model in which households chose their community, and public services are

¹³⁶ While funding is centrally determined in New Zealand, there is still variation in the financial resources available to schools. In addition to differences in state support, schools in different communities also receive different levels of private contributions (Fiske and Ladd 1999).

¹³⁷ Research on the impact of local school quality on housing prices, has found broad evidence that households pay more to live in communities with higher quality schools. In a recent review of the literature, Crone (1998) separates these studies into three categories based on the measure of school quality they employ: direct surveys of parental estimation of school quality, school expenditures, and educational outcome measures such as standardized test measures. In all three categories of studies, a correlation is found between levels of the school quality measure and housing prices after controlling for other characteristics of the home and community.

Recent work by Bogart and Cromwell (1997) and Black (1998) gives a sense of the magnitude of the investment made in children's education through housing payments. Bogart and Cromwell estimate the response of housing prices to different school quality (and district tax) levels. They study houses in Cuyahago County, Ohio, with similar physical attributes, located in the same municipality, but on different sides of school district boundaries. They find housing price differentials as large as \$18,000 (1987 U.S. dollars). Black (1998) uses homes that are close to intra-district school attendance boundaries in suburban Boston to assign a value of \$4,000 to a 5% increase in elementary school scores.

¹³⁸ The model builds on work in Epple, Filimon and Romer (1984), Epple, Filimon and Romer (1993) and Epple and Romer (1991).

financed through a local property tax. The model implies that even when households have the same underlying taste for educational attainment, communities will be stratified by income and the level of local public good.¹³⁹ Another set of models use peer effects, rather than local finance determination, to explain why we may see income and quality stratification. Benabou (1993) models the interaction of residential choice, human capital investment decisions and community production in the presence of peer effects. Highly skilled workers, who will eventually have higher incomes, price low-skill workers out of their communities in order to lower their own costs of education. Benabou finds that equilibria exist in which there is at least partial segregation by skill type. Benabou (1994) expands on this model to include explicit modeling of families, local school financing decisions, capital market imperfections and direct intergenerational wealth transfers. The model implies similar results as Benabou (1993) with the further implication that residential segregation will lead to the persistence of income and educational inequality across generations.

At first glance, this theoretical work may seem largely irrelevant for the New Zealand context where funding for schools is centralized. While such differences do need to be recognized, we think this work is still informative. Nechyba (1999), for instance, demonstrates that residential stratification along income lines is not inconsistent with central school financing, and such stratification is more likely if peer effects are strong. Furthermore, although New Zealand schools are funded centrally, local spending is affected by parental contributions and local school fundraising (including the increasing practice of finding foreign students who pay tuition in excess of marginal cost) (Fiske and Ladd 1999). Such local funding differences would further contribute to the types of segregation analyzed in these models. In fact, the main difference between the U.S. and New Zealand that might make these models less relevant is that New Zealand has abandoned district zoning and opened the public schools to competition while this has largely not been the case in the U.S. While residential location is still linked to school choice in that parents will want to live relatively close to the schools they choose, parents in New Zealand no longer have to live in a particular geographic district assigned to the school of their choice. To that extent, recent work suggests that residential segregation should decline as a result of the lifting of district zoning while school segregation by income should increase (Nechyba forthcoming). Preliminary analysis of New Zealand's education reforms seem to suggest that the latter (increasing school segregation) may in fact have happened (Fiske and Ladd 1999), while the former prediction (declining residential stratification) remains to be tested.

7.3.2 Residential Segregation by Race

Residential segregation by race could also affect the level of "investment" in community inputs by different ethnic groups. The literature on residential segregation by race suggests that racial segregation is fairly pervasive and persistent.¹⁴⁰ In Borjas (1995), the author establishes strong patterns of ethnic and immigrant residential clustering using 1970 US Census data and data from 1979 reported in the National Longitudinal Surveys of Youth (NLSY). These patterns are exhibited across educational levels and are present even amongst recent movers. Borjas (1998) finds a strong empirical correlation in segregation across generations. He also finds that more highly educated members of low skill groups move to less segregated neighborhoods, while higher education levels for members of high skill

¹³⁹ Fernandez and Rogerson (1996) present a similar model in which high income communities select higher taxes or larger houses with higher tax payments and have higher levels of schooling inputs. Nechyba (Nechyba 1996; 1997; Nechyba 1999), however, demonstrates that even under equal taxation and school funding, such stratification can arise from the existence of zoning or a pre-existing housing stock.

¹⁴⁰ See Massey and Denton (1987) and McKinney and Schnare (1989)

groups has an insignificant effect on the probability that they will live in a segregated neighborhood.

Continued residential segregation may be generated by preferences of households to live near members of the same ethnic group.¹⁴¹ Residential segregation may also be generated by the combination of similar preferences within ethnic groups over community types and by restricted residential choices. In either case, segregation implies that households of different ethnic groups, in a sense, select from different choice sets in their location decision. When household characteristics associated with positive peer effects, such as education and income, are differentially distributed across ethnic groups, segregation will tend to lower the ability or propensity of members of the disadvantaged group to select themselves into high quality communities. In other words, families who belong to a group that has, on average, lower educational and income levels, will face a trade-off between high levels of neighborhood characteristics positively associated with good student outcomes on the one hand, and sharing a residential community with other members of their same cultural and ethnic background on the other. Members of the advantaged cultural/ethnic group are not forced to make the same trade-off.

7.3.3 *Residential Peer Groups and Education*

There are a few recent studies that attempt to directly measure the effect of community educational peer group and school quality on the location decision. In Nechyba and Strauss (1998), the authors build on a model¹⁴² of simultaneous housing and community choice to estimate of the influence of community characteristics on housing choice. Unfortunately the estimation does not treat either the endogeneity of housing prices or that of neighborhood characteristics. Nechyba and Strauss find that a 1% increase in expenditures on local schools can increase the probability that any given household will locate in a community by 1.65% to 3.06%. Epple and Sieg (1999) test the income distribution predictions of a model of community choice with endogenous public good provision. Like Nechyba and Strauss, they use expenditures as a proxy for school quality. The author's decompose variation in household income within and across households and interpret the high level of income variation within communities as evidence of high levels of unobserved preference heterogeneity for local public goods.

Bayer (1999) contains the only treatment of the simultaneous determination of public school quality and the distribution of families across communities. In preliminary results, Bayer (1999) finds that African American and Hispanic households tend to locate in lower quality school districts. He is able to examine the separate impacts on this result of family income, work location, differences in tastes and the limited choices available in the housing market. He finds that part of this difference in school quality across ethnic groups is driven by differences in group income levels and in work location. African American and Hispanic households tend to chose lower quality school districts in part because they tend to have less to spend on a home and they tend to work in communities near low quality districts. Bayer finds that only half of the remaining portion is attributable to group differences in tastes for school quality. The other half is attributable to differences in preferences for community and housing characteristics that are typically linked to community school quality. Because families are limited by the available combinations of housing and community characteristics present in the market (neighborhoods with good schools also tend to have large high quality homes), preferences over one quality, such as housing size may affect the equilibrium consumption of school quality.

¹⁴¹ Clark (1991) summarizes the literature on racial preferences in determining community composition.

¹⁴² Also see, Nechyba (1997)

To clarify the forces at work, imagine that houses are either small or large and that small houses are all located in bad school districts while large houses are all in located in good districts. A family that wants both a large house and a good school simply buys the large, expensive home. A family that would like to locate in a good school district just as much as the first family, but that does not want a large home must, in a sense, pay for something it does not want just for the opportunity to buy the item it does want. There is a chance that this family will live in the poor quality district even though it would have been willing to pay for a higher quality school if it had had the option to purchase a smaller home in the better district.

Bayer's estimation suggests that low demand for housing size and other characteristics leads to lower consumption of high schooling quality since small homes tend to be located in low quality districts. The finding that African American households have lower housing demand agrees with findings in the literature.¹⁴³

7.3.4 *Conclusion*

Although direct studies are limited, theoretical and empirical work would suggest that endogenous residential choice leads to lower levels of community educational (not necessarily monetary) inputs by low income households and by ethnic minorities. The full effect on children's attainment of residential location patterns will, of course, depend on the extent to which neighborhood peer groups and local variations in school quality matter for educational outcomes. In the context of New Zealand, it will further depend on the extent to which the competitive choice system introduced over the last decade works and to what extent low SES children actually gain access to good schools outside their neighborhoods. Furthermore, it depends on the extent to which the promise of competitive gains from increased competition is realized. These are issues beyond the scope of our review.

¹⁴³ See Rapaport (1997)

8 Research Implications for New Zealand

In this section we concentrate on implications of our findings for New Zealand researchers. Overall, we think the results from the international literature are informative for New Zealand in great part because, although we recognize the differences, the evidence suggests that many of the problems faced in New Zealand are mirrored in other contexts and are likely to have similar causal explanations. For this reason, we suggest that the applicability of results obtained by social scientists should not be judged so much by the particular context from which the data are drawn as the quality with which the analysis is conducted. The research findings in the international literature therefore enable New Zealand researchers to focus on problems and issues particularly unique to New Zealand, issues involving the Maori and Pacific Island cultures, the innovative school reforms of the past decade, and the policy of reviving the Maori language and culture, but to do so with the lessons of the international literature in mind. This should enable researchers to avoid common methodological missteps that lead to inaccurate policy conclusions, and to interpret existing New Zealand evidence correctly.

Throughout much of this review, almost all of our focus has been on the international and often the U.S. literature. As mentioned before, this was done not by design and was primarily dictated by the fact that the bulk of the literature in this area has used largely U.S. and European data. The natural question that arises, then, is how this research can be of most benefit to researchers. In various places in the review, we have made an effort to indicate when recent studies using New Zealand data were either consistent or inconsistent with the international literature. Broadly speaking, we find few inconsistencies and suspect that many of the results from the international literature would be similar if similar analysis were replicated in New Zealand. We therefore think that the results from the international literature are informative for New Zealand researchers. At the same time, we recognize certain aspects of New Zealand to be quite unique. These aspects provide a role for researchers to not only learn about New Zealand but also to have the rest of the world learn from New Zealand. An analysis of unique aspects of New Zealand should still, however, be informed by the general lessons learned from the international literature.

Before we begin, we offer a short disclaimer. Clearly our knowledge of New Zealand, its institutions and its peoples remains relatively limited even after spending time in New Zealand and reading the work of New Zealand researchers. Were the primary purpose of this review to assess New Zealand in particular, clearly we would not have been the right team to do so. Rather, our task was to focus on summarizing and evaluating research linking family and community resources to child outcomes in general, and then to do our best to help the Ministry assess how all this best fits with the needs of New Zealand researchers. We do our best in this section to accomplish the latter task, and we hope the broad background offered in the previous seven sections provides sufficient information for those more familiar with New Zealand to correct our mistakes whenever we stumble over our limited knowledge.

Given our critical tone of the methodologies often employed by social scientists, one easy implication for researchers in New Zealand would be to use the available data from New Zealand and conduct properly controlled statistical analysis that avoids the methodological pitfalls outlined in Section 3. While this would indeed be a laudable endeavor, and while data sets like the Christchurch Health and Development Study and the emerging “Competent Children” study provide ample opportunity to do so, it is also unrealistic to think of limited research resources stretching far enough to accomplish this goal. We therefore suggest that one way of approaching the question of possible research implications for New Zealand is to do so in the context of a binding research budget constraint, a constraint under which researchers have to recognize tradeoffs between all the possible ends they could ideally be pursuing. This is especially true in light of the fact that the kinds of more sophisticated analyses we have held up as the better studies in our review are time and resource intensive.

At the same time, knowledge gained from them is more valuable than a hundred simpler studies replicating the same methodological errors.

With this in mind, we suggest that research implications for New Zealand fall into two broad categories. On the one hand, much of the research we have reported focuses on problems that, while studied with data from other countries, are not altogether dissimilar from those faced by New Zealand. Such problems include but are not limited to rising divorce rates, increasing teenage fertility, differences in educational outcomes among ethnic groups, rising inequality, etc. If research budgets are constrained, it would seem that New Zealand research attempting to find causal links between family/community resources and child outcomes ought not to focus on these general common problems. Causal links are likely to be similar for problems that display such similar symptoms, especially if they are replicated with data from different contexts, even if none of them is actually related to New Zealand itself. On the other hand, other problems and challenges are particularly unique to New Zealand, especially in areas where New Zealand is a policy leader in the world (particularly in education) and where New Zealand faces unique policy challenges with respect to certain cultures. Not only are New Zealand researchers who focus on these issues likely to provide unique insights for New Zealand, they are also likely to teach the rest of the world lessons that cannot be learned in other contexts, in particular lessons about policies not in force elsewhere. At the same time, such research can interact with the international literature in important ways. Most particularly, general research findings regarding causal links as well as correlational links that are no longer thought to be causal can aid researchers in New Zealand in properly interpreting their findings as they investigate situations particular to New Zealand.

In Section 8.1 we then speculate on how the international literature gives insights and results that are likely to be quite similar were analogous New Zealand studies to replicate their methodologies using New Zealand data and how this literature can serve to help in the interpretation of New Zealand evidence. Then, in Section 8.2, we suggest some problems and challenges that are particularly unique to New Zealand and that may benefit from research efforts using New Zealand data.

8.1 The Usefulness of International Findings for New Zealand Researchers

While we started in Section 4 with a review of the work by behavioral geneticist that seemed inconsistent with conventional wisdom arising out of the social sciences, we were surprised to find fewer inconsistencies than we expected. A careful examination of the evidence underlying the “conventional wisdom” makes it appear that social scientists are arriving at conclusions that are not altogether incongruent with the results arrived at through the very different methodology used by behavioral geneticists. Not only are we finding results of similar flavor in various areas, we are also finding these with data sets from different contexts, even if these contexts rarely include New Zealand. Many issues have clearly not been sorted out, and much debate remains. Still, it is difficult to come away from this literature with a feeling that different human beings in an increasingly global world are sufficiently different to raise doubts regarding at least the partial applicability of findings in one setting to another. The New Zealand research we have cited tends to confirm this feeling, frequently demonstrating very similar correlations and increasingly supporting the notion that groups within New Zealand are quite related. For this reason, we suggest *that the applicability of results obtained by social scientists should not be judged so much by the particular context from which the data are drawn as the quality with which the analysis is conducted*. Research that meets the methodological standards set by the best studies in the field is hard to come by. To insist that it further make use of New Zealand data before being judged applicable sometimes raises the bar beyond the reachable.

