Matched Sampling Methods To Reduce Bias In An Observational Study

by

Mark Gregory Duggan

B.S., Electrical Engineering
M.I.T.
(1992)

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

Author ....
Department of Electrical Engineering and Computer Science
May 9, 1994

Certified by ....
Arnold Barnett
Professor of Operations Research
Thesis Supervisor

Certified by ....
Robert M. Solow
Institute Professor of Economics
Thesis Supervisor

Accepted by ....
Frederic R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

Matched Sampling Methods to Reduce Bias

in an Observational Study

by

Mark Gregory Duggan

Submitted to the Department of Electrical Engineering and Computer Science
on May 9, 1994, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

Observational studies differ from randomized experiments in that the rule which governs the assignment of a treatment to individual units is not known. Because the control and experimental groups may differ systematically from one another, statistical adjustments must be made to ascribe the differences between the two groups to a particular treatment. Methods for drawing causal inferences from observational data are described in this document. Matched sampling techniques are employed to reduce bias between experimental and control groups in an effort to isolate the efffect of early childhood exposure to poverty on an individual's educational attainment. Experimental groups consisting of individuals who are initially poor and subsequently non-poor are compared with similar individuals who are consistently poor. This allows a first-pass estimate of the educational benefit a poor child would receive if his/her household income were lifted above the poverty line. Then, individuals who are temporarily poor early in life are compared with others who are similar but are never exposed to estimate the educational "cost" of being exposed to poverty. Finally, these results are compared with those obtained using traditional econometric techniques. The empirical results suggest that additional income provides a modest educational benefit to poor children and that individuals who are temporarily poor early in life do suffer a small educational cost.

Thesis Supervisor: Arnold Barnett
Title: Professor of Operations Research

Thesis Supervisor: Robert M. Solow
Title: Institute Professor of Economics

Acknowledgements

There are several individuals who have helped me immensely with my research over the last several months, and to whom I am very grateful.

While taking an intermediate macroeconomics with Robert Solow last spring, I began to develop an intense interest in economics. After I told him of my interests midway through the semester, he found an interesting research topic for me and agreed to be a part of my thesis committee. Shortly thereafter, he helped me to acquire a fellowship, which supported me throughout much of the year. Perhaps most importantly though, the conversations which I have had with Professor Solow and his genuine confidence in my abilities have helped me to believe that I will someday be able to contribute meaningfully to the field of economics. I cannot possibly thank him enough for all that he has done for me during the past year. His presence on my thesis committee is something which I will one day tell my grandchildren about.

Shortly after speaking with Professor Solow about this research, I met Arnold Barnett, an extremely popular professor about whom I'd heard so much during my years at M.I.T. "Arnie", as most of his students liked to call him, immediately struck me as a person with whom I wanted to work. Because he had done an enormous amount of research in applied mathematics and statistics, his background was particularly appropriate for helping me with sophisticated empirical analyses of large data sets. During the last year, I've been fortunate enough to TA a couple of Arnie's classes, and to meet with him on a regular basis. His words of encouragement and genuine belief in my abilities have meant a great deal to me, and have helped me to conquer many of my feelings of self-doubt. Despite my nagging sense during the past year that my research was not of a sufficiently high quality, Arnie helped me to believe, time and time again, that I was actually doing excellent work. I look forward to working with him again this summer and a lifelong friendship with him.

I am especially grateful to Harvard Professor Donald Rubin for his time and effort during the past year. Though he was not formally a member of my thesis committee, he regularly met with me to discuss my research. I have come to realize that the matched sampling methodology on which he has done so much research is an extremely powerful tool, and I am certain that my understanding of it will help me in my future research as an economist. I'm very pleased that I will be at Harvard for the next several years, for we will have many more chances to talk during that time.

I am also very thankful for the financial support I have received during the last several months from the William T. Grant Foundation. This support enabled me to concentrate on my research and conduct a rigorous empirical analysis of the effect of childhood exposure to poverty on a person's educational attainment.

Thanks to the generosity of University of Wisconsin Professors Robert Haveman and Barbara Wolfe and their graduate student, Kathy Wilson, I did not need to spend an enormous amount of time extracting the relevant data for my research from the Panel Study on Income Dynamics. Their hard work in filtering out information about hundreds of children from the Panel Study on Income Dynamics enabled me to start immediately on the empirical analysis.

During my six years at M.I.T., my family has been a vital source of emotional support for me. Though it's been extraordinarily difficult here, they have instilled a belief in me that I can handle any challenge, regardless of how formidable it may be. Throughout their adult lives, my parents have repeatedly sacrificed to provide as much as they possibly could for my two brothers, my sister, and me. I can't begin to thank them enough for all that they have done.

Contents

Chapter 1

Introduction

During the mid-1960s, the United States launched a War on Poverty which was intended to "provide a hand up and not a handout" to the nation's most impoverished citizens.[11, p. 10] Experts from a variety of disciplines were called in to define poverty, assess the extent of poverty in the nation, and design programs to combat it. To evaluate the effectiveness of these programs, large-scale data collection efforts were started for the first time. These data sets enabled researchers to rigorously investigate many important issues in the following years.

Despite the extraordinary growth in antipoverty programs in the United States since the start of the War on Poverty, the incidence of poverty in the population has remained stubbornly high. The composition of this impoverished population has changed substantially during the last few decades. Children make up a disproportionately large share of impoverished citizens, with almost 22% of those individuals under the age of eighteen living below the poverty line (as officially calculated by the Census Bureau) in 1992. This percentage is the largest since poverty was first measured in 1965. [18, p. A32]

Many antipoverty programs were intended to improve the opportunities open to poor children so that they might escape from poverty and become productive members of society. Implicit in this effort was the belief that childhood exposure to poverty had a pernicious effect on a child's health, emotional development, educational attainment, and subsequent economic productivity. Many also seemed to feel that all of the nation's children deserved a fair start, and that impoverished children should be compensated somehow for the myriad disadvantages with which they had to deal.

In this document, I will investigate the effect which childhood exposure to poverty has on an individual's educational attainment. More specifically, I will employ multivariate matched sampling methods to estimate the effect that an income infusion to an impoverished family could have on the total years of schooling a child in that family receives. For the purposes of this analysis, I will use the *Panel Study on Income Dynamics*, a large-scale, longitudinal database which was started in 1968 to evaluate the effectiveness of nascent antipoverty programs. This database has extensive information about thousands of households and each individual within these households, and is a particularly rich set of information because of its longitudinal nature.

When a family's income is lifted above the poverty line, children within that household may subsequently receive more nutritious food, better health care and housing, and more time with their parents. This may provide an environment more conducive to intellectual development.[28, p. 370] On the other hand, income transfers of this kind may reduce the incentive of a child to improve his/her economic position, which is to say that a child may be less inclined to continue with school beyond a certain point.[1, p. 71] Of course, there are many more possible mechanisms through which family income during childhood can affect a child's educational attainment.

The empirical results described in the upcoming chapters suggest that childhood exposure to poverty does indeed have an adverse effect on the academic achievement of children. More specifically, individuals who are initially poor and whose household incomes subsequently rise *and remain* above the poverty line tend to remain in school for a longer period of time than do similar individuals whose income remains below the poverty line. The estimates of the educational "cost" of continued exposure to poverty are not statistically significant, perhaps because of the relatively small number of individuals whose longitudinal income patterns satisfy the criteria whose effect we estimate. Nevertheless, all of the best estimates suggest that the rise in income does benefit those children who were initially poor. Also, when comparing individuals who are initially poor and subsequently non-poor with similar individuals whose income is consistently above the poverty line, one finds an educational "cost" incurred by those individuals who are temporarily exposed to poverty.

Chapter 2

Preliminaries

To address meaningfully the effect childhood exposure to poverty has on an individual's educational attainment, one must first construct a reasonable definition of poverty. An official definition was prepared for the Social Security Administration in 1965 by Mollie Orshansky.[47, p. 1075] For each household, the cost of a diet which the Department of Agriculture deemed "nutritionally adequate" was multiplied by three. This multiplier was selected because of a 1955 survey which suggested that approximately 35% of the average household's aftertax income was spent on food. Implicit in this calculation is the assumption that impoverished families should not have to devote a larger fraction of their income to food than a typical family. To account for diseconomies of small scale, a slightly higher multiplier is used in calculating the poverty line for one and two person households. The poverty thresholds are recalculated each year by indexing them with the Consumer Price Index (CPI). Therefore, if the CPI rises by 5% in a given year, then all of the poverty thresholds are increased by 5% after that year. These poverty thresholds adjust for the number of individuals in the household, the age of the head of the household, and the number of individuals under the age of eighteen in the household.[13, p. 34] For many people, this attempt to define poverty symbolized the nation's newfound commitment to raising the living standards of its poorest citizens.

Unfortunately, this "official" measure is one of an infinite number of possible methods of calculation, none of which are unambiguously superior to the others. In defining the poor population, the Census Bureau has specified which resources will be considered as income, has selected the household as the most appropriate income-sharing unit, and has chosen one year as the accounting period. As is the case with many areas, the most appropriate poverty measure is likely to depend on the question at hand. Some of the most contentious issues in the definition of poverty include the treatment of health care, in-kind transfers, wealth holdings, and geographic location. In the words of Harold Watts, "our official measures are not grounded in some self-evident principle or expert consensus but are simply a collection of more or less arbitrary and eminently vulnerable rules."[49, p. 30] Despite the imperfections of the current poverty measures, the majority of the literature has used the official definition and I will adopt this convention in the following pages, too.

2.1 Measuring Educational Attainment

During the last few decades, American educationists have become increasingly concerned with the quality of schooling which America's children are receiving. Numerous measures of a child's educational attainment exist. Standardized test scores, years of education, academic grades, and attendance records are just a few yardsticks by which we can compare the educational attainment of different children. The years of education measure will be used throughout this document. This is mainly because of the unavailability of the other information in the *Panel Study on Income Dynamics.* There are, of course, obvious disadvantages to focusing on this measure. For example, this measure does not consider the performance level or work ethic of the individual while he/she was in school. Additionally, it does not account for substantial differences in the quality of schools throughout the country.

Despite these and other disadvantages, concentrating on an individual's years of education is defensible. It clearly distinguishes those individuals who have graduated from high school from those who have not. It points out whether an individual attended college and, if so, whether or not he/she finished. Because educational attainment is such an important part of a person's human capital, it is extremely useful in estimating a person's future earnings.[23] For example, college graduates earn, on average, 77% more than high school graduates, and this disparity in incomes is increasing. Education seems to have become a proxy for skills in the American workforce, and those individuals who do not make the effort to educate themselves may be in for a life of stagnant real income.[15, p. 27]

2.2 Compensatory Education Versus Income-Transfer Programs

The level of income inequality is far greater in the United States than it is in most industrialized nations.[19] No two people are likely to agree on the optimal level of inequality. Despite this, many policymakers believe that those individuals who are born into poverty should be given the same opportunities to achieve as those who are not. This is a formidable challenge indeed, and one that is virtually impossible to achieve in any practical way. Successful parents will tend to pass on their successes to their children, whereas unsuccessful parents are likely to pass on some of their failures to their children.[22, p. 138] A variety of antipoverty programs were designed in an effort to give people of low socioeconomic status a chance at escaping from poverty. Two general ways to approach this problem are discussed below.

2.2.1 Educational Programs for Impoverished Children

In the spirit of equalizing opportunities as opposed to outcomes, many policymakers involved in the War on Poverty felt that education was the best available route to overcoming poverty. As a result, dozens of compensatory education programs were created in the mid-1960s. The hope was that impoverished children would improve their economic circumstances, and become more productive members of society. Though the effectiveness of these programs was not overwhelmingly impressive, there seemed to be, in the aggregate, a small positive effect on the academic performance of disadvantaged students.[11, p. 172] Programs like Head Start were created to provide both short and long-term benefits to impoverished youth, but the short-term gains were rarely sustained. These programs certainly did not eliminate disparities in educational achievement between poor and non-poor children, but they did help to reduce the gap.

A study completed in 1977 assessed the lasting effects of preschool intervention programs on the long-term academic achievement of low-income children.[26] Ninety-two percent of the children were black and forty percent had no father at home. By choosing experimental groups who had participated in these programs and control groups who had not, researchers found that some substantial differences existed between experimental and control individuals ten years after the completion of the programs. Though there were no long-term differences in achievement scores or IQ, there were large differences in the extent to which the participating children were retained in grade (held back a grade) or assigned to special education classes. The median rate of failure, defined as falling behind by a grade or being assigned to a special education class, was forty-five percent in the control groups and twenty-four percent in the experimental groups.[11, p. 157] A wide variety of preschool programs were evaluated, including Head Start curricula, traditional nursery preschool, and language and cognitive development programs. Though the results of this study were indeed modest, they do suggest that some long-term benefits can be gained from intervention programs.

A more comprehensive study, the Sustaining Effects Study of Compensatory and Elementary Education, was mandated by Congress in 1975 to investigate the effect of Title I services on poor children. The primary focus was a three-year longitudinal study of children who received Title I services and children who were eligible for, but did not receive, Title I services (defined here as needy children). Approximately 120,000 students, drawn from a representative sample of 300 schools, were tested. Nearly 60% of poor children received Title I services. The children receiving these services lived in both large cities and rural areas. Statistical analysis showed that Title I students had significant gains relative to needy students for the Mathematics portion of the Comprehensive Tests of Basic Skills. These results held for children in grades one through six. Significant gains on the reading portion were found for children in grades one through three, but not for those in grades four, five, or six.[7] The study found that the programs were less effective for older children than for younger ones. By the time the Title I students had reached junior high, though, no sustained effects were observed. The bulk of the literature suggests that, if greater resources were indeed available, they would be most usefully spent at the preschool and elementary school level.[11, p. 160]

Compensatory education programs confront the education-poverty relationship by leaving the poverty untouched and focusing on learning opportunities. In the past, education and not income transfers have been the preferred instrument for dealing with the needs of disadvantaged children.[23, p. 14] In other words, these programs were created to improve disadvantaged children's education, in the hope that this would have economic benefits to them later on. Here, I try to assess the effect which an improvement in a poor child's economic position would have on his/her educational performance.

2.2.2 The Effect of Income Transfers

The architects of the War on Poverty and the American public in general have tended to prefer, in principle, providing opportunities rather than handouts in their efforts to ameliorate poverty.[11, p. 15] Despite this, most of the increases in public spending since 1965 have taken the form of income or in-kind transfers. These programs have actually been far more successful in helping to reduce poverty than have compensatory education programs, as they have elevated many households' incomes above their respective poverty thresholds. Unfortunately, many households remain dependent on these welfare payments, and the percentage of hard-core poor in the country is not declining.

One important question to ask about these income redistribution measures is: what effect do they have on the children within these households? More specifically, do children tend to achieve more academically if the households in which they reside are lifted out of poverty? Researchers from many different disciplines have examined the determinants of a child's educational development. One conclusion which these studies have in common is that a child's home environment is an immensely important determinant of his/her educational attainment.[9] Factors which seem to be correlated with a child's educational attainment include the parents' educational level, the family's income, the family's socioeconomic status, durable goods ownership, proper nutrition and health care, and adequate housing.

Though there is a strong positive correlation between family income during an individual's childhood and his/her educational attainment, it is not clear that there is a causal relationship between the two. A family's income depends on the parents' abilities, motivation, and attitudes. These may be the actual causal determinants of a child's academic achievement, not income per se. If this is indeed the case, then providing an income infusion to an impoverished family without changing the attitudes or abilities of the parents may not lead to improved educational attainment by the children.[29, p. 465] On the other hand, if the increased income leads to a home environment more conducive to learning, then these extra dollars may be a boon to the child's intellectual development.

Making reasonable inferences about the effects of an income infusion on a child's educational attainment from observational data is exceedingly difficult. For this reason, the study described in the next section is quite attractive, because its inferences were based on a relatively well-controlled randomized experiment.

2.2.3 The Effect of a Negative Income Tax on the School Performance of Impoverished Children

In the mid-1970s, an experiment was designed to determine the effect of a negative income tax program on the education of children. Known as the Rural Income Maintenance Experiment, a sample of 847 children from North Carolina and Iowa were used in the subsequent analysis. These children were members of families who participated in the negative income tax experiment and for whom pre- and post-enrollment data are available. The sample is by no means representative of the nation's population of schoolchildren but rather of an "intellectually impoverished population." [28, p. 372] Generally speaking, the children tend to have larger families, lower incomes, less educated parents and be much more at risk of school failure than an average American child.

Despite the amount of time which has elapsed since this analysis was conducted (in the mid-1970s), it deserves special attention because it is based on data from a controlled experiment and it analyzes the effect of an income infusion on impoverished children's education. For the negative income tax program, the two relevant parameters are G, the guaranteed annual income and t, the income tax rate. If a family's income falls below some threshold level, then the family is provided with an income subsidy which depends on the parameters and the family's income. For those families whose incomes fall below the threshold value, the expected effects were (1) an increase in the family's total income and (2) a reduction in the parents' labor force participation. The expected result was an improvement in the children's school performance because of the effects of the negative income tax on the parents' time and income allocations.[28, p. 371]

Four different measures of educational attainment, the child's attendance record, the comportment grade point average (a behavioral measure), the academic grade point average, and a standardized test score, were used to evaluate the effect of the NIT program. Pre- and post-enrollment performance for the children were compared, and there was no non-participating control group. The results of the experiment were mixed. The most significant responses were found in the performance of the second through eighth grade North Carolina schoolchildren. These children experienced a 30.5% reduction in absenteeism, a 6.7% increase in comportment GPA, a 6.2% increase in academic GPA, and an 18.9% reduction in the gap between achievement test scores and the corresponding expected grade equivalent score. All four of the experimental effects were statistically significant for the second through eighth grade schoolchildren. However, the older children in the North Carolina sample did not exhibit any significant experimental responses, which may indicate that the behavior of the younger children is much easier to modify. Finally, the sample of schoolchildren from Iowa provided no support for the hypothesis that a negative income tax will result in improved school performance by impoverished children. Maynard contends that the Iowa schoolchildren were better performers prior to the experiment than their North Carolina counterparts and that the quality of the Iowa school environment data was not nearly as good as the North Carolina data.

The results of this study suggest that a negative income tax may significantly improve the academic achievement of impoverished youth and that elementary schoolchildren are more likely to benefit from the NIT program than individuals enrolled in high school are. Though the results of the study are mixed, this Rural Income Maintenance Experiment is a reasonable way to investigate the causal effects of a negative income tax on the educational attainment of impoverished youth.

2.3 What's To Come

Though scores of compensatory education and income-transfer programs have been established since the United States embarked on the War on Poverty three decades ago, the percentage of hard-core poor in the country remains high. Approximately 7% of the nation's citizens live in households which have yearly incomes below the poverty line more than 80% of the time. This structural poverty problem may persist if the costs of an educational transition for the nation's poorest people remains prohibitively high.[1, p. 70] Possible reasons for high rates include an inability to forgo income while investing in education, lenders' biases against impoverished people due to their seeming lack of creditworthiness, myopia on the part of the poor, and a considerable penalty for noncompletion of the degree. If poor families are permanently lifted above the poverty line by means of a negative income tax or an income subsidy, this could lead to substantial increases in the length of time that their children spend in school.

How much would it benefit poverty-exposed children if the government were permanently to lift their incomes above the poverty line? How would any benefits attenuate as the years of childhood exposure to poverty went up? In this document, I will attempt to address these issues using multivariate matched sampling techniques by filtering out relevant data from the *Panel Study on Income Dynamics* to isolate "experiments" which have been conducted by nature. This approach will allow me to make some first-pass estimates as to the effect of childhood exposure to poverty on an individual's educational attainment.