With this in mind, we think that New Zealand researchers can be confident in the general findings from well done international studies, and can use those findings to (1) fine-tune

research targeted specifically to issues particular to New Zealand and (2) help in properly interpreting their own research findings. While we leave the discussion of issues regarding the targeting of research to New Zealand contexts to the next section, we want to emphasize how important the international findings are for properly interpreting both past New Zealand research and new research efforts. An understanding of the likely role that genetics plays, and an understanding of how this might impact the interpretation given to certain regression coefficients in the analysis, for example, seems vital. Similarly, the declining importance in the international literature of such shared home environmental factors as family income should give researchers who (properly) control for family income in their analysis pause in interpreting high regression coefficients on income as causal. Various family background characteristics such as family income, the international literature suggests, serve to a large extent as proxies for other unobserved characteristics that are more likely to be the underlying causal channels. Significant correlations do not, therefore, imply that changing the particular environmental factor that is correlationally significant through policy will necessarily yield large changes in outcomes. Similarly, the statistical problems international researchers have identified when attempting to isolate peer and neighborhood effects should cause researchers everywhere including New Zealand to hesitate when interpreting the correlation of group characteristics with individual outcomes as causal effects. Selection bias, correlated unobservables, omitted variables, and specification errors are all real issues, and recognition of them by social scientists in their work is increasingly driving what we think will be a new emerging consensus.

Therefore, the biggest implication of the international literature for New Zealand is that, as New Zealand researchers focus on unique aspects of New Zealand (discussed in Section 8.2 below), they should not leave behind the lessons learned by social scientists in general. The correlational research done in earlier periods and still conducted today (oftentimes yielding useful results) has its limits: it gives us important pictures of the world, but it should not be over-interpreted, over-generalized and used in a rushed search for “big answers.” To use a metaphor, this correlational research has peeled the first layer of a complex onion often without acknowledging the existence of other layers, only to move on to the next onion to peel the same first layer once again and be pleased about finding that it is not all that different. As it turns out, onions drawn from different settings often have similar first layers, which is important information. That is also the useful message delivered by much of the correlational research already done on New Zealand. Similarly, our own review of broad general trends and statistics in New Zealand, of overall trends and those affecting particular minority groups, suggests correlations that are familiar from other settings. More importantly, some of the most insightful New Zealand research is increasingly confirming that Maori/Pakeha outcome differences decline or disappear once other background variables are controlled for (Ferguson, Lloyd et al. 1991; Ferguson, Horwood et al. 1993; Barker and Maloney 1999). This implies insights from other contexts are probably more applicable than often acknowledged, and the lessons from the international literature can be applied with less hesitation than otherwise feared.

8.2 Unique Aspects of New Zealand Calling for New Zealand Specific Research

While New Zealand researchers can thus learn much about causal links from the international literature, and while this knowledge can be quite helpful in properly interpreting past and future research, some issues are so specific to New Zealand as to make research focusing on these issues of major importance. Thus, recognizing that human beings in different contexts are often quite similar does not obviate the need for country specific research. Such research is particularly important to the extent that unique problems exist and unique policies are in place. In our limited experience with New Zealand, we think that we have identified at least some such unique circumstances, and we suspect there are others we might have missed. We discuss three classes of such circumstances: Maori and Pacific Island cultures, the role of unique school reforms, and current efforts to revive the Maori language and culture.

8.2.1 *Maori and Pacific Island Cultures*

While we have argued that human beings are similar enough to enable us to make inferences from one group to the next, it is also important to recognize that differences might play a role, especially when certain groups seem particularly and consistently disadvantaged. New Zealand, we believe, offers a unique opportunity to study the role of cultures and of “cultures in transition” and how this impacts children in minority groups. At the same time, we think that the concept of culture must be thought through carefully, and that typical distinctions may sometimes be uninformative or may cause researchers to miss important sub-cultural effects. Much of what we have read regarding Maori, for instance, indicates to us that the stark distinction between Maori, Pakeha and Pacific Islanders that is often made by commentators is somewhat artificial. Rather, the degree of intermarriage between these groups suggests a melding of cultures, a melding that, while it is happening in other societies, has its own particular flavor in New Zealand.¹⁴⁴ A continuing study of this evolution may hold important implications for links between family/community resources and child outcomes, but care must be taken not to be so focused on artificial categorizations so as to miss important distinctions within coarsely defined cultural groupings.

In particular, there is a strong sense among certain groups in New Zealand that a great part of the explanation for the relatively poorer performance of Maori and Pacific Island children lies in the fact that they are members of minority cultures lost in institutions shaped primarily by Pakeha needs.¹⁴⁵ As we have already suggested, the most recent research in New Zealand seems to indicate that this point may have been overstated. In analyses of juvenile crime, cognitive ability and reading skills, evidence from the Christchurch Health and Development Study consistently suggests that group differences in child outcomes can be fully explained by variables other than culture (Ferguson, Lloyd et al. 1991; Ferguson, Horwood et al. 1993; Barker and Maloney 1999). Furthermore, some measures of intelligence were entirely uncorrelated with ethnicity even before other factors were taken into account (Ferguson, Lloyd et al. 1991). Nevertheless, these studies are all derived from the same data set, and one of the weaknesses of that data set lies in the relative under-representation of minority children. This point is freely acknowledged by the authors of these studies, and it calls for further analysis of data from the North Island.

At the same time, the exploration of culture as undertaken in current New Zealand studies may miss the important issue of the potential formation of sub-cultures. With New Zealanders of different cultural backgrounds increasingly mixing through intermarriage (Ministry of Maori Development 1998) at the same time as some segments of these populations are segregating into certain residential neighborhoods and schools within urban areas, are we over-generalizing when we speak of “Maori” or of “Pacific Island” children? We suspect, for example, from the evidence discussed in this review that norms and attitudes can often have a

¹⁴⁴ The outside observer cannot help but be struck by the degree to which Maori words, for example, frequently flow from Pakeha lips in New Zealand, with foreign visitors often lost as they are too timid to show their ignorance of these words. Similarly, of course, Pakeha culture has impacted Maori, especially as interactions and mixing has increased with the dramatic urbanization of Maori. Ferguson et al. (1993), for example, report that 83% of those classified Maori and 52% of those classified Pacific Islander in the Christchurch longitudinal study in fact have one parent of Pakeha ethnicity, although these numbers seem to be somewhat smaller for the general population. Still, from 1991 through 1996, those classified in the general population as sole Maori declined from 74% to 52% of the Maori population, while those classified as mixed Maori rose from 26% to 48% of the Maori population. Furthermore, Ferguson et al. (1993) present evidence that a majority of those labeled Maori only report a weak association with Maori culture.

¹⁴⁵ This evidence is documented thoroughly in a number of government reports see for example, (Ministry of Education 1999).

“local” character (i.e., they can differ within ethnic groups along lines other than the usual coarse cultural and racial distinctions). A deeper understanding of the differences between Maori that succeed in the Pakeha-dominated New Zealand and those that do not is required, and this may entail a much finer understanding of relevant “neighborhood” and cultural “peer” groups, not just the traditional focus on family background and broad cultural categorizations. In that sense, current New Zealand studies demonstrating the lack of explanatory power of culture (with respect to student outcomes) may be overstating their case.

Finally, there remains a strong suggestion in the international literature that frequent mobility may be detrimental to child outcomes, whether through peer disruptions or other channels. New Zealand offers a good opportunity to further test whether this is true in general and how much it is impacting New Zealanders in particular. Certain segments of the New Zealand population (including important segments of the Maori population) are significantly more mobile, often changing neighborhoods and schools yearly or even more frequently. Caution, of course, needs to be taken to not interpret correlations as causations. Those segments of the population that are more mobile differ from others in many ways, and it will be difficult to determine how much of the possibly lower outcomes for these populations is due to mobility. However, this remains an important and still unresolved issue in the international literature that may have particular relevance for some of the most pressing issues on the New Zealand policy agenda.

8.2.2 *New Zealand’s Innovative School Reforms*

New Zealand has, in many ways, become a leader in policy reforms, and this is certainly true in the area of education policy. Not only are these reforms unique in many ways, they were also dramatic at the time of their implementation. The combination of decentralization of control and largely centralized funding, the introduction of competition and choice, and the abandonment of residential districting is rather unique in the world. While evaluation of the impact of these reforms is certainly important for New Zealand, we would be dishonest if we did not admit to being excited about the potential international implications of such evaluations as well. In particular, we are intrigued by three distinct research questions arising from these reforms: First, what impact has decentralization of control had on the variety of schools and approaches within schools, and what can be learned from these differences? Second, has the introduction of competition lived up to its promise, and what unintended consequences have flowed from it? And third, has the removal of residential districting had an important impact on the way in which families choose residences, and has this impacted levels of residential segregation that might be important for child outcomes?

The issue of decentralization is, of course, one of increasing global interest, and the combination of centralization of funding and decentralization of control has often been difficult to achieve in other contexts (McKinnon and Nechyba 1997). New Zealand, however, seems to have managed to accomplish this and thus offers an example for study that differs from studies of centralization in school funding within U.S. states where centralization of funding is typically accompanied by centralization of decision making. The hope, of course, is that decentralization will foster innovation as well as responsiveness to local needs, and that it will encourage (in part through institutions specifically designed to do so) parents to become more involved in schools. Has this been accomplished in New Zealand? Is this true across the board, or is it more true for high SES parents and schools than low SES parents and schools? Can lessons be drawn regarding possible modifications to decentralization reforms that might be useful in making these reforms beneficial for larger segments of the population?

When combined with the introduction of competition, decentralization becomes of course even more potent. While the issue in other countries is largely one of introducing competition

through the fostering of private education, much of the flavor of such competition is contained in the New Zealand reforms. Parental choice of schools has been expanded dramatically, and anecdotes abound as to the efforts schools put into actively competing (Fiske and Ladd 1999). Has this intense competition resulted in gains in efficiency as argued by the proponents of choice for decades? The rest of the world often merely speculates on these questions, or attempts to use data from relatively non-competitive settings to infer what competition *might* do (Hoxby 1998; Rouse 1998; McMillan 1999; Hoxby forthcoming). New Zealand offers researchers an opportunity to see what competition *has* done in regard to efficiency. Furthermore, it may give insights into the ways in which schools respond to competition, whether they primarily raise efficiency or, for example, whether they themselves become involved in choosing students and possibly discriminating among them in ways that benefits some at the expense of others. Again, this analysis, properly conducted, is challenging methodologically because of the many confounding effects, but it is exceedingly worth doing.

Finally, the interaction of residential neighborhood choice and school choice may be of great importance if peer effects matter both within schools and within neighborhoods. Theory suggests that residential choice is quite different in systems in which eligibility to attend schools is based on which district one lives in from systems in which there is no link between where one lives and where one attends school. New Zealand has moved from one system to the other. While residential location is a long run issue, and few would expect dramatic changes in residential location patterns over night, the New Zealand reforms have now been in place for some period and are expected to continue. New Zealand therefore offers a clear opportunity to evaluate how the introduction of choice (whether of the kind present in New Zealand or the kind proposed through vouchers in the U.S.) impacts neighborhoods. To the extent that inflated housing prices in good school districts under residential districting capture the value of such schools, we should see, for example, a change in relative property values across districts as districting is ended. This change might lower barriers for residential integration along income lines, and that may be an important unintended consequence of choice based reforms.

8.2.3 *Policies Aimed at Revitalizing the Maori Language and Culture*

Finally, one the most unique New Zealand policies is that involving current efforts to revive a language (and culture) that, by all accounts, had been largely lost or forgotten to many people of Maori decent. Those places where languages have been revived through active government intervention in the past have tended to be places where populations speaking those languages are quite segregated from others. Examples include those often mentioned by New Zealanders: Basques and Israelis, to mention some. The Maori in New Zealand are different in too many ways to enumerate here. One important distinction is that they represent a group that has become largely residentially integrated into New Zealand Pakeha society. Were Maori to be the dominant group in one particular geographical region, the effort to revive the Maori language would be more analogous to other such efforts in the world. But, this is not the case, and it is therefore widely acknowledged that Maori children will have to be fluent in Pakeha ways in order to succeed in modern New Zealand.

If it were the case that Maori were in general of higher SES and more economically privileged, the joint acquisition of Maori and Pakeha cultural and language skills could be easily envisioned.¹⁴⁶ Unfortunately, however, historical circumstances have yielded a situation in which Maori are, in general, less economically privileged than the average New Zealander, and Maori have argued that it is precisely because of the decline in the Maori culture and

¹⁴⁶ In the U.S., for example, it is not uncommon for high-income, English speaking parents to send their children to schools in which they are immersed in another language, say French.

language that Maori children are “at-risk”. If the reclamation of Maori culture succeeds in supplanting elements of a culture of poverty and lowered expectations with norms conducive to academic and scholastic achievement, the policy of reviving Maori language and culture in part through the schools may indeed lead to improved life success for Maori children. Continued research on the progress of Maori children is therefore important not only to New Zealand but also to broader international insights regarding the role of culture and language in contributing to child outcomes.

9 Broader Implications for the Social Sciences

We conclude our review with a more general assessment of our findings and a few speculative thoughts on where to go from here. The first lesson that clearly emerges from our analysis is that correlational analysis of the past is insufficient to provide firm evidence on causal linkages we most care about. Social science has advanced greatly through the use of more sophisticated techniques. Although these techniques have removed the illusion that there are a few easy “big answers” with accompanying easy policy solutions, they have helped create a new environment in which the consistent finding of small environmental effects are meaningful. Stepping back from what we have learned, we in fact think that early expectations of what the social sciences could deliver were unrealistic, and the new research environment that is emerging seems healthier and more conducive to increasing our knowledge. We genuinely view this as progress and are optimistic that better data combined with better techniques will continue to provide clearer pictures of these smaller but nevertheless important causal channels linking family/community characteristics to child outcomes. Still, we think it also pays to step back once again and judge whether other plausible explanations for the relatively small causal findings (especially with respect to home environment) exist, and whether new approaches may be needed in the future to further broaden our knowledge.

Having briefly discussed implications for research particular to the New Zealand context, we think it appropriate to add a few words regarding the broader research implications of our findings. More precisely, we attempt to address three points. The first and most obvious is addressed in Section 9.1 and concerns the use of statistics and econometrics in the social sciences and lessons learned thus far. Second, we assess in Section 9.2 the general finding of this report that statistical techniques have quantified a relatively small part of what the behavioral genetics evidence suggests is a sizable role for environment in generating child and student outcomes. Finally, we step back in Section 9.3 and briefly assess where social scientists might look to next.

9.1 *The Use of Statistics and Econometrics*

One message we hope comes through clearly in Sections 5 through 7 is just how misleading results from statistical analysis can be if the analysis does not correct for the kinds of potential problems outlined in Section 3. Again and again throughout our review, we begin with older statistical analyses showing strong potential causal effects only to find that misspecification, correlated unobservables, omitted variables, measurement error, or selection bias render these conclusions suspect or invalid. It is difficult to overstate how much further along social science might be had efforts replicating old errors gone instead into finding new and statistically more valid ways of testing old and speculative conclusions. It is equally daunting to imagine how much better informed policy makers might have been in the meantime.

Therefore, the first broad implication that probably needs no further explicit justification is that social scientists who use non-experimental data sets should use the available statistical methods to investigate findings that, while often replicated, rest on shaky foundations. Often these findings are taken for granted, as for example findings that increasing family income will improve student achievement, and equally often the statistical evidence in their support is thin once the often-repeated correlational exercises are done more carefully. Mere publication of such studies, while giving some confidence that the results have been reviewed by disinterested third parties, should not obviate the need for further critical examination. Furthermore, those researchers that continue to provide valuable correlational and statistically descriptive studies should not over-interpret their findings. We know too much to continue implying that mere correlations, even those adjusting for some factors and even those that provide very interesting and informative descriptive pictures of the world, automatically imply causations.