Chapter 3

The Panel Study of Income Dynamics

Shortly after President Johnson's War on Poverty was launched in the mid-1960s, the U.S. Bureau of the Census was asked by the Office of Economic Opportunity (OEO) to assess the extent of poverty in the country and the effectiveness of the new antipoverty programs. As a result, a large-scale census study, known as the Survey of Economic Opportunity (SEO), finished interviews with approximately 30,000 households in 1966 and 1967. Realizing the abundance of information which a study like this one could provide, the OEO approached the Survey Research Center at the University of Michigan and asked the SRC to continue this study of the nation's overall economic well-being. Because the study's primary objective was to investigate the dynamics of poverty, the SEO wanted intensive interviews to be conducted with 2,000 low-income households from the original national sample of 30,000 households. Researchers at the University of Michigan argued that a randomly selected cross-section of the original 30,000 should also be included in the study, so that the sample would continue to be representative of the nation's individuals and families.

For this study, the *Panel Study of Income Dynamics*, 4,802 household interviews were successfully conducted in 1968. Of these, 1,872 were low-income households from the SEO while 2,930 were selected from the SRC national sample. The PSID continues to this day, and has become one of the most frequently used and influential data sets for research in the social sciences. Through its annual interviews, the PSID obtains extensive information about families and the individuals who make up those families. The data provide substantial detail about employment, income, education, and family composition for each of the households interviewed. The rules for following household members since the PSID's inception were created to maintain a sample of families which was representative of the population. New PSID families were created when children grow up and establish their own households or when married partners go their separate ways. Compensatory weights for each family are included to adjust for unequal selection probabilities and the variation in attrition rates between different socioeconomic groups. Because those families which drop out of the study may differ systematically from those which remain, these weights do not remove all of the bias due to attrition. Several studies have provided reassuring evidence that there is not substantial nonresponse bias in the PSID,[16] though, which makes national estimates calculated using the probability-of-selection weights all the more reasonable.[20, p. 25]

3.1 The PSID Children

Thanks to the generosity of Robert Haveman and Barbara Wolfe from the University of Wisconsin, a filtered portion of the PSID data was made available for the purposes of this analysis. This sample of the data includes extensive information about the 1705 original PSID members (i.e. since 1968) who were born sometime between 1962 and 1968, which is to say that the children were zero through six years old when the PSID started. These individuals were selected primarily because they have been PSID members since the start of the study in 1968; they are now adults and there exists an abundance of longitudinal information about them. Ideally, the filtered portion would consist only of people who were the same age in 1968 (i.e. all of the individuals born to PSID families in 1968). Unfortunately, with this constraint, the number of individuals drops to approximately 350, which is not enough people to allow the kind of rigorous analysis which Haveman and Wolfe have been doing.

Longitudinal income, family, and geographic information is available for those years during which the individuals were six through fifteen years old. The lower end of the age range was selected primarily because this represents the lowest age for which longitudinal information is available for all 1705 people. Havemann and Wolfe's selection of fifteen as the upper end is somewhat more arbitrary, though it does seem reasonable to choose sixteen as the age at which a child begins to become independent, and therefore more of an adult.

3.2 The Available Information About Each of the 1705 Children

Extensive information is available for each of the 1705 individuals, including data about each individual, his or her parents and family, the kind of home environment in which the child lived, his/her educational attainment, and the income and poverty status of the individual's household. In all, there are nearly fifty pieces of information for each individual, and several of the most important ones are described below.

3.2.1 Background Information about the Individual

When the PSID was launched in 1968, policymakers were particularly interested in the dynamics of poverty. As a result, the sample of 1705 individuals includes a disproportionately large number of households with incomes below the poverty line. This oversampling of impoverished families resulted in a large subsample of black households, with nearly half of the 1705 individuals regarded as nonwhite in the PSID data. For each individual, the *NONWHITE* variable takes on a value of one if the person is black or hispanic, and zero otherwise. Unweighted and weighted distributions for the *NONWHITE* variable are given in figure 3-1. Assuming that the weighted distribution truly is representative of the nation's children who were six and under in 1968, the fraction of this population which was nonwhite is roughly 16.7%. The *FEMALE* variable reveals the gender of each of the sampled individuals, while *FIRSTBORN* takes on a value of one if the person was the firstborn child in his/her family and zero otherwise. Finally, *YEARSED* reveals the years of education for each individual. If this variable takes on a value greater than eleven, then the individual did graduate from high school. The weighted distribution for the *YEARSED* variable is shown in figure 3-2. Weighted and unweighted statistics for these four variables are provided in the following table.

Variable	Unweighted Statistic	Weighted Statistic
NONWHITE	48.0%	16.7%
FIRSTBORN	22.5%	28.1%
FEMALE	51.2%	50.5%
YEARSED	12.614 yrs	12.905 yrs

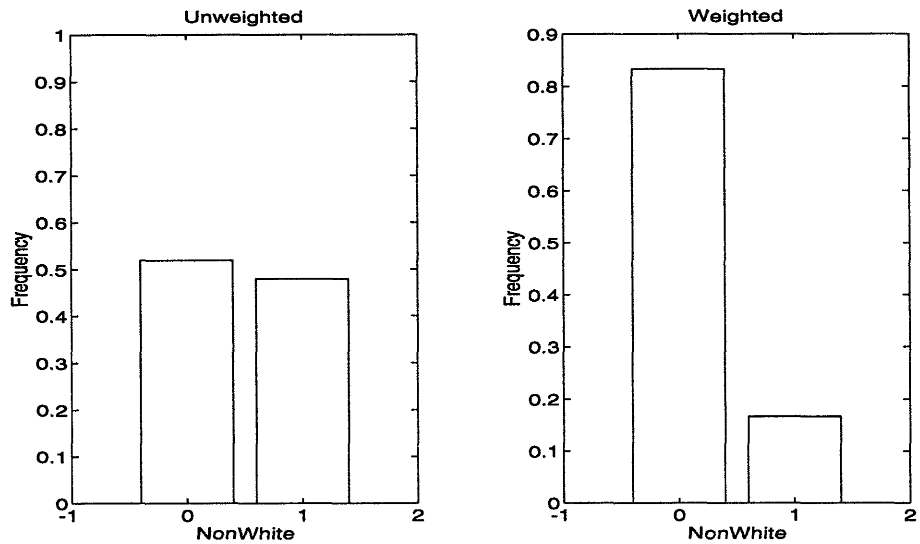

Figure 3-1: Weighted and Unweighted Distributions for *NONWHITE*

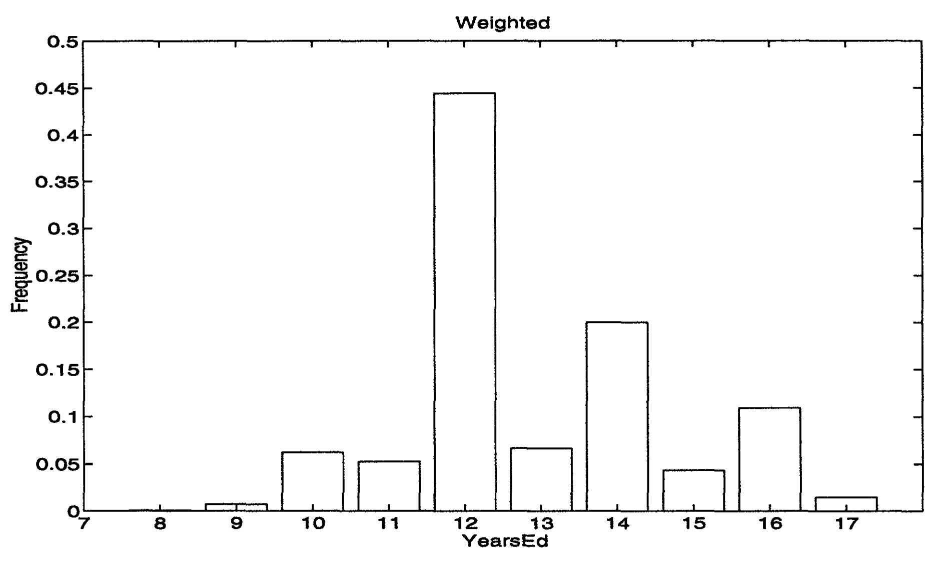

Figure 3-2: Weighted Distribution for *YEARSED*

The disparity in the weighted and unweighted average for the YEARSED and FIRSTBORN individuals is due to the overrepresentation of poor children in the sample of 1705 children. Because poor children are more likely to have many siblings, one would expect a smaller percentage of them to be firstborn children. Thus, it makes intuitive sense that the weighted estimate for this variable is larger than the unweighted one. Similarly, because there is a positive correlation between income and educational attainment, one might anticipate that the sample mean for YEARSED would be less than the estimate for the national average.

3.2.2 Background Information About the Individual's Family

The nature of a child's home environment is believed to have a significant impact on his/her educational attainment. [9] Effects of the family must therefore be taken into account if one hopes reliably to assess the effect of poverty on an individual's academic achievement. One potentially important determinant of the quality of a child's home environment is the number of parents living at home. For each individual, the number of biological parents living at home with him/her in 1968 are included. Weighted and unweighted statistics for this variable are included in the following table.

Number of Parents in 1968	Unweighted Statistic	Weighted Statistic
Two	78.0%	90.3%
One	18.9%	8.6%
Zero	3.1%	1.1%

Given the greater likelihood that a poor child will have fewer than two parents at home, it makes sense that the proportion of one-parent and zero-parent households is greater for the sample than for the national average.

Information concerning the educational attainment of each individual's parents are included in the data set. Numerous studies have shown that children whose parents are educated are more likely to do well than those whose parents are not. Using four categories of educational attainment, the variables MOM_YRS and DAD_YRS reveal whether each parent graduated from high school, attended some college, or finished a four-year college degree.