Related to this, it is important for social scientists in particular disciplines to occasionally look beyond their own discipline to ascertain whether their analysis might not be leaving out important variables or falling into methodological problems. The current fragmentation of the social sciences often causes findings in one discipline to seep into other disciplines only slowly, and aspects such as the literature by behavioral geneticists seem to go entirely ignored. When social scientists take too narrow a view, they can get locked into one particular way of thinking of a particular problem; it then often does not even occur to them that they may be leaving an important part of the problem out of the analysis. Often, for instance, it may be the case that researchers are perfectly aware of the abstract problem of omitted variables, but are unaware which of the variables they exclude are considered important in other disciplines.

9.2 *Do We Really Know So Little?*

A deeper question, however, involves the nature of the underlying findings in this review and the inability of decades' worth of research to causally account for a significant fraction of the large portion of the variance in child outcomes that is not accounted for by heredity. We therefore first ask whether we should be surprised that, when the statistical analysis properly accounts for the methodological difficulties outlined in Section 3, we find relatively weak evidence for many of the hypotheses we previously took for granted. Our conclusion is that previous expectations that social scientists would "discover" the "big answer" as to how the environment shapes children in general set an unreasonably high standard, and that more realistic expectations of what we can find using the methods reviewed in this paper have finally emerged.

In part this is due to the sheer number of environmental channels that probably all contribute somewhat to child development. When examining the model we laid out in Figure 1, for example, one is soon struck by the number of environmental "arrows", while only one modest arrow encompasses the heredity channel. That alone means that the roughly fifty percent of cognitive ability not accounted for by heredity may be spread between many different possible features of child environments. Even counting up the arrows representing causal environmental channels in Figure 1 is misleading, as each of those arrows has a multitude of possible sub-channels. When researchers look for evidence that any one of them contributes causally to child outcomes, it is then unlikely that they will find "large" effects in any one place. The first point to be made, therefore, is that social science researchers who search for causal environmental channels contribute much to our knowledge even if they can consistently demonstrate (after traversing the methodological minefield outlined in Section 3) effects for particular channels that are relatively small in magnitude. In the recent literature on parental involvement in schools, for instance, researchers are documenting some positive impacts from parental involvement even after carefully accounting for a myriad of statistical issues previously ignored. Even though this effect explains only a small fraction of the variance in outcomes that is not due to heredity, it is one piece of the puzzle. Others will follow, as, for example, from the experimental evidence on class size or the more carefully analyzed evidence on income and divorce.

A second and related point concerns the nature of the evidence supporting conclusions on the environment and on heredity. While the evidence regarding heredity is derived from almost-controlled experiments (involving twins, siblings and adoptions), very little of the evidence on environment is truly experimental. Social science research is messy and difficult because so much relies on evidence from non-experimental and non-controlled settings. Researchers only observe certain imprecise measures of what they are trying to capture, and they must constantly find ways of implementing statistical techniques correcting for various biases, techniques that themselves require more and better data. Given these factors, it may be difficult to document small causal effects of the environment even if they are quite real, which means that consistent findings, even if small, of particular environmental causal channels are

all the more important. The fact that the newer statistical techniques often result in statistically insignificant results where previous correlational analysis had found significance may therefore mean that the effects we are trying to measure are too small to be accurately picked up by our methods without getting better data. Their cumulative effect, however, may nevertheless be large.

Better data thus holds promise for more conclusive results on particular environmental channels, even if those results are likely to be of small magnitude. Furthermore, wherever possible, properly controlled experiments may yield insights substantially more persuasive and powerful than those derived from badly constructed non-experimental data. Within the “class size” debate reviewed in Section 6, for example, it seems a new consensus is emerging that class size does indeed have impacts on student performance, a finding that seemed quite absent from much of the literature using non-experimental data. Similarly, the relatively controlled within-classroom peer grouping experiments seem to yield much more significant insights than do other studies, and results from Gautreaux experiments seem to be picking up neighborhood effects that are difficult to find elsewhere. Of course, in many cases in the social sciences, such experiments raise serious ethical issues, which probably means that the role of experiments in answering these questions is likely to (and should) remain limited.

Finally, we should note that the relatively stronger results we can find for more extreme households that are not well represented in behavioral genetic studies are not only informative but also quite important for policy. Part of the reason that we can find such stronger effects in more extreme situation may be that multiple effects are present in these situations. In homes that are “average” on most dimensions (middle-income, average parenting style, etc.), for instance, an unusual family structure or a divorce event may have such a minor average impact on children that we have difficulty measuring it. In homes that are less average on many dimensions, on the other hand, such effects might be larger as they are important not only in and of themselves but as they also reinforce other negative aspects of home environments. Much of this is speculative, but it is consistent with the larger home environment effects found for at-risk children.

The very nature of social science research using statistical methods (i.e., the limited data available and the multitude of possible environmental channels) is then somewhat stacked against finding large average effects for any particular environmental factor. A priori, it would in fact be surprising to find some of the large effects that previous correlational analysis claimed as causal. The newest evidence indicates that our understanding of environmental channels will most likely emerge as a mosaic of small pieces rather than a few large “discoveries” of important channels which can be manipulated by policy. It further calls for better and more detailed data that makes it easier for statistical techniques to tease out even small causal channels of influence, and for modesty on the part of social science researchers searching for “the big answer”. Ruling out “big answers” that have been hypothesized is progress as well, and it may be that the more sophisticated use of data combined with a recognition of the “mosaic” approach that is emerging signifies a maturation of the social sciences that holds much promise with less glamour.

9.3 *What Next?*

Nevertheless, given that much of what we thought we knew about parents, neighborhoods and schools has yet to be persuasively demonstrated, the lack of general understanding (arising from the studies we reviewed) of the environment’s impact on child outcomes is puzzling, even stunning. It is therefore fitting to end the body of this review with some possible explanations for this puzzle, explanations that go beyond the observations regarding data limitations and multitudes of channels. We propose the possibility of four effects that may account for a continuing dearth of explanations even as our methods and data sets become more sophisticated.

First, in almost everything said in this review, there is a common underlying assumption that “environment” can be objectively quantified in a way that is similarly meaningful to all children. In the behavioral genetics studies, for instance, it is demonstrated that the sum of all “shared” home environments can at best account for a small portion of the non-genetic variance in cognitive abilities among adults. Similarly, in the more careful statistical analyses we review, we often find that various aspects of environments such as family structure do not seem to be causally linked to outcomes.

It may, however, be the case that certain objectively observed aspects of environments impact different children quite differently. We made the point in Section 4, for instance, that behavioral genetic studies do not rule out a large impact of home environments if such environments differ across children within a household. While one interpretation of this is that we should focus more on objective differences in treatments of different children within the home (such as the different treatment of first borns), another might be that perceptions of home environments differ. We cited evidence in our review, for example, that such perceptions might have a heritable component (i.e., children of different genetic backgrounds may report different impressions of the same environment). Similarly, certain shared home environment differences that do not contribute to an explanation of differences in outcomes may cause differences in peer and neighborhood effects. It may be, for example, that divorce causes some children to experience stress and anxiety that takes away from their cognitive development, while it causes other children to divert their attention from the unpleasant situation by concentrating more on schools or peers. If different children, for whatever reasons, experience the same environment differently, then it will be difficult to find statistical evidence for environmental causal channels without finding significantly different types of data. Even though they are not within the scope of this review, we are aware of non-statistical work by psychologists that proceeds from this perspective.

Second, as mentioned in our review, Harris (1995) proposed the theory that children are more active in creating their own environment than previously thought. This process starts early in life and becomes more important as time passes, the theory goes. This would be consistent with behavioral genetic and statistical evidence that certain home environments become less important with time, and that the non-genetic portion of variances in outcomes becomes more difficult to explain as children enter adolescence and adulthood. It further opens the possibility for a level of interaction between genes and environment that is a great deal more complex than what we have presented in this review, with children of different genetic background creating different environments for themselves in part due to these genetic differences. As Plomin and Petrill (1997) comment, “genes (Rodgers, Rowe et al.) contribute to environment itself”. This interaction between genes and environments holds much promise for more definitive findings and policy implications in the future.

Third, and here we really are reaching far beyond our expertise, scientists have learned much in the past decades about the natural world that can be explained by “chaos theory”, a theory that essentially states that small differences in initial conditions may cause large differences in outcomes, and that certain influences may cause dramatically different paths depending on the initial conditions. In the case of child development, it may be the case that children who are not all that different initially (young siblings who tend to resemble their parents and each other) are sufficiently different to cause very different impacts of the same environmental factors (such as divorce, school quality, peers) even if these factors are perceived similarly. Some children, for example, may react poorly if they witness their parents unemployed, while others might find in the unpleasant state reasons to achieve. This example is overly simplistic, but it is not difficult to think of more complex ways in which similar events in one’s life may cause dramatic impacts in some cases but not in others. If this is a large part of the reason for

the relative lack of findings on home environments, for instance, it is unlikely that statistical methods of the kind we review will ever be able to uncover them.¹⁴⁷

Finally, at least some of the evidence cited in this review suggests that current statistical methods may miss important effects known as “threshold effects.” For certain environmental factors, for instance, small changes in the factor may have little impact on outcomes while large changes do over time. In the case of neighborhood and culture effects, for instance, we have cited theoretical evidence suggesting such effect. Models in the social sciences that formalize this are often called “tipping models”. Certain healthy communities, for instance, may “tip” and become unhealthy ghettos at some critical point. While such effects are discussed throughout the psychology, sociology and economics literatures, only now are statistical methods being developed that will allow for empirical testing.

¹⁴⁷ At the same time, it is unclear to us how exactly one would go about studying either of the two above possibilities (i.e., that objectively similar environments are perceived differently, or that similarly perceived environments have dramatically different impacts). One method, for example, involves “critical incidents studies” where subjects are asked to discuss important events in their lives that have impacted them. While this can certainly be an interesting exercise, it is easy to misread results. Answers are likely to depend on popular trends of what is currently thought to matter, with people in one time and place blaming parents for problems they may have and people in another time and place finding fault with teachers or communities. We are therefore skeptical of any methodology that is based on introspection on the part of subjects.

10 Summary

Our review started in Section 2 with the development of a broad conceptual model linking family and community characteristics to child outcomes, a model intended to capture the main causal avenues through which such links might be established. The picture summarizing this model by itself should give us pause because the model contains so many possible ways that linkages might occur, it seems difficult to imagine how to unravel the messy data social science researchers work with to come to conclusions regarding which causal links are of significance. We therefore proceeded in Section 3 with a discussion of the general methodological issues that enable researchers to overcome the problem of being faced with data that is typically not from controlled experiments and to extract useful information mirroring the kind of information such experiments would provide. This section makes clear that the mere finding of correlations between family/community characteristics and child outcomes, even if these hold some factors constant, is far from sufficient to attribute causal significance to most links identified in our conceptual model. There are many statistical tools to assist in obtaining greater certainty that correlations are indeed causal.

Section 4 then proceeded with a review of some relatively unique evidence that comes close to the kind of evidence derived from controlled experiments. This evidence on twins, siblings and adopted children has provided insights that, without the need for the application of many of the statistical tools discussed in Section 3, suggests a strong role (50 percent) for non-genetic factors in explaining variations in cognitive ability. At the same time, this evidence also indicates that home environmental factors which are shared by siblings within a home together account for very little of this non-genetically determined variation in outcomes, at least for outcomes measured in later childhood and adolescence in circumstances where home environments were not “extreme”. Although early correlational analysis of non-experimental data seemed at odds with this conclusion, we argue in Section 5 that much of the more recent social science evidence on home environments is actually quite consistent with it. This more recent evidence uses modern statistical techniques to go beyond painting descriptive pictures of the data and toward finding more causal links. Factors such as family income, long thought to be crucial for child development, are now thought to be of minor significance, at least once family income has reached a minimal threshold level. Other factors, family structure, supervised child care, etc. , are also shown to have smaller effects than previously thought, but the effects are larger for at-risk children from more “extreme” backgrounds.

All this suggests that the large role for the environment to explain outcomes measured in adolescence and adulthood may, at least in great part, come from environments outside the home. Candidates for such environmental influences include schools and broadly conceived notions of communities, notions that include a role for race, ethnicity and culture as well as residential neighborhood. Sections 6 and 7 explore these possibilities. We find some recent evidence that parental involvement in schools may have positive effects, possibly for all children in the school or possibly only for the child whose parents become involved, although these effects are not large. We also find evidence that at least some aspects of financial resources within schools matter for outcomes, and that low SES children might perform better when mixed with high SES children. However, while the latter finding is important, it is unclear whether it represents the existence of peer effects or the impact of teacher expectations and quality. Finally, while we join researchers in the area in being optimistic that more evidence in favor of more general neighborhood effects will be found in the future, we found little such evidence that has accounted for various statistical biases in the existing literature. Some such evidence, however, is emerging from controlled experiments.

Although much of the evidence we report on is derived from studies conducted outside New Zealand, we argue strongly that the results are quite applicable and informative for New Zealand researchers. In particular, we think the evidence can help to fine-tune research that investigates unique features of New Zealand and to interpret existing evidence correctly.

Mostly, we urge New Zealand readers to not dismiss the evidence we present simply because some aspects of New Zealand are different or unique, and we suggest that the evidence is consistent with our belief that, despite cultural and temperamental differences, similarities of human beings across contexts often outweigh differences. We concluded our review with a final section on more general research implications and an optimistic assessment of the progress social science has made in the past decade. While what we think we know may be less, what we actually know has increased dramatically.