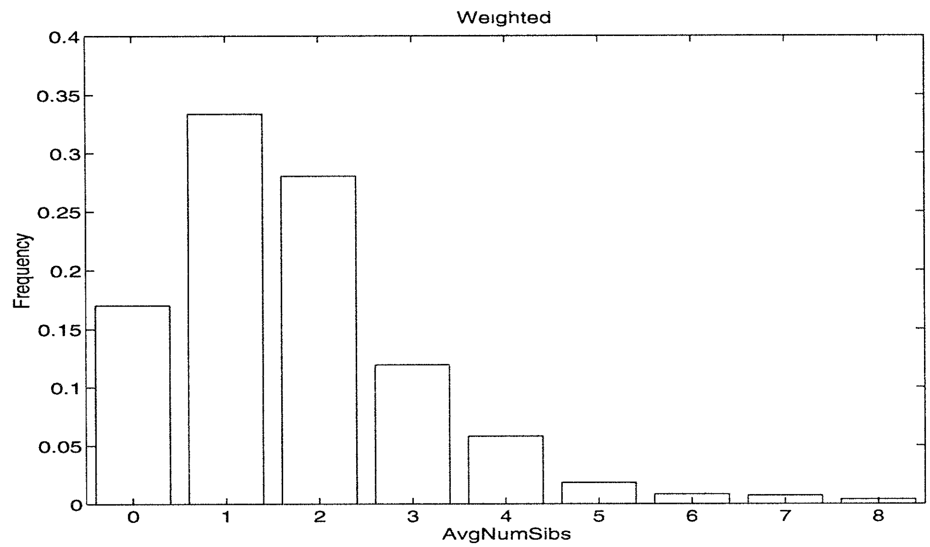

Figure 3-3: Weighted Distribution for *AVG_NUM_SIBS*

$$MOM\_YRS = \begin{cases} 1 & \text{if Mother did not graduate from high school} \\ 2 & \text{if Mother graduated from high school} \\ 3 & \text{if Mother attended college but didn't finish four year degree} \\ 4 & \text{if Mother finished a four year degree or more} \end{cases}$$

A similar definition applies for DAD_YRS. Weighted statistics for these two variables are provided in the following table.

DAD_YRS	Weighted	MOM_YRS	Weighted
1	39.4%	1	34.2%
2	28.5%	2	47.2%
3	15.6%	3	10.7%
4	16.5%	4	7.9%

The number of siblings a child has while growing up may also play a role in determining his/her educational attainment. The variable AVG_NUM_SIBS reveals the average number of siblings each child had over the ten-year period we are considering. Figure 3-3 gives the weighted distribution for this variable. For AVG_NUM_SIBS, the sample mean is 2.52 whereas the estimate for the population mean is 2.07.

3.2.3 Home Environment and Geographic Information

Several additional variables are included in the filtered portion of the PSID to provide more information about the environment in which the sample children grew up. NUM_SEPARATIONS gives the number of separations which each individual's parents had during the ten-year period under consideration. For this variable, the sample and population means were 0.24 and 0.23 respectively, and the unweighted and weighted distributions were not strikingly different from one another. YRS_HEAD_DISABLED provides the number of years during which the head of the individual's household was disabled. For this variable, the sample and population distributions were quite dissimilar, reflecting the relatively high probability that a poor person would be disabled. The sample and population means were 1.56 and 1.03, respectively.

The number of years which each individual spent living with one parent is represented by the variable YRS_WITH_ONE. If it is the case that a child was living with no biological parents in a particular year, then this is still considered living with one parent, mainly because the child must have lived with some guardian during that time. The sample and population means for this variable are 2.79 and 1.65 respectively, and the weighted distribution for YRS_WITH_ONE is shown in figure 3-4. YRS_MOM_WK reveals the number of years that the individual's mother worked, with sample and population means of 5.74 and 5.75, respectively. Finally, information is provided about the number of years each individual spent in the South during the ten year period. YRS_SOUTH has a sample mean of 4.56 years and a population mean of 2.78 years.

3.2.4 Income and Poverty Information

Longitudinal income information is available for each of the 1705 individuals during the ten years of interest. As opposed to giving actual income information though, the longitudinal data contains the ratio of the household income to the household poverty line, also known as the income-to-needs ratio. The poverty line used is the official poverty measure described in the previous chapter and depends on the size of the family. The AV_INC_NEEDS variable provides the average of the ten income-to-needs ratios of interest. Of the 1705 individuals in the study, 314 have an average income-to-needs ratio during these ten years which is below one. Using the compensatory weights, the average percentage of the population which is,

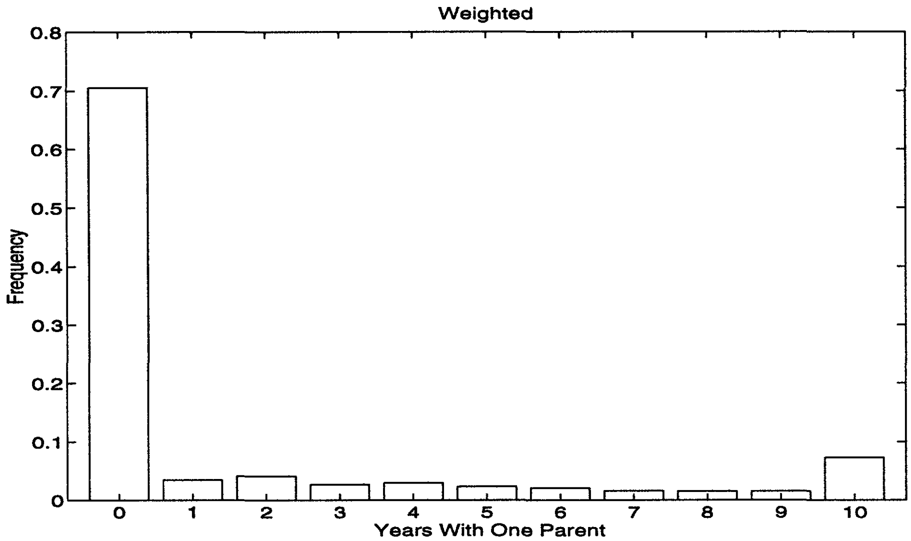

Figure 3-4: Weighted Distributions for *YRS_WITH_ONE*

on average, in poverty during the ten years is calculated to be 7.2%.

From this longitudinal data, one can easily construct the variable YRS_BELOW, which gives the number of years during which each individual's household income was below the poverty line. Weighted and unweighted distributions for this variable are provided in figure 3-5, and help to show the overrepresentation of impoverished households in the filtered PSID sample. One final poverty measure, CUM_POV_DEF gives the cumulative poverty deficit for each individual's household for the ten years of interest. More specifically, if we define x_i to be the number of dollars below the poverty line that the child's household is when he/she is i years old, then:

$$CUM\_POV\_DEF = x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} + x_{13} + x_{14} + x_{15} \quad (3.1)$$

It is important to note that x_i can never be negative, which means that, if an individual's household income does not drop below the poverty line during these ten years, his/her CUM_POV_DEF is zero.

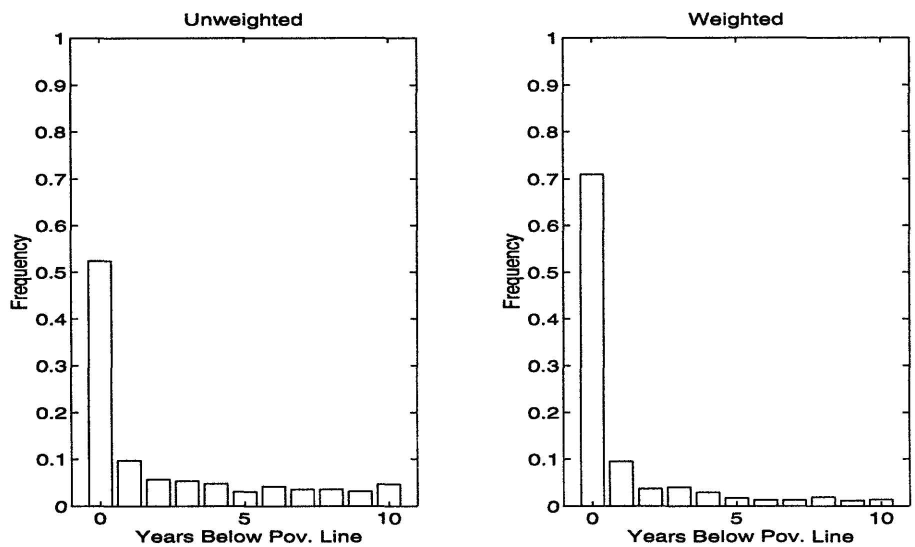

Figure 3-5: Weighted and Unweighted Distributions for *YRS_BELOW*

3.3 Correlation Between Income and Educational Attainment

Few people would question the assertion that poor children are less likely to excel academically than their more affluent counterparts. Numerous studies have convincingly established a significant relationship between a person's economic status during childhood and his/her subsequent educational attainment. One can also use the filtered portion of the PSID to see this correlative relationship. The following tables give estimates for the average years of education and high school graduation rates for individuals with different AV_INC_NEEDS, YRS_BELOW, and CUM_POV_DEF values. The compensatory weights are used in these calculations, yielding estimates for the true national averages. (Unless otherwise specified, these probability-of-selection weights will be used for the remainder of this document, so as to estimate national, as opposed to sample, parameters.) The *% of Population* column gives the estimated fraction of the population which satisfy the specified criterion (i.e. YRS_BELOW = 0.0).

AV_INC_NEEDS	YEARS_ED	HS GRAD. RATE	Sample Individuals	% of Population
0 - 1	11.91	75.2	314	7.2
1 - 2	12.04	73.9	568	24.0
2 - 3	12.82	90.4	373	25.1
3 - 4	13.17	93.7	246	21.9
4 - 5	13.95	99.3	104	10.7
5 and up	14.09	97.0	100	11.1

YRS_BELOW	YEARS_ED	HS GRAD. RATE	Sample Individuals	% of Population
0	13.22	92.6	893	71.0
1 - 3	12.29	79.0	354	17.3
4 - 6	11.90	69.2	205	5.9
7 - 10	11.85	73.9	253	5.8

CUM_POV_DEF	YEARS_ED	HS GRAD. RATE	Sample Individuals	% of Population
0	13.22	92.6	893	71.0
1 - 10,000	12.20	76.7	457	20.4
10,001 - 20,000	11.98	74.3	170	4.1
20,001 and up	11.91	74.5	185	4.5

If one chooses instead to compare the income and poverty information of individuals with different amounts of education, one sees a similarly strong positive correlation between income and educational attainment. Let an individual's categorical years of education, CAT_YRS, be defined as follows:

$$CAT\_YRS = \begin{cases} 1 & \text{if individual did not graduate from high school} \\ 2 & \text{if individual graduated from high school} \\ 3 & \text{if individual attended college but didn't finish four year degree} \\ 4 & \text{if individual finished a four year degree or more} \end{cases}$$

The following table gives income and poverty information for these four mutually exclusive groups. As one would expect after seeing the previous tables, the negative correlation between educational attainment and poverty is again quite strong.

CAT_YRS	AV_INC_NEEDS	YRS_BELOW	CUM_POV_DEF	Sample Individuals	% of Population
1	2.01	2.26	5,502	268	12.2
2	2.63	1.30	3,330	819	44.4
3	3.51	0.59	1,540	479	31.0
4	4.43	0.21	532	139	12.4

Though there is an unambiguous relationship between poverty and educational attainment, it is not so clear that income has any causal relationship with academic achievement. As statisticians have always asserted, *correlation does not imply causation.* Other variables which are simply correlated with income may be the actual determinants of a child's educational attainment.

Analyzing the relationship between an individual's years of education and other background variables, one can find similarly strong correlations. For example, children whose parents have done well academically are more likely to do well than those whose parents have not done so well. This may be due to the increased emphasis which better-educated parents tend to place on their children's academic performance. In the following table, PARENT_YRS represents the categorical years of education for the child's more educated parent. Household income and poverty information are also provided to show the correlation between parents' education and economic well-being.

PARENT_YRS	YEARS_ED	HS GRAD. RATE	YRS_BELOW	AV_INC_NEEDS
1	11.91	72.1	2.55	1.77
2	12.74	89.3	0.93	2.73
3	13.32	95.1	0.34	3.62
4	14.15	97.7	0.09	4.81

Assuming for the moment that income and parents' education are the only two possible determinants of a child's academic achievement, it is far from obvious from the data given above which of the two is more important. It may be the case that increased income leads to better nutrition, health care, and housing for the child, thereby producing a home environment more well-suited for concentrating on schoolwork, and that the parents' education is not an important factor. On the other hand, the parents' education may be the dominant factor, which is to say that the parents may coach the child through school, help him/her with homework, etc. In this scenario, income may not play a very big role. With appropriate statistical techniques, one could *attempt to* isolate the effect which both income and parents' education have on the child's future academic success. Depending on the result, one could then develop a strategy to improve the educational attainment of poor children.

Of course, there are many more possible determinants of a child's academic success than these two, and they must also be considered. For example, children who do not live with two parents tend to drop out of school sooner than the children of two-parent households do. This may be due to the lack of stability in a one-parent environment, the reduced attention which the child is likely to receive from an adult, or to some outside factor which is correlated with the number of parents in the household. Analyzing the PSID sample, one can easily see the negative correlation between the years a child spends with one parent (out of the ten years of interest) and his/her educational attainment. It is also the case though, that the years an individual spends with one parent is negatively correlated with the household income. The following table provides household income and poverty information for those individuals who have certain YRS_WITH_ONE values. This measure includes the years during which a child lived with no biological parents, because he/she must have been living with some parental guardian.

YRS_WITH_ONE	YEARS_ED	HS GRAD. RATE	YRS_BELOW	AV_INC_NEEDS
0	13.10	90.9	0.55	1.66
1 - 3	12.48	80.6	1.15	2.62
4 - 6	12.63	82.5	1.71	2.71
7 - 10	12.24	78.7	3.65	3.38

A similar correlative relationship seems to exist between the number of parental separations which a child experienced during the ten years of interest and his/her subsequent educational attainment. It is widely believed that such an event can have a pernicious effect on a child's emotional state. The variable NUM_SEPS gives the number of parental separations which took place in each child's household while his/her age ranged from six to fifteen, and appears to be negatively correlated with both income and educational attainment.

NUM_SEPS	YEARS_ED	HS GRAD. RATE	YRS_BELOW	AV_INC_NEEDS
0	12.99	89.2	0.91	3.19
1	12.60	83.3	1.61	2.54
2 or 3	12.47	72.8	1.88	2.47

Similar correlations with income and with educational attainment exist for other background variables included in this PSID sample. The number of siblings an individual has during his/her childhood may be an important determinant of his/her educational attainment. With many children in the household, the parent(s) may have less time to spend with each individual. Additionally, there are more people to feed, clothe, and house, which means that an income which is more than adequate for a small family may not be sufficient for a larger one. So, it seems reasonable that children with many siblings tend not to remain in school for as long as children with relatively few siblings. The following table provides educational and income information for children with different family sizes. As explained earlier in this chapter, the AV_NUM_SIBS variable gives the average number of siblings the individual had during the ten years of interest.

AV_NUM_SIBS	YEARS_ED	HS GRAD. RATE	YRS_BELOW	AV_INC_NEEDS
0 - 1	13.05	91.9	0.54	4.06
1 - 2	13.10	90.7	0.58	3.42
2 - 3	12.94	88.5	1.00	2.78
3 - 4	12.60	79.8	2.08	2.26
4 - 5	12.24	71.6	2.54	1.75
5 and up	12.18	88.3	2.69	1.70

The high school graduation rate for those individuals with five or more siblings appears to be peculiarly high, though the corresponding years of education seems consistent with the downward trend.

Another background variable which appears to be correlated with a person's educational attainment is his/her race. Using the filtered portion of the PSID, it appears that nonwhite children remain in school for a shorter period of time than white children do. As one would intuitively expect given the abundance of information regarding black-white earnings differentials, it is also the case that race is correlated with income. The following table provides national estimates for years of education and household income for nonwhite (black and hispanic) and white individuals who were between zero and six years old in 1968.

RACE	YEARS_ED	HS GRAD. RATE	YRS_BELOW	AV_INC_NEEDS
NonWhite	12.50	81.9	3.38	1.77
White	12.99	88.9	0.60	3.30

Finally, the number of years during which the head of an individual's household is disabled is correlated with that person's educational attainment and the family's income. The negative correlation between YRS_HEAD_DISABLED and household income may be due to the loss of income a family is likely to suffer after a debilitating injury, whereas the relationship with the child's education may stem from the inability of the injured parent to spend as much "quality time" with the child as he/she otherwise would. The following table shows these correlative relationships.

YRS_HEAD_DISABLED	YEARS_ED	HS GRAD. RATE	YRS_BELOW	AV_INC_NEEDS
0	13.13	91.8	0.50	3.43
1 - 3	12.49	82.7	2.06	2.24
4 - 6	12.58	78.3	2.23	2.23
7 - 10	11.72	60.7	3.77	1.66

This is by no means a complete list of all of the important correlative relationships which one must bear in mind when attempting to assess the effect of poverty on an individual's educational attainment. Instead, these relationships were chosen to point out that, if one hopes to establish a causal relationship between poverty and education, one must make a concerted effort to adjust for as many of the potentially confounding factors as possible. If an analysis is not done carefully, one may well interpret a correlative relationship as a causative one.

Chapter 4

Three Methods for Determining the Effect of Poverty on Educational Attainment

In this section, I will discuss the methodology employed by Haveman and Wolfe in an initial analysis of the PSID data. Including all 1705 sample individuals in several econometric analyses, they find statistically significant results for the effect of poverty on educational attainment. Then, I will begin to introduce a multivariate matching technique known as propensity score matching. Developed by Don Rubin, a professor of statistics at Harvard University, propensity score matching helps one to make reasonable causal inferences from observational data. By matching on the propensity score, one can attempt to control for systematic differences between treated and control groups (i.e. poor and non-poor). Finally, I will describe Mahalanobis metric matching, which eliminates bias between two groups by matching individuals who have similar background variables. Unlike propensity score matching, which matches individuals with similar propensity scores, this technique attempts to pair individuals who are close on all matching variables. I will employ both matching techniques in the subsequent empirical analysis.

4.1 A Previous Analysis

4.1.1 Modelling Techniques

Haveman and Wolfe used a data set with the same 1705 individuals to estimate the effect of poverty on an individual's educational attainment. They used three income/poverty measures for the analysis, all of which have been previously mentioned in this document (the AV_INC_NEEDS, YRS_BELOW, and CUM_POV_DEF variables). Additionally, they constructed three indicators of an individual's educational attainment: the number of years of schooling completed, the categorical years of education (as was described earlier in this chapter), and a variable HS_GRAD, which takes on a value of one if the person graduated from high school and zero otherwise. Then, they constructed models to explain the educational attainment of the sample children. For these models, least-squares multiple linear regressions were performed, and coefficients were estimated for each of the dependent variables. Two general types of models were constructed. The first type, known as the parsimonious models, contains fewer variables than the second type. An example of a parsimonious model follows, with the t-statistic corresponding to each coefficient enclosed in parentheses below it.

$$\begin{array}{lcccc} Years\_Ed = & .73 * AvIncNeeds & +.03 * NonWhite & +.02 * Female & +.56 * (Female * NonWhite) \\ & (8.9) & (0.2) & (0.2) & (3.8) \\ & -.02 * YrsWithOne & +.64 * MomEducation & -.04 * AvgNumSibs & +11.8 \\ & (1.9) & (7.7) & (1.4) & (87.6) \end{array}$$

Eight other parsimonious models are constructed in a similar fashion. The coefficients for some of these models, along with the corresponding t-statistics, are listed in the following table.