Tables

Table 6.1.2: Summary of studies relating family background to parental involvement

Table 6.1.3: Summary of studies that relate parental involvement to student outcomes

Table 6.2.1: Summary of studies that relate continuous measures of peer characteristics to student achievement

Table 6.2.2: Summary of studies that relate continuous measures of peer characteristics to student attainment

Table 6.3.1: Percentage distributions of estimated effects of selected school inputs on student performance

Table 7.2.1: Summary of studies that relate neighbourhood characteristics to cognitive ability or achievement

Table 7.2.2: Summary of studies that relate neighbourhood characteristics to student attainment

Table 6.1.2
Summary of Studies Relating Family Background to Parental Involvement

Study	Data	Grade level(s)	Time period	Estimation method	Measure of parental involvement	Effect of 1 S.D. increase in measures of family background on parental involvement			
						Measure of SES	Effect	Race or ethnicity (relative to whites)	Effect
(McMillan 1999)	National Education Longitudinal Survey (NELS) (738 public schools)	Grade 8	1988	Ordinary least squares	% of parents in school who are active in Parent-Teacher Association	Income	0.37	---	---
						Parental education	0.29		
(Zellman and Waterman 1998)	Los Angeles, California (193 students)	Grades 2 and 5	?	Ordinary least squares	Index (frequency of attendance at scheduled events, participation on school committees, volunteer activity, employment at school, attendance at Parent-Teacher Association meetings)	Index of SES	NS	Latino	0.22
						Child IQ	NS	Black	0.19
(Sui-Chu and Willms 1996)	National Education Longitudinal Survey (NELS) (24,599 public and private school students)	Grade 8	1988	Hierarchical linear model	Index of school communication (parent contacts school and school contacts parent)	(communication) Index of SES	0.18	(communication) Asian	-0.23
								Latino	NS
								Black	0.06
								Native American	-0.14
		Index of school participation (volunteers at school, attends Parent-Teacher Association meetings)	(participation) Index of SES	0.16	(participation) Asian	-0.19			
					Latino	-0.07			
					Black	NS			
					Native American	-0.07			
(Kerbow and Bernhardt 1993)	National Education Longitudinal Survey (NELS) (22,150 public and private school students)	Grade 8	1988	Hierarchical linear model	Frequency of parent-initiated contact with school	(contact) Index of SES	0.32	(contact) Asian	-0.12
								Latino	0.08
								Black	0.09
								Involvement in parent-teacher association	(PTA) Index of SES
					Latino	NS			
					Black	NS			

(Stevens on and Baker 1987)	National U.S. (179 students)	Grades K-12	1981-82	Ordinary least squares	Involvement in Parent-Teacher Association and parent-teacher conferences	Mother's education	0.27	---	---
(Baker and Stevenson 1986)	41 students	Grade 8	?	Ordinary least squares	Index (met with one of child's teachers, attended teacher conference, other measures)	Mother's education	Positive ^a	---	---

^a Unstandardized regression coefficient could not be transformed into standard deviation units.

Note: NS indicates that effects were not statistically distinguishable from zero. We report effects from the most fully-specified regression. Effects are presented in standard deviation units (when unstandardized regression coefficients were reported, they were transformed using the reported standard deviations of dependent and independent variables).

Table 6.1.3
Summary of Studies that Relate Parental Involvement to Student Outcomes Study

Study	Data	Grade level(s)	Time period	Estimation method	Measure of student outcome	Family background control variables	Measure of parental involvement	Effect of 1 S.D. increase in measures of parental involvement on student outcome
(McMillan 1999)	National Education Longitudinal Survey (NELS) (12 140 public school students)	Grade 8	1988	Instrumental variables ^a	Reading	Index of socioeconomic status, parental	Individual student's parents are active in parent-teacher association	0.36
							% of parents in school who are active in parent-teacher association	0.18
(Zellman and Waterman 1998)	Los Angeles, California (193 students)	Grades 2 and 5	?	Ordinary least squares	Reading	Race, index of socioeconomic status, child IQ	Index (frequency of attendance scheduled events, participation on school committee, volunteer activity, employment at school, attendance at Parent-Teacher Association meetings)	NS
(Griffith 1997)	Suburban U.S. district (41 schools)	Grades 1-6	?	Ordinary least squares	General test battery (reading, mathematics, science, etc.)	Race, % free-and-reduced lunch	Index (frequency of parental participation in volunteer activities, attending Parent-Teacher Association meetings, attending school activities)	0.54
(Sui-Chu and Willms 1996)	National Education Longitudinal Survey (NELS) (24,599 public and private school students)	Grade 8	1988	Hierarchical linear model	Reading Mathematics	Race, socioeconomic status	Index of school communication (parent contacts school and school contacts parent) (individual level)	-0.06 (reading) -0.05 (math)
							Index of school participation (volunteers at school, attends parent-teacher association meetings) (individual level)	0.03 (reading) 0.03 (math)
							Index of school participation (school level)	0.08 (reading) NS (math)
(Paulson 1994)	Indiana/Virginia (247 students)	Grade 9	?	Ordinary least squares	Self-reported grades	None	Index (attendance at parent-teacher conferences and school activities, volunteer work, other measures)	Positive for mother's involvement in subsamples of ^b boys and girls

(Muller 1993)	National Education Longitudinal Study (NELS) (approximately 18,000 public and private school students)	Grade 8	?	Ordinary least squares	Composite of reading and mathematics achievement	Race, parental education, income	Frequency of parental contact with school	-0.06 (achievement) -0.11 (grades)
					Composite of self-reported grades		Participation in Parent-Teacher Association	NS (achievement) 0.06 (grades)
							Parent volunteers at school	NS (achievement) 0.02 (grades)
(Reynolds 1992)	Chicago, Illinois (481 students)	Grade 2	1988	Ordinary least squares	Composite of reading and mathematics (year 2)	Parental education, free and reduced lunch	Index of parent ratings (communicates with school, participates in school activities, helps in child's classroom, talks with teacher, attends parent meetings, other measures)	0.10 (year 2) -0.09 (year 3)
					Composite of reading and mathematics (year 3)		Index of teacher ratings of parents (provides learning experiences, communicates with school, participates in school activities)	0.29 (year 2) 0.30 (year 3)
							Child reports that parent comes to school	-0.15 (year 2) -0.14 (year 3)
(Epstein 1991)	Baltimore, Maryland (293 students)	Grades 3 and 5	1980-81	Ordinary least squares	Reading Mathematics	Index of socioeconomic status	Index (parent reports of teacher requests for involvement and parents' ratings of quality of assigned homework)	0.13 (reading) NS (math)
(Stevenson and Baker 1987)	National U.S. (179 students)	Grades K-12	1981-82	Ordinary least squares	Student performance (teacher rating)	Mother's education	Parental involvement in Parent-Teacher Association and parent-teacher conferences	0.36
(Iverson, Brownlee et al. 1981)	U.S. (398 students)	Grades 1-8	?	Ordinary least squares	Reading	Race	Frequency of parent-teacher contacts	Positive ^b

^a Excluded instruments included the percentage of parents active in non-school organizations

^b Unstandardized regression coefficients could not be transformed into standard deviation units, or regression coefficients were not reported

Note: NS indicates that effects were not statistically distinguishable from zero. We report effects from the most fully-specified regression. Effects are presented in standard deviation units (when unstandardized regression coefficients were reported, they were transformed using the reported standard deviations of dependent and independent variables).

Table 6.2.1
Summary of Studies that Relate Continuous Measures of Peer Characteristics to Student Achievement

Study	Data	Grade level(s)	Time period	Estimation method	Measure of student outcome	Teacher/school control variables	Level of aggregation of peer characteristic	Measure of peer characteristic	Effect of 1 S.D. increase in measures of peer characteristics on student outcome
(McEwan 1999)	Chilean <i>Sistema de Medición de Calidad de la Educación</i> (62,684 public students)	Grade 8	1997	Ordinary least squares	Spanish Mathematics	None	Classroom	Mean mother's education Percent Native American	0.18 (Spanish) 0.19 (math) -0.03 (Spanish) -0.02 (math)
(Caldas and Bankston 1997)	Louisiana Graduate Exit Examination (42,041 students)	Grade 10	1989-90	Ordinary least squares	Composite achievement in mathematics, English, composition	None	School	Percentage of student receiving free-and-reduced lunch Mean social status	0.08 0.12
(Zimmer and Toma 1997) ^a	IEA (International Association for the Evaluation of Educational Achievement): Belgium, France, New Zealand Ontario, and United States (public school events)	Grade 8 (or grade equivalent of ages 13-14)	1981	Ordinary least squares	Mathematics	Teacher gender and experience for some countries	Classroom	Percentage black Mean achievement	-0.14 NS (Belgium) NS (France) Positive (New Zealand) _b Positive (Ontario) _b Positive (United States)
(Robertson and Symons 1996) ^a	UK National Child Development Study (approximately 4,500 students in tracked and untracked schools)	Ages 7-33	1965-91	Instrumental variables ^c	Reading Mathematics Earnings	Teacher experience, class size	School	Percent high SES	Positive (reading in tracked schools) _b NS (reading in untracked schools) _b Positive (math in tracked schools) _b Positive (math in untracked schools) NS (earnings for males and females)

(Link and Milligan 1991)	Study of the Sustaining Effects of Compensatory Education of Basic Skills	Grades 3-6	1976-77	Ordinary least squares	Reading Mathematics	Class size	Classroom	Mean achievement	0.05 (third-grade reading, black students)
									0.11 (third-grade reading, Latino students)
									0.06 (third-grade reading, white students)
									0.12 (third-grade math, black students)
									0.13 (third-grade math, Latino students)
									0.08 (third-grade math, white students)
								Percent of same race	0.12 (third-grade reading, black students)
									-0.06 (third-grade reading, Latino students)
									-0.06 (third-grade reading, Latino students)
									NS (third-grade reading, white students)
									NS (third-grade math, black students)
									NS (third-grade math, Latino students)
									0.02 (third-grade math, white students)

(Bryk and Driscoll 1988)	High School and Beyond	Secondary	1980-82	Hierarchical linear model	Mathematics	School size	School	Mean achievement	-0.25
								Mean social class	0.56
								Minority concentration	0.30
(Gamoran 1987)	High School and Beyond (12,418-13,254 students)	Secondary	1980-82	Ordinary least squares	Mathematics Science Reading Vocabulary Writing Civics	Advanced courses offered	School	Mean SES	NS (mathematics) NS (science) NS (reading) NS (vocabulary) NS (writing) NS (civics)
								Mean achievement	0.07 (mathematics) NS (science) NS (reading) NS (vocabulary) NS (writing) NS (civics)
								Percentage black	NS (mathematics) -0.01 (science) NS (reading) -0.01 (vocabulary) NS (writing) NS (civics)
								Percentage Latino	NS (mathematics) NS (science) NS (reading) NS (vocabulary) NS (writing) NS (civics)
(Willms 1986)	Scottish School Leavers Survey (23,151 students)	Secondary	1981	Ordinary least squares	Composite of "O-grade" examinations	None	School	Mean SES	0.40
(Henderson, Mieszkowski et al 1978)	French-speaking students in Montreal (7,000 students)	Grades K-3	1976?	Ordinary least squares	French Mathematics	Class size, teacher experience, teacher education	Classroom	Mean IQ	Positive ^b (French and math in grades 1-3)

(Summers and Wolfe 1977)	Philadelphia (627 students)	Grade 6	1970-71	Ordinary least squares	Composite achievement score	Quality of teacher education, teacher experience, teacher exam score, class size, school size, library books per pupil	School	Percent high achievers	Positive ^b
								Percent low achievers	Negative ^b
								20-40 percent black students (relative to 0-20 percent)	Positive ^b
								40-60 percent black students (relative to 0-20 percent)	Positive ^b
(Winkler 1975)	Urban U.S. school district (773 students)	Secondary	1964-65	Ordinary least squares	Reading	Expenditures per pupil, teacher salaries, quality of teacher education, student-teacher ratio	School	Percent low SES	-0.31 (white students) NS (black students)
								Percent black	0.62 (white students) NS (black students)

^a These studies also test for non-linearities in peer effects. Henderson et al. (1978) find diminishing marginal returns to peer quality (quadratic peer terms that are negative and statistically significant). Zimmer and Toma (1997) find diminishing marginal returns in the public school samples of New Zealand, Ontario, and the United States (but not in Belgium and France). Robertson and Symons (1996) find diminishing marginal returns only in tracked schools in math. 125

^b Unstandardized regression coefficients could not be transformed into standard deviation units.

^c Instrumental variables include area of birth in 11 regions of the United Kingdom.

Note: NS indicates that effects were not statistically distinguishable from zero. We report effects from the most fully-specified regression. Effects are presented in standard deviation units (when unstandardized regression coefficients were reported, they were transformed using the reported standard deviations of dependent and independent variables). Link and Mulligan (1991) also report results for grades 4, 5, and 6. 126

Table 6.2.2
Summary of Studies that Relate Continuous Measures of Peer Characteristics to Student Attainment

Study	Data	Grade level(s)	Time period	Estimation method	Measure of student outcome	Teacher/school control variables	Level of aggregation of peer characteristic	Measure of peer characteristic	Effect of 1 S.D. increase in measures of parental involvement on student outcome ^b
(Gaviria and Raphael 1997)	National Education Longitudinal Survey (8,790 students)	Grade 10	1990	Instrumental variables ^a	Whether student has dropped out of high school	Private/public	School	Percentage of other students in school who have dropped out	Positive
(Mayer 1991)	High School and Beyond (26,425 students)	Secondary	1982	Maximum likelihood	Whether student has dropped out of high school	None	School	Mean SES	-0.025 (white) -0.014 (black) -0.023 (Latino)
(Bryk and Driscoll 1988)	High School and Beyond	Secondary	1980-82	Hierarchical linear model	Whether student has dropped out of high school	School size	School	Mean achievement	0.008
								Mean social class	NS
								Minority concentration	0.009

^a Instrumental variables include school averages of selected individual variables (such as the percentage in single-parent families and the percentage who have some college education).

^b The standard deviation for the independent variable was not provided.

^c Effects are taken from Mayer's Table 4 (the estimated change in probability of dropping out if the average student moves from the average SES school to the high-SES school, an increase of one standard deviation in mean SES).

Note: NS indicates that effects were not statistically distinguishable from zero. We report effects from the most fully-specified regression. In regressions with dichotomous dependent variables (such as dropping-out), the effects are presented as the change in probability.