Model Number	1	2	6	7	9
Education Variable	HS_GRAD	HS_GRAD	YEARS_ED	CAT_YRS	CAT_YRS
Income Variable	YRS_BELOW	AV_INC_NEEDS	CUM_POV_DEF	YRS_BELOW	CUM_POV_DEF
Income Coeff. and t	-.05 (-3.4)	.42 (4.8)	.09 (2.2)	-.04 (-4.4)	-.06 (2.9)
NonWhite Coeff. and t	.16 (1.3)	.20 (1.6)	-.15 (-1.2)	-.04 (-0.7)	-.07 (-1.2)
Female Coeff. and t	.002 (0.0)	-.02 (0.2)	.08 (0.8)	.07 (1.4)	.07 (1.4)
Female*NonWhite Coeff. and t	.34 (2.2)	.36 (2.3)	.49 (3.3)	.24 (3.3)	.24 (3.3)
YrsWithOne Coeff. and t	-.04 (-3.3)	-.03 (3.0)	-.05 (4.8)	-.02 (3.4)	-.03 (4.8)
MomEducation Coeff. and t	.57 (6.6)	.49 (5.6)	.84 (10.3)	.39 (9.6)	.41 (10.2)
AvgNumSibs Coeff. and t	-.07 (2.8)	-.05 (1.8)	-.11 (4.1)	-.05 (4.2)	-.06 (4.4)
Constant and t	1.05 (9.8)	.64 (4.7)	12.6 (118.8)	2.28 (43.2)	2.27 (42.8)

The second group of models were constructed with several more explanatory variables. Two variables not previously mentioned, YRS_IN_SMSA and RELIGIOUS (*find out exact* *definition*), are included in these models. An example is provided below.

$$
\begin{array}{lllll}
Years\_Ed = & .53 * AvIncNeeds & +.07 * NonWhite & +.01 * Female & +.59 * (Female * NonWhite) \\
 & (5.8) & (0.6) & (0.1) & (4.1) \\
 & -.03 * YrsWithOne & +.46 * MomEducation & -.05 * AvgNumSibs & -.003 * Yrs\_In\_SMSA \\
 & (-2.0) & (5.4) & (-2.0) & (-0.4) \\
 & +.63 * DadEducation & +.26 * Religious & +.01 * Yrs\_Mom\_Wk & -.04 * Yrs\_Head\_Disabled \\
 & (6.5) & (2.0) & (0.0) & (-2.8) \\
 & +11.6 & & & \\
 & (6.0) & & &
\end{array}
$$

As they did with the parsimonious-type model, the researchers constructed nine models of this type using least-squares multiple regression analysis, yielding a model for each possible combination of the three income measures with the three educational indicators.

4.1.2 Results of the Study

Following calculation of these models, Haveman and Wolfe estimated the effect of a reduction in poverty on educational attainment. For these simulations, the coefficients from some of the parsimonious and the more extensive models were used. Their results are summarized in the table below.

	HS GRAD. RATE	YEARS_ED	CAT_YRS
Original Unweighted Average	84.3%	12.61	2.29
If reduce YRS_BELOW by half	86.0%	12.68	2.33
If reduce YRS_BELOW to zero	87.5%	12.75	2.37
If reduce CUM_POV_DEF by half	85.2%	12.64	2.30
If reduce CUM_POV_DEF to zero	86.0%	12.67	2.32

The results of the study suggest that reducing the poverty which many of these sample households confront could have modest educational benefits for the children of these households. For example, if one were to reduce the number of years during which the sample households were exposed to poverty to zero, then it appears that the high school dropout rate (in the sample) would drop from 15.7% to 12.5% and that the average years of education would increase from 12.61 to 12.75. These preliminary results suggest that increased income may well improve the academic achievement of poor children.

In fact, the gains listed above are averaged over the entire population, and therefore understate the gains for impoverished children. Roughly 18% of the PSID sample children

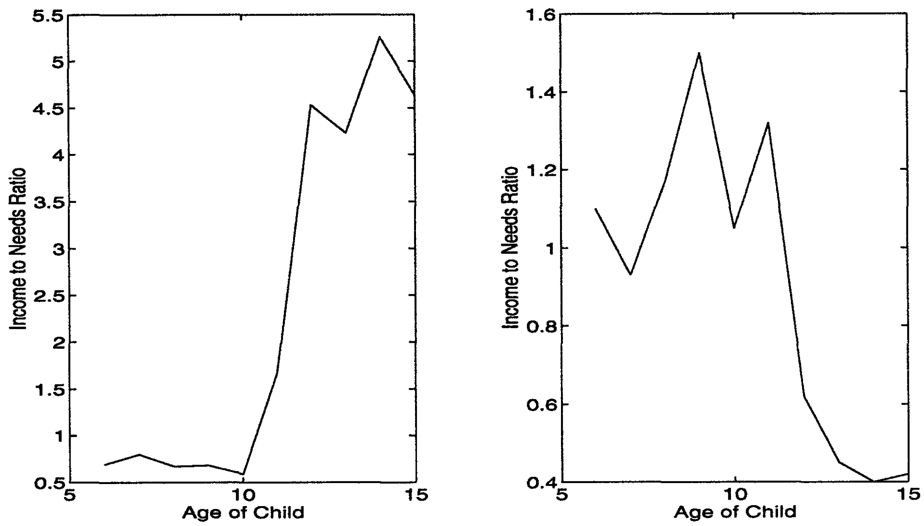

Figure 4-1: Income/Needs for two Individuals With $YRS\_BELOW = 5$

are, on average, poor. Because these are the individuals who would directly benefit from a reduction in poverty, their gains will be larger than is implied by the chart above. If the overall graduation rate would improve by 3.2% after all of the children's household incomes were lifted above the poverty line, then the graduation rate among poor children would increase by roughly 18% (= .032 / .18). Similarly, an increase of 0.14 in the number of years of education implies that the poor PSID sample children would receive, on average, 0.78 (= .14 / .18) years more of education.

4.1.3 Comments on the Analysis

Models such as those described in the previous section are useful because they allow the analyst simultaneously to consider a number of background variables while analyzing more than 1700 individuals. The price of such a model, however, is that one must accept as true a number of important assumptions. For example, the assumption that there exists a linear relationship between the educational variable and the independent variables may well be violated. Consider the independent variable, YRS_BELOW. The child's age when he/she is exposed to poverty is likely to play a role in what effect this poverty will have on his/her subsequent educational attainment. Figure 4-1 shows the household income-to-needs ratios for two different individuals included in the analysis. Each of the individuals whose household income patterns are described in these graphs have YRS_BELOW values of five. But the economic circumstances in which these children found themselves during these ten years are quite dissimilar. So, the YRS_BELOW variable fails to distinguish between the economic circumstances of these two individuals. It would not be superfluous to examine the income patterns of individuals more closely than one possibly can using the YRS_BELOW, AVG_INC_NEEDS, and CUM_POV_DEF measures. Similarly, a child who spends the first five years (i.e ages six through ten) with one parent is likely to be affected differently from one whose parents are separated while his/her age is between eleven and fifteen. Problems such as these make a closer look at the data desirable. Other possible problems with the models described above include the exclusion of important background variables and the correlations between the independent variables.

Perhaps the most important shortcoming in this study, though, is the possibility that correlative relationships are being interpreted as causative ones. Correlative relationships tell everything that one needs to know about how a group of variables are statistically related, but says next to nothing about how the variables are causally related.[4, p. 30] Causation is rarely settled by statistical arguments alone, but is made more plausible when three criteria are satisfied. Consider two variables x and y. First, there should be a consistent and unambiguous relationship between x and y. Second, it should be shown that there exist no possible common causes of x and y or alternatively, that the relationships between the possible common causes,x, and y are not enough to explain the clear relationship between x and y. Finally, the assumed direction of causality (i.e. x causes y) should be reasonable, which is to say that the analyst should demonstrate that y could not cause x.[30, p. 261] In Haveman and Wolfe's models, the second criterion may be violated. Looking at the extended model, if it is the case that a higher number of parental separations tends to cause a reduction in household income and a reduction in the number of years of the child's education, then the positive correlation between income and educational attainment may actually be best explained by this omitted variable. If it is, then the conclusions which one might draw from their preliminary analysis could be erroneous.

Here, we are trying to determine what effect a reduction in childhood exposure to poverty would have on an individual's educational attainment. The responsiveness of a person's educational attainment to a change in his/her household's economic well-being, holding all other things equal, would ideally be determined by an experiment. In engineering, chemistry, or physics, experiments are designed and subsequently conducted to determine causative relationships. For situations in which only observational data are available, one must proceed very carefully to achieve a judicious analysis of the effect of one variable on another.

4.2 Propensity Score Matching

In order to make causal inferences about the effect one variable *X* has on some other variable *Y*, one would ideally design an experiment which held constant any potentially confounding factors. For example, in randomized experiments, the results between the treated and control groups can usually be compared because the individuals will, on average, be similar with respect to the distribution of important background variables. This is not always feasible though, so observational data must frequently be used to estimate the effect of a treatment on some background variable. In these circumstances, direct comparisons between a group which is "exposed" to some treatment and a control group which is not exposed may be misleading because of systematic differences between the two groups.[36, p. 33] Matching techniques aim to group treated and control individuals so that direct comparisons are more meaningful. For the PSID sample, one possible group of treated individuals could be composed of those children whose household incomes are, for the first few years, below the poverty line, and then above for the remaining several years. These individuals might then be compared with similar individuals who remain in poverty to make a first-pass estimate of the effect of a permanent income infusion to the household on a child's educational attainment.

4.2.1 The Propensity Score: The Coarsest Balancing Score

A balancing score, *B(X)*, is a function of the observed background variables X such that the conditional distribution of the background variables, given *B(X)*, is the same for both the treated and the control groups. The most trivial of all balancing scores is the vector *X*, whereas the coarsest balancing score is the propensity score, which is simply the probability of exposure to the treatment given the vector of background variables.[35, p. 42] In a randomized experiment, the propensity score is the same for all units, because each individual is equally likely to fall into the treated group. When using observational data, though, the exact form of the propensity score is not known, and must be estimated from the available data.

Propensity scores can be used to adjust for systematic differences between treated and control groups. Treated and control individuals with the same value of the propensity score $e(X)$ will, on average, have the same distributions of background variables X. Therefore, exact matching on $e(X)$ will tend to balance the distributions of the background variables in the treated and control groups. Matching is a method of sampling individuals from a large group of controls to form a group more appropriate for direct comparison with the treated group.