Table 6.3.1
 Percentage Distributions of Estimated Effects of Selected School Inputs on Student Performance

Resources	Number of estimates	Statistically significant		Statistically insignificant		
		Positive	Negative	Positive	Negative	Unknown sign
Teachers						
Teacher-pupil ratio	277	15%	13%	27%	25%	20%
Teacher education	171	9%	5%	33%	27%	26%
Teacher experience	207	29%	5%	30%	24%	12%
Teacher test-score	41	37%	10%	27%	15%	12%
Financial aggregates						
Teacher salary	119	20%	7%	25%	20%	28%
Expenditure per-pupil	163	27%	7%	34%	19%	13%
Other inputs						
Administrative inputs	75	12%	5%	23%	28%	32%
Facilities	91	9%	5%	23%	19%	44%

Source: Adapted from Tables 3 and 4, Hanushek (1997). 128

Table 7.2.1
Summary of Studies that Relate Neighbourhood Characteristics to Cognitive Ability or Achievement

Study	Major data source	Grade level(s) or age(s)	Time period	Estimation method	Measure(s) of student outcome	Level of aggregation of neighbourhood characteristic	Measure(s) of neighbourhood characteristic	Effect of 1 S.D. increase in measure(s) of neighbourhood characteristics on student outcome
(Chase-Lansdale, Gordon et al. 1997)	Infant Health and Development Program (IHDP) (793 children)	Ages 3-4 Ages 5-6	1985	Ordinary least squares	Vocabulary and verbal ability IQ	Census tract	Low SES (Index of female-headed families, poverty, other measures)	NS (IQ, ages 3-4) NS (vocab., ages 3-4) NS (IQ, ages 5-6) NS (vocab., ages 5-6)
							High SES (Index of schooling, professional occupations, income)	0.08 (IQ, ages 3-4) NS (vocab., ages 3-4) 0.07 (IQ, ages 5-6) NS (vocab., ages 5-6)
							Male joblessness (Index of unemployed and part-time employment)	NS (IQ, ages 3-4) NS (vocab., ages 3-4) NS (IQ, ages 5-6) NS (vocab., ages 5-6)
							Ethnic diversity (Index of foreign-born individuals, percent Latino, other measures)	NS (IQ, ages 3-4) NS (vocab., ages 3-4) -0.17 (IQ, ages 5-6) -0.16 (vocab., ages 5-6)
							Family concentration (Index of ratio of persons to occupied dwellings, percent aged 0-17 and 65+)	NS (IQ, ages 3-4) NS (vocab., ages 3-4) NS (IQ, ages 5-6) NS (vocab., ages 5-6)

(Chase-Lansdale, Gordon et al. 1997)	Children of the National Longitudinal Survey of Youth (NLSY) (862 preschoolers and 697 early school-age children)	Ages 3-4 Ages 5-6	1985	Ordinary least squares	Vocabulary and verbal ability Reading recognition Mathematics	Census tract	Low SES (Index of female-headed families, poverty, other measures) High SES (Index of schooling, professional occupations, income) Male joblessness (Index of unemployed and part-time employment) Ethnic diversity (Index of foreign-born individuals, percent Latino, other measures) Family concentration (Index of ratio of persons to occupied dwellings, percent aged 0-17 and 65+)	NS (IQ, ages 3-4) NS (vocab., ages 3-4) NS (IQ, ages 5-6) NS (vocab., ages 5-6) NS (IQ, ages 3-4) NS (vocab., ages 3-4) NS (IQ, ages 5-6) NS (vocab., ages 5-6) NS (IQ, ages 3-4) NS (vocab., ages 3-4) NS (IQ, ages 5-6) -0.12 (vocab., ages 5-6)
(Halpern-Felsher, Connell et al. 1997)	Upstate New York school district (1040 students)	Grades 3-5	1990-91	Ordinary least squares	Index of educational risk (low attendance, low test scores, other measures)	Census tract	Neighbourhood risk factor (Index of distribution of SES and male employment patterns)	NS (white males) NS (black males) NS (white females) NS (black females)
(Halpern-Felsher, Connell et al. 1997)	Atlanta (346 black students)	Grades 6-8	1989-90	Ordinary least squares	Iowa Test of Basic Skills	Census tract	Neighbourhood risk factor (Index of distribution of SES and male employment patterns)	NS (black males) NS (black females)
(Halpern-Felsher, Connell et al. 1997)	New York City/Baltimore/Washington DC (669 students)	Grades 5-6 and 8-9	1988-89	Ordinary least squares	Composite of reading and mathematics	Census tract	Neighbourhood risk factor (Index of distribution of SES and male employment patterns)	NS (white males) NS (black males) -0.16 (white females) NS (black females)

(Halpern-Felsher, Connell et al. 1997)	Upstate New York school district (3406 students)	Grades 6-8	1989-91	Ordinary least squares	Index of educational risk (low attendance, low test scores, other measures)	Census tract	Neighbourhood risk factor (Index of distribution of SES)	0.13 (white males) 0.06 (black males) 0.08 (white females) 0.07 (black females)
(Halpern-Felsher, Connell et al. 1997)	Upstate New York school district (1,797 students)	Secondary	1987-88	Ordinary least squares	Index of educational risk (low attendance, low test scores, other measures)	Census tract	Neighbourhood risk factor (Index of distribution of SES and male employment patterns)	0.13 (white males) 0.07 (black males) NS (white females) 0.06 (black females)
(Entwisle, Alexander et al. 1994)	Baltimore (455 students)	Grade 1	1982	Hierarchical linear model	Mathematics	26 "Regional Planning Districts"	Median neighbourhood income	NS ^b
(Brooks-Gunn, Duncan et al. 1993)	Infant Health and Development Program (IHDP) (765 children)	Age 3	1985	Ordinary least squares	IQ	Census tract	% families with income < \$10,000 % families with income > \$30,000	NS 0.19
(Dornbusch, Ritter et al. 1991)	All students in 6 San Francisco schools (3,084 students)	Secondary	1987-88	Ordinary least squares	Self-reported grades	Census tract	Community socioeconomic status (Index of per capita income, families above poverty, other measures)	0.09 (white students) 0.19 (black students)
(Garner and Raudenbush 1991)	95 percent sample of one Scottish educational authority (2,500 students)	Secondary	1983-84	Hierarchical linear model	Composite of "O-grade" examinations	Census enumeration district	"Deprivation score" (Index of unemployment, single-parent families, earnings, illness and other variables)	Negative

^a Effect sizes were calculated using the standard deviations of the dependent variables (Chase-Lansdale, Gordon et al. 1997), and the standard deviations of the independent variables for blacks reported by Duncan and Aber (1997), given that standard deviations for the entire sample are not reported.

^b Unstandardized regression coefficients could not be transformed into standard deviation units.

Note: NS indicates that effects were not statistically distinguishable from zero. We report effects from the most fully-specified regression. Effects are presented in standard deviation units (when unstandardized regression coefficients were reported, they were transformed using the reported standard deviations of dependent and independent variables).

Table 7.2.2
Summary of Studies that Relate Neighbourhood Characteristics to Student Attainment

Study	Major data source	Grade level(s) or age(s)	Time period	Estimation method	Measure(s) of student outcome	Level of aggregation of neighbourhood characteristic	Measure(s) of neighbourhood characteristic	Effect of 1 S.D. increase in measure(s) of characteristics on student outcome
(Cutler and Glaeser 1997)	1990 1% US Census Public Use Micro Sample	Ages 20-24 and 25-30	1990	Instrumental variables (various)	High school graduation	Metropolitan Statistical Area (MSA)	Index of Segregation (high school graduation)	<u>I.V.: fiscal variables</u> Positive ^a (white) Negative ^a (black)
					College graduation		(college graduation)	<u>I.V.: topography</u> NS (white) Negative ^a (black)
(Cutler and Glaeser 1997)	National Longitudinal Survey of Youth	Ages 26-33	1990	Instrumental variables (lagged fiscal variables)	High school graduation	Metropolitan Statistical Area (MSA)	Index of Segregation (high school graduation)	<u>I.V.: lagged residence</u> Positive ^a (white) Negative ^a (black)
					College graduation		(college graduation)	<u>I.V.: fiscal variables</u> Positive ^a (white) Negative ^a (black)
(Cutler and Glaeser 1997)	National Longitudinal Survey of Youth	Ages 26-33	1990	Instrumental variables (lagged fiscal variables)	High school graduation	Metropolitan Statistical Area (MSA)	Index of Segregation (high school graduation)	<u>I.V.: topography</u> NS (white) NS (black)
					College graduation		(college graduation)	<u>I.V.: lagged residence</u> Positive ^a (white) Negative ^a (black)

(Halpern-Felsher, Connell et al. 1997)	Panel Study of Income Dynamics (PSID) (3,395 individuals)	Ages 16-21	1968-91	Ordinary least squares	Years of completed schooling	Census tract	Neighbourhood risk factor (Index of distribution of SES and male employment patterns)	-0.12 (white males) -0.08 (black males) -0.16 (white females) -0.09 (black females)
(Ensminger, Lamkin et al. 1996)	Chicago (Woodlawn) sample (277 males and 288 females)	Teen-agers	1975-76	Logit	Dropping out of high school	Census tract	% below poverty line	NS (males) NS (females)
							% white-collar	Negative ^a (males) Negative (females)
				Ordinary least squares	Years of completed schooling		% below poverty line	NS (males) NS (females)
							% white-collar	Positive ^a (males) Positive (females)
Duncan (1994)	Panel Study of Income Dynamics (PSID) (3,439 individuals)	Ages 16-22	1968-91	Ordinary least squares	Years of completed ^b schooling	Census tract	% families with income<\$10,000	0.23 (white males) NS (black males) NS (white females) 0.21 (black females)
							% families with income>\$30,000	0.25 (white males) NS (black males) 0.20 (white females) 0.20 (black females)
							% black	NS (white males) -0.18 (black males) NS (white females) NS (black females)
							% female-headed families	-0.04 (white males) NS (black males) NS (white females) -0.23 (black females)
							% adult woman working 26+ weeks	0.08 (white males) NS (black males) NS (white females) -0.10 (black females)
(Brooks-Gunn, Duncan et al. 1993)	Panel Study of Income Dynamics (PSID) (1,980 females)	Ages 14-19	1968-85	Logit	Dropping out of high school	Census tract	% families with income<\$10,000	NS
							% families with income>\$30,000	-0.05

(Borjas 1992)	General Social Surveys (6,756 individuals)	Ages 18-64	1977-89	Ordinary least squares	Years of education	Ethnic categories	Mean education of fathers	Positive ^a
(Borjas 1992)	National Longitudinal Survey of Youth (NLSY) (5,619 individuals)	Ages 14-29	1979-87	Ordinary least squares	Years of education	Ethnic categories	Mean education of fathers	Positive ^a
(Crane 1991)	15% Neighbourhood Characteristic File of Public Use Microdata Samples (92,512 individuals)	Ages 16-19	1970	Logit	Dropping out of high school	Neighbourhoods of 4000-5000 people	% professional or managerial workers	Negative ^a (whites, blacks, Latinos)
(Datcher 1982)	Panel Study of Income Dynamics (PSID) (552 males)	Ages 23-32	1978	Ordinary least squares	Years of schooling	Zip (postal) codes	Average neighbourhood income % white	Positive ^a (white) NS (black) NS (white) NS (black)

^a Unstandardized regression coefficients could not be transformed into standard deviation units, or insufficient information was provided to estimates the change in probability due to a one standard deviation increase in the peer characteristic.

^b In another set of analyses, the author relates dichotomous variables (high school dropping out and college attendance) to the same family and neighbourhood measures.

Note: NS indicates that effects were not statistically distinguishable from zero. We report effects from the most fully-specified regression. In regressions with continuous dependent variables, effects are presented in standard deviation units (when unstandardized regression coefficients were reported, they were transformed using the reported standard deviations of dependent and independent variables). In regressions with dichotomous dependent variables (such as dropping-out), the effects are presented as the change in probability.

References

- Alexander, K. L., E. L. McDill, et al. (1979). "School SES Influences--Composition or Context?" Sociology of Education **52**: 222-237.
- Alwin, D. F. (1991). "Family of Origin and Cohort Differences in Verbal Ability." American Sociological Review **56**: 625-638.
- Amato, P. R. (1993). "Family Structure, Family Process, and Family Ideology." Journal of Marriage and Family **55**(1): 50-54.
- Amato, P. R. and B. Keith (1991). "Parental Divorce and Adult Well-Being: A Meta-Analysis." Journal of Marriage and the Family **53**(1): 43-58.
- Ambert, A. M. (1994). "An International Perspective on Parenting: Social Change and Social Constructs." Journal of Marriage and the Family **56**: 529-543.
- Angrist, J. and V. Lavy (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." The Quarterly Journal of Economics **114**(2): 533-576.
- Anh, T. S., J. Knodel, et al. (1998). "Family Size and Children's Education in Vietnam." Demography **35**(1): 57-70.
- Argys, L. M., D. I. Rees, et al. (1996). "Detracking America's Schools: Equity at Zero Cost?" Journal of Policy Analysis and Management **15**(4): 623-645.
- Ashenfelter, O. and A. Krueger (1994). "Estimates of the Economic Returns to Schooling from a New Sample of Twins." American Economic Review.
- Ashenfelter, O. and C. Rouse (1998). "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins." Quarterly Journal of Economics **113**: 253-284.
- Astone, N. M. and S. S. McLanahan (1991). "Family Structure, Parental Practices and High School Completion." American Sociological Review **56**: 309-320.
- Axinn, W. and A. Thornton (1996). "The Influence of Parents' Marital Dissolutions on Children's Attitudes toward Family Formation." Demography **33**: 66-81.
- Baker, D. P. and D. L. Stevenson (1986). "Mothers' Strategies for Children's School Achievement: Managing the Transition to High School." Sociology of Education **59**: 156-166.
- Baldwin, A. L., Baldwin, C., Cole, R. E. (1990). Stress Resistant Families and Stress Resistant Children. Risk and Protective Factors in the Development of Psychopathology. J. Rolf, Masters, A.S., Cicchetti, D., Neuchterlein, K., Weintraub, S. Cambridge, UK, Cambridge University Press.
- Barker, G. and T. Maloney (1999). Final Regression Results on the Cognitive Achievement of Children in the Christchurch Health and Development Study with Corrections for Attrition from this Longitudinal Study. Wellington, NZ, The New Zealand Treasury.
- Barnett, W. S. (1995). "Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes." Future of Children **5**(3): 25-50.

- Baumrind, D. (1971). "Current patterns of parental authority." Developmental Psychology **4**: 1-103.
- Baydar, N., Brooks-Gunn, J. (1991). "Effects of Maternal Employment and Child-Care Arrangements on Preschoolers' Cognitive and Behavioral Outcomes: Evidence from the Children of the National Longitudinal Survey of Youth." Developmental Psychology **27**: 932-45.
- Bayer, P. J. (1999). *An Empirical Analysis of the Equilibrium in the Education Market*, Stanford University.
- Becker, G. and N. Tomes (1976). "Child Endowments and the Quantity and Quality of Children." Journal of Political Economy **84**: S143-S162.
- Becker, G. S. (1991). A Treatise on the Family. Cambridge, Harvard University Press.
- Becker, G. S., K. M. Murphy, et al. (1990). "Human Capital, Fertility and Economic Growth." Journal of Political Economy **98**(5): S12-S37.
- Behrman, J. and M. Rosenzweig (1999). "'Ability' Biases in Schooling Returns and Twins: A Test and New Estimates." Economics of Education Review **18**: 159-67.
- Behrman, J. and P. Taubman (1986). "Birth Order, Schooling and Earnings." Journal of Labor Economics **4**: S121-45.
- Behrman, J. and P. Taubman (1989). "Is Schooling 'Mostly in the Genes'? Nature-Nurture Decomposition Using Data on Relatives." Journal of Political Economy **97**: 1425-46.
- Behrman, J. R., R. A. Pollack, et al. (1989). "Family Resources, Family Size, and Access to Financing for College Education." Journal of Political Economy **97**(2): 398-419.
- Belsky, J., Eggebeen, D. (1991). "Early and Extensive Maternal Employment and Young Children's Socioemotional Development: Children of the National Longitudinal Survey of Youth." Journal of Marriage and Family **53**: 1083-1110.
- Benabou, R. (1993). "Workings of a City: Location, Education and Production." Quarterly Journal of Economics **108**(3): 619-652.
- Benabou, R. (1994). "Education, Income Distribution, and Growth: The Local Connection." NBER Working Paper(4798): 1-34.
- Benasich, A. A. and J. Brooks-Gunn (1996). "Maternal Attitudes and Knowledge of Child-Rearing: Associations with Family and Child Outcomes." Child Development **67**(3): 1186-1205.
- Besharov, D. (1992). "A New Start for Head Start." American Enterprise **3**: 52-57.
- Betts, J. R. (1995). "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." Review of Economics and Statistics **77**: 231-250.
- Betts, J. R. (1996). *Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature. Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success.* G. Burtless. Washington, D.C., Brookings Institution Press: 141-191.

- Black, S. E. (1998). "Measuring the Value of Better Schools." Federal Reserve Bank of New York Economic Policy Review **4**(1): 87-94.
- Blau, D. M. (1997). "The Production of Quality in Child Care Centers." Journal of Human Resources **32**(2): 354-387.
- Blau, D. M. (1999). "The Effect of Income on Child Development." Review of Economics and Statistics **81**(2): 261-276.
- Blau, D. M. (forthcoming). "The Effect of Child Care Characteristics on Child Development." Journal of Human Resources.
- Blau, F. and A. J. Grossberg (1992). "Maternal Labor Supply and Children's Cognitive Development." The Review of Economics and Statistics **74**(3): 474-481.
- Bogart, W. T. and B. A. Cromwell (1997). "How Much More is a Good School District Worth?" National Tax Journal **50**: 215-32.
- Boggess, S. (1998). "Family Structure, Economic Status, and Educational Attainment." Journal of Population Economics.
- Boozer, M. and C. Rouse (1995). "Intraschool Variation in Class Size: Patterns and Implications." NBER Working Paper(5144).
- Borjas, G. (1992). "Ethnic Capital and Intergenerational Mobility." Quarterly Journal of Economics **107**(1): 123-150.
- Borjas, G. (1995). "Ethnicity, Neighborhoods, and Human Capital Externalities." American Economic Review **85**: 365-390.
- Borjas, G. J. (1998). "To Ghetto or Not to Ghetto: Ethnicity and Residential Segregation." Journal of Urban Economics **44**: 228-253.
- Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., Tellegen, A. (1990). "Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart." Science **250**: 223-50.
- Bouchard, T. J. J. (1994). "Genes, Environment and Personality." Science **264**: 1700-01.
- Bound, J. and G. Solon (1999). "Double Trouble: On the Value of Twins-Based Estimation of the Returns to Schooling." Economics of Education Review **18**: 159-68.
- Brewer, D. J., D. I. Rees, et al. (1995). "Detracking America's Schools: The Reform Without Cost?" Phi Delta Kappan: 210-215.
- Brooks-Gunn, J., G. J. Duncan, et al. (1993). "Do Neighborhoods Influence Child and Adolescent Development?" American Journal of Sociology **99**(2): 353-395.
- Brooks-Gunn, J., P. K. Klebanov, et al. (1996). "Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics." Child Development **67**: 396-408.
- Brown, B. B., N. Mounts, et al. (1993). "Parenting Practices and Peer Group Affiliation in Adolescents." Child Development **64**: 467-82.

- Brunner, E. and J. Sonstelie (1999). "School Finance Reform and Voluntary Fiscal Federalism." University of California - Santa Barabar working paper.
- Bryant, D. M., Burchinal, M., Lau, L. B., Sparling, J. J. (1994). "Family and Classroom Correlates of Head Start Children's Developmental Outcomes." Early Childhood Research Quarterly **9**: 289-309.
- Bryk, A. S. and M. E. Driscoll (1988). *The High School as Community: Contextual Influences and Consequences for Students and Teachers*, National Center on Effective Secondary Schools, University of Wisconsin, Madison.
- Burchinal, M. R., Ramey, S. I, Reid, M.K., Jaccard, J. (1995). "Early Child Care Experiences and Their Association with Family and Child Characteristics During Middle Childhood." Early Childhood Research Quarterly **10**: 33-61.
- Butcher, K. F. and A. Case (1994). "The Effects of Sibling Sex Composition on Women's Education and Earnings." Quarterly Journal of Economics **109**(3): 531-563.
- Cairns, R. B., M. Leung, et al. (1995). "Friendships and Social Networks in Childhood and Adolescence: Fluidity, Reliability and Interrelations." Child Development **66**: 1330-45.
- Caldas, S. J. and C. Bankston (1997). "Effect of School Population Socioeconomic Status on Individual Academic Achievement." Journal of Educational Research **90**(5): 269-277.
- Cameron, S. and J. Heckman (1999). "The Dynamics of Educational Attainment for Blacks, Hispanics and Whites." NBER working paper 7249.
- Card, D. and A. Krueger (1996). "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina." NBER Working Paper(5708).
- Card, D. and A. B. Krueger (1992). "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." Journal of Political Economy **100**(1): 1-40.
- Card, D. and A. B. Krueger (1996). *Labor Market Effects of School Quality: Theory and Evidence. Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. G. Burtless. Washington, DC, Brookings Institution Press.
- Carey, N., L. Lewis, et al. (1998). Parent Involvement in Children's Education: Efforts by Public Elementary Schools. Washington, D.C., U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.
- Case, A. and A. Deaton (1999). "School Inputs and Educational Outcomes in South Africa." Quarterly Journal of Economics **114**: 1047-84.
- Case, A. C. and L. F. Katz (1991). "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." NBER Working Paper(3705).
- Caughy, M. O., DiPietro, J.A., Strobino, D.M. (1994). "Day Care Participation as a Protective Factor in the Cognitive Development of Low-Income Children." Child Development **65**: 457-71.
- Chase-Lansdale, P. L., R. A. Gordon, et al. (1997). *Neighborhood and Family Influences on the Intellectual and Behavioral Competence of Preschool and Early School-Age Children*.

- Neighborhood Poverty. J. Brooks-Gunn, G. J. Duncan and J. L. Aber. New York, Russell Sage Foundation. **1**: 79-118.
- Cherlin, A., P. Chase-Lansdale, et al. (1998). "Effects of Parental Divorce on Mental Health Throughout the Life Course." American Sociological Review **63**: 239-49.
- Cherlin, A., K. Kiernan, et al. (1995). "Parental Divorce in Childhood and Demographic Outcomes in Young Adulthood." Demography **32**: 299-318.
- Cherlin, A. J., Furstenberg, F.F., Chase-Lansdale, P.L., Kiernan, K.E., Robins, P.K., Morrison, D.R., Teitler, J.Q. (1991). "Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States." Science **252**: 1386-89.
- Chipuer, H. M. and R. Plomin (1990). "LISREL Modelling: Genetic and Environmental Influences on IQ Revisited." Intelligence **14**: 11-29.
- Clark, W. A. V. (1991). "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model." Demography **28**(1): 1-19.
- Clarke-Stewart, K. A. (1991). "A Home is not a School: The Effects of Childcare on Children's Development." Journal of Social Issues **47**(2): 105-23.
- Coleman, J. S., E. Q. Campbell, et al. (1966). Equality of Educational Opportunity. Washington, D.C., U.S. Government Printing Office.
- Comer, J. P. and N. M. Haynes (1991). "Parent Involvement in Schools: An Ecological Approach." Elementary School Journal **91**(3): 271-277.
- Cook, T. D. (1984). What Have Black Children Gained Academically from School Integration?: Examination of the Meta-Analytic Evidence. School Desegregation and Black Achievement. T. Cook, D. Armor, R. Crainet al. Washington, D.C., National Institute of Education.
- Crane, J. (1991). "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." American Journal of Sociology **96**(5): 1226-1259.
- Crone, T. M. (1998). "House Prices and The Quality of Public Schools: What Are We Buying?" Federal Reserve Bank of Philadelphia Business Review: 3-14.
- Currie, J. and T. Duncan (1995). "Does Head Start Make a Difference?" American Economic Review **85**(3).
- Cutler, D. M. and E. L. Glaeser (1997). "Are Ghettos Good or Bad?" Quarterly Journal of Economics **112**(3): 827-872.
- Daniels, M., B. Devlin, et al. (1997). Of Genes and IQ. Intelligence, Genes, and Success: Scientists Respond to The Bell Curve. B. Devlin, S. Fienberg, D. Resnick and K. Roeder. New York, Springer Verlag: 45-70.
- Datcher, L. (1982). "Effects of Community and Family Background on Achievement." Review of Economics and Statistics **64**: 32-41.
- Densem, P. and B. Keeling (1988). "Family, Residential, and Community Indicators of Psychosocial Retardation." Mental Hcp NZ **11**: 4-26.

- DeParle, J. (1999). Project to Rescue Needy Stumbles Against the Persistence of Poverty. New York Times. New York: A1.
- Devlin, B., S. Fienberg, et al. (1997). Intelligence, Genes and Success: Scientists Respond to The Bell Curve. New York, Springer Verlag.
- Dodge, K. A., G. S. Pettit, et al. (1994). "Socialization Mediators of the Relation Between Socioeconomic Status and Child Conduct Problems." Child Development **65**: 649-65.
- Dornbusch, S. M., P. L. Ritter, et al. (1991). "Community Influences on the Relation of Family Statuses to Adolescent School Performance: Differences between African Americans and Non-Hispanic Whites." American Journal of Education: 543-567.
- Downes, T. A. and J. Horowitz (1995). "An Analysis of the Effect of Chicago School Reform on Student Performance." Federal Reserve Bank of New York Economic Perspectives **19**(3): 13-35.
- Downey, D. B. (1995). "When Bigger is Not Better: Family Size, Parental Resources and Children's Educational Performance." American Sociological Review **60**: 746-61.
- Downey, D. B., B. Powell, et al. (1999). "Much Ado About Siblings: Change Models, Sibship Size, and Intellectual Development." American Sociological Review **64**: 193-198.
- Dreeben, R. and R. Barr (1988). "Classroom Composition and the Design of Instruction." Sociology of Education **61**: 129-142.
- Duncan, G., W. Yeung, et al. (1998). "How Much Does Childhood Poverty Affect the Life Chances of Children." American Sociological Review **63**: 406-23.
- Duncan, G. J. (1994). "Families and Neighbors as Sources of Disadvantage in the Schooling Decisions of White and Black Adolescents." American Journal of Education **103**: 20-53.
- Duncan, G. J. and J. L. Aber (1997). Neighborhood Models and Measures. Neighborhood Poverty. J. Brooks-Gunn, G. J. Duncan and J. L. Aber. New York, Russell Sage Foundation. **1**: 62-78.
- Duncan, G. J. and J. Brooks-Gunn, Eds. (1997). Consequences of Growing up Poor. New York, Russel Sage Foundation.
- Duncan, G. J., J. Brooks-Gunn, et al. (1994). "Economic Deprivation and Early Childhood Development." Child Development **65**: 296-318.
- Duncan, G. J., J. P. Connell, et al. (1997). Conceptual and Methodological Issues in Estimating Causal Effects of Neighborhoods and Family Conditions on Individual Development. Neighborhood Poverty. J. Brooks-Gunn, G. J. Duncan and J. L. Aber. New York, Russell Sage Foundation. **1**: 219-250.
- Dunn, L. (1993). "Proximal and Distal Features of Day Care Quality and Children's Development." Early Childhood Research Quarterly **8**: 167-92.
- Eccles, J. S. and R. D. Harold (1993). "Parent-School Involvement during the Early Adolescent Years." Teachers College Record **94**(3): 568-587.
- Ensminger, M. E., R. P. Lamkin, et al. (1996). "School Leaving: A Longitudinal Perspective Including Neighborhood Effects." Child Development **67**: 2400-2416.

- Entwisle, D. R., K. L. Alexander, et al. (1994). "The Gender Gap in Math: Its Possible Origins in Neighborhood Effects." American Sociological Review **59**: 822-38.
- Epple, D., R. Filimon, et al. (1984). "Equilibrium Among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice." Journal of Public Economics **24**: 281-304.
- Epple, D., R. Filimon, et al. (1993). "Existence of Voting and Housing in a System of Communities with Property Taxes." Regional Science and Urban Economics **23**: 585-610.
- Epple, D. and T. Romer (1991). "Mobility and Redistribution." Journal of Political Economy **99**(4): 828-858.
- Epple, D. and H. Sieg (1999). "Estimating Equilibrium Models of Local Jurisdiction." Journal of Political Economy.
- Epstein, J. L. (1987). "Parent Involvement: What Research Says to Administrators." Education and Urban Society **19**(2): 119-136.
- Epstein, J. L. (1991). "Effects on Student Achievement of Teachers' Practices of Parent Involvement." Advances in Reading/Language Research **5**: 261-276.
- Epstein, J. L. (1992). School and Family Partnerships, Center on Families, Communities, Schools, and Children's Learning.
- Epstein, J. L. and S. L. Dauber (1991). "School Programs and Teacher Practices of Parent Involvement in Inner-City Elementary and Middle Schools." Elementary School Journal **91**: 291-305.
- Ernst, C., Angst, J. (1983). Birth Order: Its Influence on Personality. Berlin, Springer-Verlag.
- Evans, W. N., W. E. Oates, et al. (1992). "Measuring Peer Group Effects: A Study of Teenage Behavior." Journal of Political Economy **100**(5): 966-991.
- Falbo, T. and T. L. Poston (1993). "The Academic, Personality, and Physical Outcomes of Only Children in China." Child Development **64**: 18-35.
- Feldman, S. S., K. R. Wentzel, et al. (1990). "Marital Satisfaction of Parents of Preadolescent Boys and Its Relationship to Family and Child Functioning." Journal of Family Psychology **4**(2): 213-34.
- Ferguson, D., L. Horwood, et al. (1993). "Ethnicity, Social Background and Young Offending: A 14-Year Longitudinal Study." Australian and New Zealand Journal of Criminology **26**: 155-77.
- Ferguson, D., M. Lloyd, et al. (1991). "Family Ethnicity, Social Background and Scholastic Achievement -- An Eleven Year Longitudinal Study." New Zealand Journal of Educational Studies **26**: 49-63.
- Fernandez, R. and R. Rogerson (1996). "Income Distribution, Communities and the Quality of Public Education: A Policy Analysis." Quarterly Journal of Economics **111**: 135-164.
- Finn, J. D. (1998). Class Size and Students At Risk: What is Known? What is Next?, U.S. Department of Education.

- Fiske, E. and H. Ladd (1999). When Schools Compete. manuscript, Sanford School of Public Policy, Durham, NC.
- Flaks, D. K., I. Fichter, et al. (1995). "Lesbians Choosing Motherhood: A Comparative Study of Lesbian and Heterosexual Parents and Their Children." Developmental Psychology **31**: 105-14.
- Flynn, J. R. (1980). Race, IQ and Jensen. London, Routledge and Kegan Paul.
- Freese, J., B. Powell, et al. (1999). "Rebel Without a Cause or Effect: Birth Order and Social Attitudes." American Sociological Review **64**: 207-31.
- Fuller, B. and P. Clarke (1994). "Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy." Review of Educational Research **64**(1): 119-157.
- Furstenberg, F. F. and J. O. Teitler (1994). "Reconsidering the Effect of Marital Disruption: What Happens to Children of Divorce in Early Adulthood." Journal of Family Issues **15**(2): 173-90.
- Galambos, N. L. and J. L. Maggs (1991). "Out-of-School Care of Young Adolescents and Self-Reported Behavior." Developmental Psychology **27**: 644-55.
- Gallagher, J. J. (1995). "Comments on 'The Reform Without Cost?'" Phi Delta Kappan: 216-217.
- Gamoran, A. (1987). "The Stratification of High School Learning Opportunities." Sociology of Education **60**: 135-155.
- Gamoran, A. and M. Berends (1987). "The Effects of Stratification in Secondary Schools: Synthesis of Survey and Ethnographic Research." Review of Educational Research **57**: 415-436.
- Gamoran, A. and R. D. Mare (1989). "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?" American Journal of Sociology **94**(5): 1146-1183.
- Garasky, S. (1995). "The Effects of Family Structure on Educational Attainment: Do the Effects Vary by the Age of the Child." American Journal of Economics and Sociology **54**(1): 89-105.
- Garner, C. L. and S. W. Raudenbush (1991). "Neighborhood Effects on Educational Attainment: A Multilevel Analysis." Sociology of Education **64**: 251-262.
- Gaviria, A. and S. Raphael (1997). School-Based Peer Effects and Juvenile Behavior, University of California at San Diego.
- Gephart, M. A. (1997). Neighborhoods and Communities as Contexts for Development. Neighborhood Poverty. J. Brooks-Gunn, G. J. Duncan and J. L. Aber. New York, Russell Sage Foundation. **1**: 1-43.
- Geronimus, A. T. and S. Korenman (1994). "Does Young Maternal Age Adversely Affect Child Development? Evidence from Cousin Comparisons in the United States." Population and Development Review **20**: 585-609.