Ideally, one would match control and treated units which had the same values for all background variables X. As the number of background variables increases though, this criterion becomes prohibitively difficult to meet. Fortunately, exact matching on any balancing score $B(X)$ is sufficient to obtain the same probability distributions of the background variables for both treated and control units.

Several issues must be addressed before proceeding. First, matching on the propensity score will only serve to balance the distributions of the *observed* background variables. Therefore, if there are any important background variables which have not been observed, then systematic bias may remain. The less correlated any unobservable variables are with the observed ones, the more likely it is that substantial bias will still remain after the matching has taken place.

Second, because the exact functional form of the propensity score is not known in observational studies, it must be estimated from the available data. Third, if there exist more than a few background variables on which to match, exact matches on the propensity score will rarely be available. As a result, when constructing treated and control groups for direct comparison, one must determine how close two units' propensity scores must be for a match to be appropriate. Finally, matching on the propensity score $e(X)$ will only balance the distribution of background variables on average, so adjustments may be required to account for any imbalances in the matched distributions.

4.2.2 Defining the Treated and Control Groups

When trying to meaningfully estimate the effect of childhood exposure to poverty on educational attainment, one must first determine the appropriate groups for comparison. As was pointed out in the previous chapter, there is an unambiguous positive correlation be-

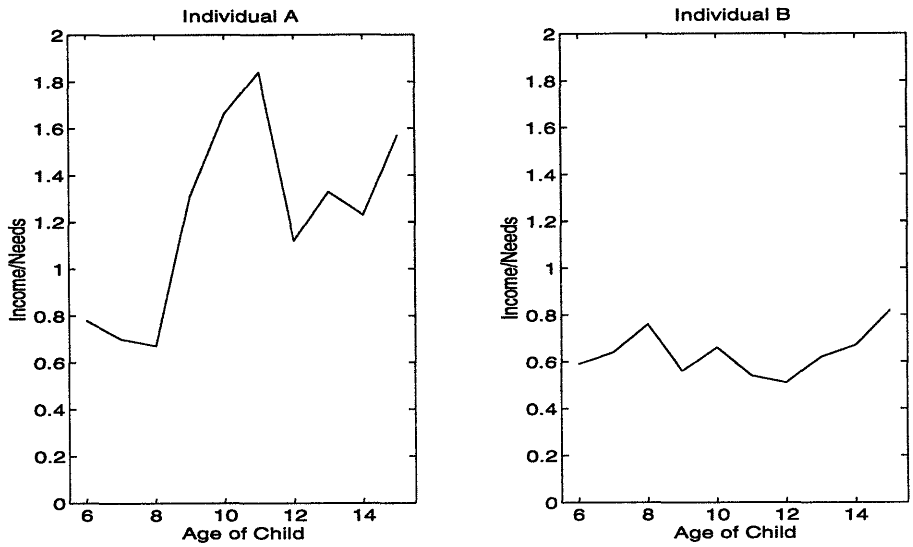

Figure 4-2: Income/Needs Patterns for a Treatment and Control Individual

tween income and educational attainment. This, of course, does not imply that childhood exposure to poverty adversely affects a person's subsequent academic achievement. Given the limited nature of the longitudinal income information, we cannot make any statements about the effects of childhood poverty before an individual reaches the age of six or seven.

However, we can make statements about the effects of exposure to poverty later in a child's life. In order to form *experimental* and *control* groups, there must exist some *treatment* which a substantial number of the sample children undergo. For example, consider two individuals, A and B, who were identical in all observed background variables and whose household income-to-needs ratios were identical during years six, seven, and eight. Also, assume that their household incomes during these three years were below the poverty line. Then, if individual A's household income suddenly rises and remains above the poverty line for the rest of his/her childhood, and if individual B's household income remains below the poverty line, then an *experiment* has been conducted. Figure 4-2 reveals possible income patterns during the ten years of interest for two such individuals. Of course, this experiment has not been conducted under ideal conditions, because the rise in income is probably not exogenous. Nevertheless, it seems a reasonable way to make a first-pass estimate as to the effects of lifting a poor child's household income above the poverty line. Instead of asking what cost an impoverished child incurs as a result of being poor, this experiment tries to estimate the benefit which lifting a child out of poverty may have.

It is important to note that families whose incomes rise above the poverty line are presumably *not* representative of the population of poor families in this country. In other words, it may be the case that a family which experiences such a sustained increase in income is more motivated and well-informed than most poor families. So, if a child in a family such as A's tends to achieve more academically than a child such as B, it may not be due to the increase in income, but rather due to unmeasured factors. Nevertheless, by attempting to match individuals who are similar with respect to all background variables, *one hopes to account for such differences.* The approach is not ideal, but is perhaps the best way to estimate, from this observational data, the effect that an income infusion (above the poverty line) would have on a poor child's future educational attainment.

So, one possible treated group could include those individuals who are exposed to poverty for the first three years and then, for each of the next seven years, have household incomes above the poverty line. Other treated groups could be constructed using similar income pattern criteria. For example, one might investigate children whose household income is below the poverty line for the first and second years, and is above for all of the eight remaining years. A potential control group for either of these treated groups could include those individuals whose household income remains below the poverty line for all ten years.

4.2.3 Calculating the Propensity Score

The propensity score *e(X)* is the conditional probability of exposure to the treatment, given the vector of background variables, X. Assume a unit's value for the random variable y takes on a value of one if a unit is in the treated group and zero if a unit falls in the control group. Therefore, the outcome variable for the propensity score is dichotomous. When this is the case, logistic regression is frequently the model-builder of choice. The two main differences between logistic regression and linear regression are: (1) the underlying assumptions which must be met and (2) the choice of the parametric model.

Assuming only one independent variable x in a typical linear regression model, the quantity of interest is typically the conditional mean of the dependent variable, given the value of the independent variable. This value is known as the conditional mean, *E(Y—x)*, and is calculated as follows:

$$E(Y|x) = \beta_0 + \beta_1 x \tag{4.1}$$

This equation implies that Y can take on any value as x ranges from negative infinity to positive infinity. If the independent variable is dichotomous, though, the conditional mean must be less than or equal to one and greater than or equal to zero. Several distribution functions have been used in the analysis of a dichotomous outcome variable. The logistic distribution is frequently chosen because of its flexibility and because it lends itself easily to a meaningful interpretation.

When the logistic distribution is used, the conditional mean can be written as $\pi(x) = E(Y—x)$. This is simply the probability, given the independent variable x, that the value of y is one. The particular form of the logistic regression model which we will employ is as follows:[10]

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{4.2}$$

$\pi(x)$ is simply the propensity score which was described above. By performing a logit tranformation on the propensity score, one can obtain an equation for the log odds, $Q(x)$, which is linear in the dependent variables.

$$Q(x) = \frac{\pi(x)}{1+\pi(x)} \tag{4.3}$$

The parameters β_0 and β_1 must be estimated from the available data. The importance of this logit transformation lies in the fact that $Q(x)$ is linear in the background variable x. If, instead of one background variable, there were several, then the functional form of the propensity score would be modified as follows:

$$\pi(x) = Prob(y = 1|x_1, x_2, ..., x_n) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_n x_n}}{1+e^{beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_n x_n}} \tag{4.4}$$

Here, the value of π(x) would yield the probability of exposure to the treated group, given the vector of background variables $(x_1, x_2, ..., x_n)$.

Suppose there exists a sample of n independent observations of the type (y_i,x_{i1},x_{i2}, . . . ,x_{in}), where y_i is the value of the outcome variable (zero or one, depending on whether the individual is in the control or treated group) for the ith unit and x_{ij} is the value of the jth background variable for the ith unit. To fit the logistic regression model described above, one must estimate the values of the parameters (β_0,β_1, . . .,β_n). The usual method of estimation under the logistic regression model is maximum-likelihood. This technique yields values for the unknown parameters which maximize the probability of obtaining the observed data. To apply this method, one must first construct a likelihood function, which expresses the probability of the observed data in terms of the unknown parameters.

With the dependent variable y coded as a zero or one (for the control and treated groups, respectively), then $\pi(x_1, x_2, ..., x_n)$ gives the conditional probability that an individual is in the treated group and $1 - \pi(x_1, x_2, ..., x_n)$ is the conditional probability that an individual is in the control group. So, for those units with y_i equal to one, $\pi(x)$ is the unit's contribution to the likelihood function. If, on the other hand, y_i is zero, then $1 - \pi(x)$ represents this unit's contribution. Thus, one can express a unit's contribution to the likelihood function with the following term:

$$\zeta(x_{i1}, ..., x_{in}) = \pi(x_{i1}, ..., x_{in})^{y_i}[1 - \pi(x_{i1}, ..., x_{in})]^{1-y_i} \tag{4.5}$$

Here we will assume that the observations are independent, so the likelihood function for all of the m units is given by $\Lambda(\beta_0, \ldots, \beta_n)$:

$$\Lambda(\beta_0, ..., \beta_n) = \prod_{i=1}^{m} \zeta(x_{i1}, ..., x_{in}) \tag{4.6}$$

By taking the log of both sides of this equation, one can obtain a more tractable mathematical expression. The *log likelihood*, $\lambda(\beta_1, \ldots, \beta_n)$ is:

$$\lambda(\beta_0, ..., \beta_n) = \sum_{i=1}^{m} y_i ln[\pi(x_{i1}, ..., x_{in})] + (1 - y_i) ln[1 - \pi(x_{i1}, ..., x_{in})] \tag{4.7}$$

To find the values of the coefficients which maximize $\lambda(\beta_0, \ldots, \beta_n)$, one must take the partial derivatives of the equation above with respect to each of the coefficients. This will yield $n + 1$ likelihood equations which are nonlinear in the parameters $(\beta_0, \ldots \beta_n)$. Special iterative methods are required to solve these equations. Fortunately, these techniques have been programmed into available logistic regression software.