- Glewwe, P. (1997). "Estimating the Impact of Peer Group Effects on Socioeconomic Outcomes: Does the Distribution of Peer Group Characteristics Matter?" Economics of Education Review **16**(1): 39-43.
- Goldberger, A. (1979). "Heritability." Economia **46**: 327-47.
- Goldberger, A. and C. Manski (1995). "Review Article: The Bell Curve." Journal of Economic Literature **33**: 762-76.
- Golombok, S., R. Cook, et al. (1995). "Families Created by the New Reproductive Technologies: Quality of Parenting and Social and Emotional Development of the Children." Child Development **66**: 285-98.
- Graham, S. (1994). "Motivation in African Americans." Review of Educational Research **64**(1): 55-117.
- Greenstein, T. N. (1995). "Are the 'Most Advantaged' Children Truly Disadvantaged by Early Maternal Employment?" Journal of Family Issues **16**(2): 149-169.
- Greenwald, R., L. V. Hedges, et al. (1996). "The Effect of School Resources on Student Achievement." Review of Educational Research **66**(3): 361-396.
- Griffith, J. (1997). "Relation of Parental Involvement, Empowerment, and School Traits to Student Academic Performance." Journal of Educational Research **90**(1): 33-41.
- Guo, G. and L. K. VanWey (1999). "The Effects of Closely Spaced and Widely Spaced Sibship Size on Intellectual Development: Reply." American Sociological Review **64**: 199-206.
- Guo, G. and L. K. VanWey (1999). "Sibship and Intellectual Development: Is the Relationship Causal?" American Sociological Review **64**: 169-187.
- Hallinan, M. T. (1990). "The Effects of Ability Grouping in Secondary Schools: A Response to Slavin's Best-Evidence Synthesis." Review of Educational Research **60**(3): 501-504.
- Halpern-Felsher, B. L., J. P. Connell, et al. (1997). Neighborhood and Family Factors Predicting Educational Risk and Attainment in African American and White Children and Adolescents. Neighborhood Poverty. J. Brooks-Gunn, G. J. Duncan and J. L. Aber. New York, Russell Sage Foundation. **1**: 146-173.
- Hanushek, E. A. (1986). "The Economics of Schooling: Production and Efficiency in Public Schools." Journal of Economic Literature **24**(3): 1141-1177.
- Hanushek, E. A. (1992). "The Trade-Off Between Child Quality and Quantity." Journal of Political Economy **100**(1): 84-117.
- Hanushek, E. A. (1995). "Interpreting Recent Research on Schooling in Developing Countries." World Bank Research Observer **10**(2): 227-246.
- Hanushek, E. A. (1996). School Resource and Student Performance. Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success. G. Burtless. Washington, DC, Brookings Institution Press.

- Hanushek, E. A. (1997). "Assessing the Effect of School Resources on Student Performance: An Update." Educational Evaluation and Policy Analysis **19**(2): 141-164.
- Harris, J. R. (1995). "Where is the Child's Environment? A Group Socialization Theory of Development." Psychological Review **102**(3): 458-489.
- Harris, J. R. (1998). The Nurture Assumption. New York, The Free Press.
- Harvey, E. (1999). "Short-term and Long-term Effects of Early Parental Employment on Children of the National Longitudinal Survey of Youth." Developmental Psychology **35**(2): 445-59.
- Haskins, R. (1989). "Beyond Metaphor: The Efficacy of Early Childhood Education." American Psychologist **44**: 274-82.
- Hauser, R., Kuo, Hsiang-Hui Daphne (1997). "Does the Gender Composition of Sibships Affect Women's Educational Attainment?" The Journal of Human Resources **33**(3): 644-57.
- Haveman, R. and B. Wolfe (1994). Succeeding Generations: On the Effects of Investments in Children. New York, Russell Sage.
- Heckman, J. (1995). "Lessons from the Bell Curve." Journal of Political Economy **103**(5): 1091-1120.
- Heckman, J., A. Layne-Farrar, et al. (1996). Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship. Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success. G. Burtless. Washington, D.C., Brookings Institution Press: 192-289.
- Hedges, L. V. and R. Greenwald (1996). Have Times Changed? The Relation Between School Resources and Student Performance. Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success. G. Burtless. Washington, DC, Brookings Institution Press.
- Henderson, A. (1987). The Evidence Continues to Grow: Parent Involvement Improves Student Achievement. Columbia, Maryland, National Committee for Citizens in Education.
- Henderson, V., P. Mieszkowski, et al. (1978). "Peer Group Effects and Educational Production Functions." Journal of Public Economics **10**(1): 97-106.
- Herrnstein, R. J. and C. Murray (1994). The Bell Curve: Intelligence and Class Structure in American Life. New York, Free Press.
- Hidalgo, N. M., S.-F. Siu, et al. (1995). Research on Families, Schools, and Communities: A Multicultural Perspective. Handbook of Research on Multicultural Education. J. A. Banks. New York, Macmillan Publishing USA: 498-524.
- Hoffer, T. B. (1992). "Middle School Ability Grouping and Student Achievement in Science and Mathematics." Educational Evaluation and Policy Analysis **14**(3): 205-227.
- Hoover-Dempsey, K. V. and H. M. Sandler (1997). "Why Do Parents Become Involved in Their Children's Education?" Review of Educational Research **67**(1): 3-42.
- Hoxby, C. (1998). "Does Competition Among Public Schools Benefit Students and Taxpayers?" American Economic Review.

- Hoxby, C. (forthcoming). "Do Private Schools Provide Competition for Public Schools?" Journal of Political Economy.
- Hur, Y. M. and T. J. Bouchard (1995). "Genetic Influences on Perceptions of Childhood Family Environment: A Reared Apart Twin Study." Child Development **66**: 330-345.
- Huston, A., G. Duncan, et al. (1999). "Work-based Anti-Poverty Programs for Parents Can Enhance the School Performance and Social Behavior of Children." working paper.
- Iverson, B. K., G. D. Brownlee, et al. (1981). "Parent-Teacher Contacts and Student Learning." Journal of Educational Research **74**(6): 394-396.
- Jencks, C. (1992). Rethinking Social Policy: Race, Poverty and the Underclass. Cambridge, MA, Harvard University Press.
- Jencks, C. and S. E. Mayer (1990). The Social Consequences of Growing Up in a Poor Neighborhood. Inner-City Poverty in the United States. J. Laurence E. Lynn and M. G. H. McGeary. Washington, D.C., National Academy Press.
- Jensen, A. R. (1969). "How Much Can We Boost IQ and Scholastic Achievement." Harvard Educational Review **39**: 1-123.
- Kamerman, S. and A. Fahn (1997). Family Change and Family Policies in Great Britain, Canada, New Zealand and the U.S. Oxford, Clarendon Press.
- Katz, L., J. Kling, et al. (1999). "Moving to Opportunity in Boston: Early Impacts of a Housing Mobility Program." Harvard University working paper.
- Kerbow, D. and A. Bernhardt (1993). Parental Intervention in School: The Context of Minority Involvement. Parents, Their Children, and Schools. B. Schneider and J. S. Coleman. Boulder, Westview Press.
- Kessler, D. (1991). "Birth Order, Family Size, and Achievement: Family Structure and Wage Determination." Journal of Labor Economics **9**: 413-26.
- Kiernan, K. and J. Hobcraft (1997). "Parental Divorce during Childhood: Age at First Intercourse, Partnership and Parenthood." Population Studies **51**: 41-55.
- Kniesner, T., M. McElroy, et al. (1988). "Getting into Poverty without a Husband, and Getting Out, With or Without." American Economic Review **78**: 86-90.
- Knodel, J. and M. Wongsith (1991). "Family Size and Children's Education in Thailand: Evidence from a National Sample." Demography **28**(1): 119-131.
- Kontos, S. J. (1991). "Child Care Quality, Family Background and Children's Development." Early Childhood Research Quarterly **6**: 249-262.
- Korenman, S., J. E. Miller, et al. (1995). "Long-Term Poverty and Child Development in the United States: Results from the NLSY." Child and Youth Service Review **17**(12): 127-155.
- Krueger, A. (1999). "Experimental Estimates of Education Production Functions." The Quarterly Journal of Economics **114**(2): 497-532.

- Kuo, H.-H. D. and R. M. Hauser (1997). "How Does Size of Sibship Matter? Family Configuration and Family Effects on Educational Attainment." Social Science Research **26**: 69-94.
- Ladd, H. and J. Ludwig (1997). "Federal Housing Assistance, Residential Relocation, and Educational Opportunities: Evidence from Baltimore." American Economic Review **87**(2): 272-277.
- Lamborn, S., Mounts, N. , Steinberg, L., Dornbush, S. (1991). "Patterns of Competence and Adjustment Among Adolescents from Authoritative, Authoritarian, Indulgent and Neglectful Homes." Child Development **62**: 1049-65.
- Lareau, A. (1987). "Social Class Differences in Family-School Relationships: The Importance of Cultural Capital." Sociology of Education **60**: 73-85.
- Lareau, A. (1989). *Home Advantage: Social Class and Parental Intervention in Elementary Education*. London, The Falmer Press.
- Levin, H. M. (1997). "Raising School Productivity: An X-Efficiency Approach." Economics of Education Review **16**(3): 303-311.
- Levin, H. M. (1998). "Educational Vouchers: Effectiveness, Choice, and Costs." Journal of Policy Analysis and Management **17**(3): 373-392.
- Link, C. R. and J. G. Mulligan (1991). "Classmates' Effects on Black Student Achievement in Public School Classrooms." Economics of Education Review **10**(4): 297-310.
- Linney, J. A. and E. Vernberg (1983). *Changing Patterns of Parental Employment and the Family-School Relationship. Children of Working Parents: Experiences and Outcomes*. Washington, D.C., National Academy Press: 73-99.
- Lockheed, M. E., B. Fuller, et al. (1989). "Family Effects on Students' Achievement in Thailand and Malawi." Sociology of Education **62**: 239-256.
- Loehlin, J. C. (1992). Genes and Environment in Personality Development. Newbury Park, CA, Sage.
- Loehlin, J. C., J. M. Horn, et al. (1989). "Modeling IQ Change: Evidence from the Texas Adoption Project." Child Development **60**: 993-1004.
- Lou, Y., P. C. Abrami, et al. (1996). "Within-Class Grouping: A Meta-Analysis." Review of Educational Research **66**(4): 423-458.
- Ludwig, J., G. Duncan, et al. (1998). *Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment*, Georgetown Public Policy Institute.
- Mahard, R. E. and R. L. Crain (1983). *Research on Minority Achievement in Desegregated Schools. The Consequences of School Desegregation*. C. H. Rossell and W. D. Hawley. Philadelphia, Temple University Press.
- Manski, C. F. (1993). "Identification of Endogenous Social Effects: The Reflection Problem." Review of Economic Studies **60**: 531-542.
- Massey, D. S. and N. A. Denton (1987). "Trends in the Residential Segregation of Blacks, Hispanics, and Asians: 1970-1980." American Sociological Review **52**: 802-825.

- Mayer, S. E. (1991). How Much Does a High School's Racial and Socioeconomic Mix Affect Graduation and Teenage Fertility Rates. The Urban Underclass. C. Jenks and P. E. Peterson. Washington, D.C., The Brookings Institution: 321-41.
- Mayer, S. E. (1997). What Money Can't Buy: Family Income and Children's Life Chances. Cambridge, Mass., Harvard University Press.
- McCartney, K., Roberson, W. W., Jordan, E., Mouradian, V. (1991). Mother's Language with First and Second-Born Children: A Within-Family Study. Parent-Child Relations Throughout Life. K. Pillemer, McCartney, K. Hillsdale, Erlbaum: 125-142.
- McEwan, P. J. (1999). The Effectiveness of Religious and Non-religious Private Schooling: Evidence from Chile, Stanford University.
- McGue, M., Lykken, D. T. (1992). "Genetic Influence on Risk of Divorce." Psychological Science **3**: 368-373.
- McGue, M. (1993). From Proteins to Cognitions,: The Behavioral Genetics of Alcoholism. Nature, Nurture and Psychology. R. Plomin and G. E. McClearn. Washington, D.C., American Psychological Association: 244-268.
- McKinney, S. and A. B. Schnare (1989). "Trends in Residential Segregation by Race: 1960-1980." Journal of Urban Economics **26**: 269-280.
- McKinnon, R. and T. Nechyba (1997). .
- McLanahan, S. (1994). Growing Up with a Single Parent: What Hurts, What Helps. Cambridge, Harvard University Press.
- McLanahan, S. and K. Booth (1989). "Mother-Only Families: Problems, Prospects, and Politics." Journal of Marriage and Family **51**: 557-580.
- McLanahan, S. S. (1997). Parent Absence or Poverty: Which Matters More? Consequences of Growing up Poor. G. J. Duncan, Brooks-Gunn, J. New York, Russel Sage Foundation.
- McMillan, R. (1999). Parental Pressure and Competition: An Empirical Analysis of the Determinants of Public School Quality, Stanford University.
- Menaghan, E. G., L. Kowaleski-Jones, et al. (1997). "The Intergenerational Costs of Parental Social Stressors: Academic and Social Difficulties in Early Adolescence for Children of Young Mothers." Journal of Health and Social Behavior **38**(1): 72-86.
- Menaghan, E. G. and T. L. Parcel (1995). "Social Sources of Change in Children's Home Environments: The Effects of Parental Occupational Experiences and Family Conditions." Journal of Marriage and Family **57**: 69-84.
- Miller, P., C. Mulvey, et al. (1995). "What do Twin Studies Reveal about the Economic Returns to Education? A Comparison of Australian and U.S. findings." American Economic Review **85**: 586-599.
- Miller, P., C. Mulvey, et al. (1996). "Multiple Regression Analysis of the Occupational Status of Twins: A Comparison of Economic and Behavioral Genetic Models." Oxford Bulletin of Economics and Statistics **58**: 227-39.

- Miller, P., C. Mulvey, et al. (1997). "Family Characteristics and the Returns to Schooling: Evidence on Gender Differences from a Sample of Australian Twins." Economica **64**(253): 119-136.
- Ministry of Education (1993). The New Zealand Curriculum Framework. Wellington, NZ.
- Ministry of Education (1999). Annual Report on Maori Education 1997/98. Wellington, NZ.
- Ministry of Maori Development (1998). Maori Towards 2000. Wellington, NZ.
- Moffitt, R. A. (1998). Policy Interventions, Low-Level Equilibria and Social Interactions, Johns Hopkins University.
- Moreland, R. L. and J. M. Levine (1992). The Composition of Small Groups. Advances in Group Processes, JAI Press. **9**: 237-280.
- Mosteller, F. (1995). "The Tennessee Study of Class Size in the Early School Grades." THE Future of Children **5**(2): 113-27.
- Mulkey, L. M., R. L. Crain, et al. (1992). "One-Parent Households and Achievement: Economic and Behavioral Explanations of a Small Effect." Sociology of Education **65**: 48-65.
- Muller, C. (1993). Parent Involvement and Academic Achievement: An Analysis of Family Resources Available to the Child. Parents, Their Children, and Schools. B. Schneider and J. S. Coleman. Boulder, Westview Press.
- Muller, C. and D. Kerbow (1993). Parent Involvement in the Home, School, and Community. Parents, Their Children, and Schools. B. Schneider and J. S. Coleman. Boulder, Westview Press.
- Murray, C. (1984). Losing Ground : American Social Policy, 1950-1980. New York, Basic Books.
- Nechyba, T. (1996). "A Computational Model of Intergovernmental Aid." Journal of Public Economics **62**: 363-97.
- Nechyba, T. (1997). "Existence of Equilibrium and Stratification in Local and Hierarchical Tiebout Economies with Property Taxes and Voting." Economic Theory **10**: 277-304.
- Nechyba, T. (1999). "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers." Journal of Public Economic Theory **1**(1): 5-50.
- Nechyba, T. (1999). "Social Approval, Values and AFDC: A Re-Examination of the Illegitimacy Debate." NBER Working Paper.
- Nechyba, T. (forthcoming). "Mobility, Targeting and Private School Vouchers." American Economic Review.
- Nechyba, T. J. and R. P. Strauss (1998). "Community Choice and Local Public Services: A Discrete Choice Approach." Regional Science and Urban Economics **28**: 51-73.
- Neumark, D. (1999). "Biases in Twin Estimates of the Return to Schooling." Economics of Education Review **18**(143-48).

- New Zealand Ministry of Education (1999). *Strengthening Education in Mangere and Otara*. Wellington, NZ.
- Nisbett, R. E. (1997). *Race, Genetics and IQ. The Black-White Test Score Gap*. C. Jencks, Phillips, M. Washington, D.C., Brookings Institution Press.
- Ongley, P., S. Carson, et al. (1998). *Maori*. Wellington, NZ., Statistics New Zealand (Te Tari Tatau).
- Pallas, A. M., D. R. Entwisle, et al. (1994). "Ability-Group Effects: Instructional, Social, or Institutional?" *Sociology of Education* **67**: 27-46.
- Parcel, T. L. and E. G. Menaghan (1994). "Early Parental Work, Family Social Capital, and Early Childhood Outcomes." *American Journal of Sociology* **99**: 972-1009.
- Paulson, S. E. (1994). "Relations of Parenting Style and Parental Involvement with Ninth-Grade Students' Achievement." *Journal of Early Adolescence* **14**(2): 250-267.
- Pavuluri, M., S. Luk, et al. (1995). "A Community Study of Preschool Behavior Disorder in New Zealand." *Australian and New Zealand Journal of Psychiatry* **29**: 454-62.
- Pedersen, N. L., Plomin, R., Nesselroade, J. R., McClearn, G. E. (1992). "A Quantitative Genetic Analysis of Cognitive Abilities During the Second Half of the Life Span." *Psychological Science* **3**: 346-53.
- Peisner-Feinberg, E. S. and M. R. Burchinal (1997). "Relations between Preschool Children's Child-Care Experiences and Concurrent Development: The Cost, Quality and Outcomes Study." *Merrill-Palmer Quarterly* **43**(3): 451-77.
- Phelps, C. D. (1998). "Gender Differences in the Long-Term Economic Consequences of Parental Divorce." *Journal of Economic Behavior and Organization* **37**(2): 151-168.
- Phillips, M. (1999). "Sibship Size and Academic Achievement: What We Know and What We Still Need to Know." *American Sociological Review* **64**: 188-92.
- Pierce, K., J. Hamm, et al. (1999). "Experiences in After-School Programs and Children's Adjustment in First Grade Classrooms." *Child Development* **70**: 756-67.
- Plomin, R. (1994). *Genetics and Experience: The Interplay Between Nature and Nurture*. Newbury Park, Sage.
- Plomin, R., D. W. Fulker, et al. (1997). "Nature, Nurture and Cognitive Development from 1 to 16 Years: A Parent-Offspring Adoption Study." *Psychological Science* **8**: 442-47.
- Plomin, R. and S. Petrill (1997). "Intelligence: What's New?" *Intelligence* **24**: 53-77.
- Popkin, S. J., J. E. Rosenbaum, et al. (1993). "Labor Market Experiences of Low-income Black Women in Middle-Class Suburbs: Evidence from the Survey of Gautreaux Program Participants." *Journal of Policy Analysis and Management* **12**: 556-74.
- Posner, J. and D. Vandell (1999). "After-School Activities and the Development of Low-Income Urban Children: A Longitudinal Study." *Developmental Psychology* **35**: 868-79.
- Posner, J. K., Vandell, D. L. (1994). "Low-Income Children's After-School Care: Are there Beneficial Effects for After-School Programs?" *Child Development* **65**: 440-456.

- Powell, B. and L. C. Steelman (1995). "Feeling the Pinch: Child-Spacing and Constraints on Parental Economic Investments in Children." Social Forces **73**(4): 1465-1486.
- Purkey, S. C. and M. S. Smith (1983). "Effective Schools: A Review." Elementary School Journal **83**(4): 427-452.
- Rapaport, C. (1997). "Housing Demand and Community Choice: An Empirical Analysis." Journal of Urban Economics **42**: 243-260.
- Retherford, R. D., Sewell, W.H. (1991). "Birth Order and Intelligence: Further Tests of the Confluence Model." American Sociological Review **56**(2): 141-158.
- Reynolds, A. J. (1992). "Comparing Measures of Parental Involvement and Their Effects on Academic Achievement." Early Childhood Research Quarterly **7**: 441-462.
- Robertson, D. and J. Symons (1996). Do Peer Groups Matter? Peer Group versus Schooling Effects on Academic Attainment, London School of Economics Centre for Economic Performance.
- Rodgers, J. L., D. C. Rowe, et al. (1994). "DF Analysis of NLSY IQ/Achievement Data: Nonshared Environmental Influences." Intelligence **19**: 157-177.
- Rodman, H., D. J. Pratto, et al. (1985). "Child Care Arrangements and Children's Functioning: A Comparison of Self-Care and Adult-Care." Developmental Psychology **21**: 413-18.
- Roe, B. (1999). "Is There Competition between Breast-Feeding and Maternal Employment?" Demography **36**: 157-71.
- Rosenbaum, J. E., N. Fishman, et al. (1993). "Can the Kerner Commission's Housing Strategy Improve Employment, Education, and Social Integration for Low-Income Blacks?" North Carolina Law Review **71**: 1519-1556.
- Rosenzweig, M. R. (1990). "Population Growth and Human Capital Investments: Theory and Evidence." Journal of Political Economy **98**(5): S38-S70.
- Rouse, C. (1998). "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." Quarterly Journal of Economics **113**: 553-602.
- Rouse, C. (1999). "Further Estimates of the Economic Return to Schooling from a New Sample of Twins." Economics of Education Review **18**: 149-58.
- Rowe, D. C. (1994). *The Limits of Family Influence: Genes, Experience, and Behavior*. New York, Guilford Press.
- Rumberger, R. W. and J. D. Willms (1992). "The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools." Educational Evaluation and Policy Analysis **14**(4): 377-396.
- Sandefur, G. D., S. McLanahan, et al. (1992). "The Effects of Parental Marital Status during Adolescence on High School Graduation." Social Forces **71**(1): 103-121.
- Scarr, S. and R. Weinberg (1978). "The Influence of 'Family Background' on Intellectual Attainment." American Sociological Review **43**: 674-692.

- Schofield, J. W. (1995). Review of Research on School Desegregation's Impact on Elementary and Secondary School Students. Handbook of Research on Multicultural Education. J. A. Banks. New York, Macmillan Publishing USA: 597-616.
- Slavin, R. E. (1987). "Ability Grouping and Student Achievement in Elementary Schools: A Best-Evidence Synthesis." Review of Educational Research **57**(3): 293-336.
- Slavin, R. E. (1990). "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis." Review of Educational Research **60**(3): 471-499.
- Smith, A. (1998). Understanding Children's Development.
- Snyderman, M. and S. Rothman (1988). The IQ Controversy, the Media, and Public Policy. New Brunswick, NJ, Transaction Books.
- Sowell, T. (1994). Race and Culture. New York, Basic Books.
- Sowell, T. (1996). Migrations and Cultures. New York, Basic Books.
- Sowell, T. (1998). Conquests and Cultures. New York, Basic Books.
- St. John, N. H. (1975). School Desegregation: Outcomes for Children. New York, Wiley.
- Stanton-Salazar, R. D. and S. M. Dornbusch (1995). "Social Capital and the Reproduction of Inequality: Information Networks Among Mexican-origin High School Students." Sociology of Education **68**: 116-135.
- Steinberg, L. (1986). "Latchkey Children and Susceptibility to Peer Pressure: An Ecological Analysis." Developmental Psychology **22**: 433-439.
- Steinberg, L., Dornbusch, S. M., Brown, B.B. (1992). "Ethnic Differences in Adolescent Achievement: An Ecological Approach." American Psychologist **47**: 723-729.
- Stevenson, D. L. and D. P. Baker (1987). "The Family-School Relation and the Child's School Performance." Child Development **58**: 1348-1357.
- Sui-Chu, E. H. and J. D. Willms (1996). "Effects of Parental Involvement on Eighth-Grade Achievement." Sociology of Education **69**: 126-141.
- Sulloway, F. J. (1995). "Birth Order and Evolutionary Psychology: A Meta-Analytic Overview." Psychological Inquiry **6**: 75-80.
- Sulloway, F. J. (1997). Born to Rebel: Birth Order, Family Dynamics, and Creative Lives. New York, Pantheon.
- Summers, A. A. and B. L. Wolfe (1977). "Do Schools Make a Difference?" American Economic Review **67**: 639-652.
- Taubman, P. (1976). "The Determinants of Earnings: Genetics, Family, and Other Environments; A Study of White Male Twins." American Economic Review **66**: 858-70.
- Taubman, P. (1976). "Earnings, Education, Genetics and the Environment." Journal of Human Resources **11**: 447-61.

- Taubman, P. (1981). "On Heritability." Economica **48**: 417-20.
- Taufua, B. (1999). Discussion of Pacific Family and Community Issues, Manukau Youth Center.
- Teachman, J. D. (1996). "Intellectual Skill and Academic Performance: Do Families Bias the Relationship?" Sociology of Education **69**: 35-48.
- Thomas, G., M. P. Farrell, et al. (1996). "The Effects of Single-Mother Families and Nonresident Fathers on Delinquency and Substance Abuse in Black and White Adolescents." Journal of Marriage and the Family **58**: 884-894.
- Thompson, L. A., Detterman, D. K., Plomin, R. (1991). "Associations Between Cognitive Abilities and Scholastic Achievement: Genetic Overlap but Environmental Differences." Psychological Science **2**: 158-165.
- Thornton, Y. S. and J. Coudert (1995). A Ditchdigger's Daughters: A Black Family's Astonishing Success Story. New York, Birch Lane Press.
- Tienda, M. (1991). Poor People and Poor Places: Deciphering Neighborhood Effects on Poverty Outcomes. Macro-Micro Linkages in Sociology. J. Huber. Newbury Park, Calif., Sage Publications.
- Tooby, J. and L. Cosmides (1992). The Psychological Foundations of Culture. The Adapted Mind: Evolutionary Psychology and the Generation of Culture. J. Barkow, L. Cosmides and J. Tooby. Oxford, Oxford University Press.
- Vandell, D. L., Ramanan, Janaki (1991). "Children of the National Longitudinal Survey of Youth: Choices in After-School Care and Child Development." Developmental Psychology **27**(4): 637-643.
- Vandell, D. L. and M. A. Corasaniti (1988). "The Relation between Third Grader's After-School Care and Social, Academic, and Emotional Functioning." Child Development **59**: 868-875.
- Wadsworth, S. J. (1994). School Achievement. Nature and Nurture During Middle Childhood. J. C. DeFries, Plomin, R., Fulker, D.W. Cambridge, Blackwell.
- Webb, N. M. (1982). "Peer Interaction and Learning in Cooperative Small Groups." Journal of Educational Psychology **74**: 642-655.
- Whitbeck, L. B., R. L. Simons, et al. (1997). "The Effects of Parents' Working Conditions and Family Economic Hardship on Parenting Behaviors and Children's Self-Efficacy." Social Psychology Quarterly **60**(4): 291-303.
- Wilcox, W. (1998). "Conservative Protestant Childrearing: Authoritarian or Authoritative?" American Sociological Review **63**: 796-809.
- Willms, J. D. (1986). "Social Class Segregation and its Relationship to Pupils' Examination Results in Scotland." American Sociological Review **51**: 224-241.
- Wilson, J. Q. and R. Herrnstein (1985). Crime and Human Nature. New York, Simon and Schuster.

Winkler, D. R. (1975). "Educational Achievement and School Peer Group Composition." Journal of Human Resources **10**(2): 189-204.

Wojtkiewicz, R. A. (1993). "Duration in Parental Structures and High School Graduation." Sociological Perspectives **36**(4): 393-414.

Wood, D., Halfon, N., Scarlata, D., Newacheck, P., Nessim, S. (1993). "Impact of Family Relocation on Children's Growth Development, School Function and Behavior." Journal of the American Medical Association **270**: 1334-1338.

Wylie, C., J. Thompson, et al. (1999). *Competent Children at 8: Families, Early Education, and Schools*. Wellington, NZ, New Zealand Council for Educational Research.

Zak, P. J. (1999). *Genetics, Family Structure, and Economic Growth*, Claremont Graduate University.

Zellman, G. L. and J. M. Waterman (1998). "Understanding the Impact of Parent School Involvement on Children's Educational Outcomes." Journal of Educational Research **91**(6): 370-380.

Zigler, E. and S. Muenchow (1992). *Head Start: The Inside Story of America's Most Successful Educational Experiment*. New York, Basic Books.

Zimmer, R. W. and E. F. Toma (1997). *Peer Effects in Private and Public Schools: Across Country Empirical Analysis*, University of Kentucky.