The values of $(\beta_0, \ldots, \beta_n)$ given by the solutions to the likelihood equations are *maximum likelihood estimates* of the true parameter values. These MLE parameters will be denoted as $(\hat{\beta}_0, \ldots, \hat{\beta}_n)$ to emphasize that they are estimates of the true values. Additionally, $\hat{\pi}(x_{i1}, \ldots, x_{in})$ is the maximum likelihood estimate of the probability that

y is equal to one. It is important to note that the sum of the predicted values of y is equal to the sum of the observed values of y:

$$\sum_{i=1}^{m} y_i = \sum_{i=1}^{m} \hat{\pi}(x_{i1}, ..., x_{in}) \tag{4.8}$$

After calculating the maximum likelihood estimates of all $n + 1$ coefficients, one can then calculate the propensity score for each unit. Then, by matching treated and control individuals with similar propensity scores, one can begin to form control and treated groups which are more appropriate for comparison.

4.3 Mahalanobis-Metric Matching

If its form is estimated accurately, the propensity score will, on average, balance the distributions of background variables between treated and control groups. Despite this, one may wish to give more importance to the individual background variables themselves by matching individuals who are "close" to one another on all background variables. More specifically, it may be the case that two individuals who are matched on the propensity score are different from one another on several background variables, but that the coefficients are such that their propensity scores are quite close to one another.

When matching pairs of treated and control units, one hopes to form matched samples which are similar with respect to the distribution of important background variables. One measure of the "distance" between the background characteristics of two units is the Mahalanobis distance.[45, p. 293] Consider two individuals, A and B, the first of whom is in the treated population and the second of whom is in the control population. For each individual, there exists not only treatment and outcome information, but also information about their background characteristics. Let $X_A = (x_{A1}, \ldots, x_{An})$ and $X_B = (x_{B1}, \ldots, x_{Bn})$ be the vector of background variables for individuals A and B.

Then, one must calculate the covariance matrix for both the treated and control populations. For both populations, the corresponding covariance matrices will be $n * n$ in dimension. If there are q individuals in the control population, then the unbiased estimate for the covariance between background variables x_i and x_j in the control population, σ_{ijC}, is calculated as follows:

$$\sigma_{ijC} = \frac{\bar{x_i y_i} - \bar{x}_i\bar{y}_i}{q-1} \tag{4.9}$$

As is obvious from the formula above, $\sigma_{ijC} = \sigma_{jiC}$. After calculating each of the covariances between the background variables of individuals *in the control population*, the covariance matrix for the control population, S_C can be constructed, and it will be of the following form:

$$S_C = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \dots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \dots & \sigma_{nn} \end{bmatrix} \tag{4.10}$$

One can proceed in a similar fashion in calculating S_T, the covariance matrix for the treated population. Assume that there are kq individuals in the treated population. Then, to find the covariance matrix S_{TC} needed for the Mahalanobis distance calculation, combine the two population covariance matrices as follows:

$$S_{TC} = \frac{(q-1)S_C + (kq-1)S_T}{q + kq - 2} \tag{4.11}$$

Then, the Mahalanobis distance, M_{AB} between two units A and B is defined as:

$$M_{AB} = (X_A - X_B) * S_{TC}^{-1} * (X_A - X_B)^T \tag{4.12}$$

The Mahalanobis distance, as defined above, is a measure of the closeness of two units from the treated and control populations. By matching individuals who are close with respect to the Mahalanobis distance, one can eliminate much of the systematic bias between the treated and control groups. This matching variable differs from the propensity score principally by its emphasis on *all* of the background variables, whereas the latter method matches on only on the propensity score (though this is a function of all of the important background variables).

Chapter 5

Assessing the Effect of a Sustained Rise in Income

When attempting to make causal inferences about the effect a particular treatment would have had on a unit which received some other treatment, one essentially confronts a missing data problem. By comparing the treatment effects of individuals who were similar in all important respects *before the start of the treatments*, one can obtain an unbiased estimate for the difference in treatment effects. For this reason, I will now compare individuals who were initially poor and whose household income subsequently rose above the poverty line with those who were consistently poor throughout the ten-year time period. The first "treatment" for this experiment is the rise in income, whereas the second is continued exposure to poverty. Employing matched sampling methods to construct treated and control groups which are appropriate for comparison, I then make a first-pass estimate of the effect of such a sustained rise in income on a child's academic achievement.

It is important to note, though, that this sustained rise in income is quite different from a government subsidy to an impoverished family. Because the parents of the "experimental" children seem to have lifted themselves above the poverty line, it is quite likely that they are more ambitious and determined than individuals who remain below the poverty line. In other words, their families may differ in unobserved respects from the control individuals, and thus the estimated treatment effect may be misleading. If the parents of the experimental children pass on their extra determination, their children may tend to achieve more academically because of this determination and not because of the additional income.

Alternatively, if the parents of the experimental children had to spend more time at work, they may spend less time with their children, and this reduction could adversely affect their children's academic performance. It is not clear which of these two effects would tend to dominate the other, so the estimated effect may be biased upwards or downwards.

Despite the obvious disadvantages of focusing primarily on those individuals whose families seem to have pulled themselves out of poverty, their income patterns follow precisely the trajectory whose effect we wish to estimate. In the absence of more detailed information (i.e. about welfare payments), it is therefore reasonable to focus on these individuals to obtain a first-pass estimate of the effect which an income infusion would have on an impoverished child's educational attainment.

5.1 Construction of the Experimental and Control Groups

The experimental group includes children whose household incomes were below the poverty line for a period of time and then, for some reason, rose *and remained* above the poverty line. Unfortunately, the available longitudinal income information includes only those years during which the children's ages were between six and fifteen. Nevertheless, by looking at individuals who were exposed to poverty only for the first one, two, or three years, and comparing them with children who were poor throughout the ten year period, one can make a first-pass estimate of the costs of continued exposure to poverty.

Three mutually exclusive experimental groups are constructed. To be included in one of the experimental groups, a PSID child's household income must follow one of three income patterns during the ten years of interest. Let M_i equal the ratio of the child's household income to the household poverty line when he/she was i years old. The selection criteria for each of the three groups are as follows:

Group One	$M_6 < 1.00$	$M_7 > 1.00, \ldots, M_{15} > 1.00$
Group Two	$M_6 < 1.00$ and $M_7 < 1.00$	$M_8 > 1.00, \ldots, M_{15} > 1.00$
Group Three	$M_6 < 1.00$, $M_7 < 1.00$, $M_8 < 1.00$	$M_9 > 1.00, \ldots, M_{15} > 1.00$

There are forty individuals whose income patterns satisfy the group one selection criteria, twenty-one individuals in group two, and thirteen in group three. The yearly income data for three individuals, one from each of the experimental groups, are shown in figure 5-1.

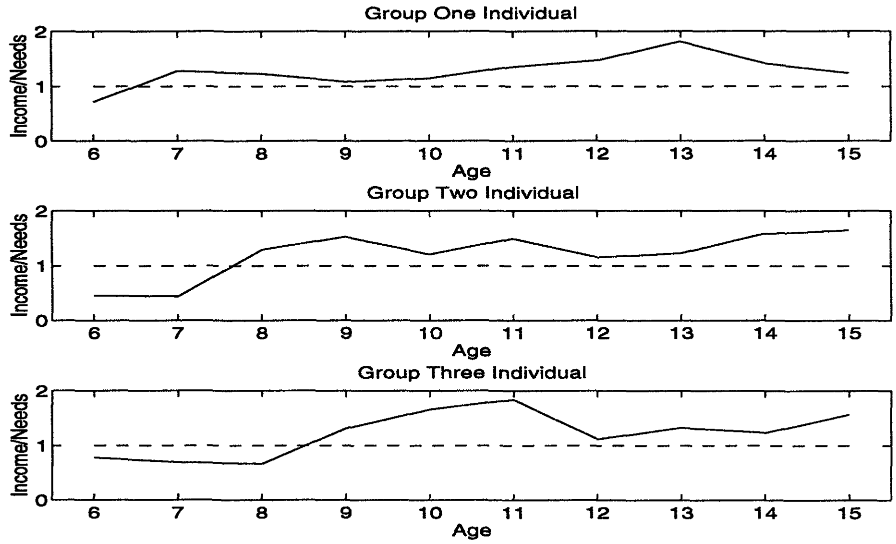

Figure 5-1: Income/Needs Data for Three Different Experimental Individuals

To estimate the effect of rising out of poverty, as those individuals described above did at different stages of their lives, I use a control group of individuals who remain impoverished throughout the ten years of interest. There are seventy-nine individuals whose household income is below the poverty line for *all* ten years. Because the number of individuals in each of the three experimental groups is quite small, the information about most of the individuals in the PSID sample will not be considered in this portion of the analysis.

5.2 The Background Variables

With control and experimental groups as defined above, I aim to estimate the effect that an income infusion to an impoverished child's family would have on his/her subsequent educational attainment. Here, the "experiment" is the income infusion. To compare individuals who are similar in all measurable respects *before* the experiment takes place, the individuals will be matched on all available, *pre-treatment* background variables. Examples of these include an individual's race, gender, and parents' education. Additionally, there exists longitudinal information concerning, for example, the number of parental separations, the number of years living with one parent, the number of siblings, and the household income for each individual. When matching individuals on the basis of these background variables, one should focus on *pre-treatment* information.

Unfortunately, some of the longitudinal variables mentioned above are not available on a yearly basis. For example, instead of knowing whether an individual lived with one parent when he/she was six, seven, and/or eight, the filtered PSID data set contains only the total (0, 1, 2, or 3) number of years during these three years while he/she was living with one parent. The variable EARLY_WITH_ONE can take on one of four possible values, depending on how many years the child spent with one parent when his/her age was between six and eight.

$$EARLY\_WITH\_ONE = \begin{cases} 0 & \text{if individual spent none of the three years living with one parent} \\ 1/3 & \text{if individual spent one of the three years living with one parent} \\ 2/3 & \text{if individual spent two of the three years living with one parent} \\ 1 & \text{if individual spent all three years living with one parent} \end{cases} \tag{5.1}$$

Data concerning the number of years the head of the child's household is disabled, the number of parental separations, and the years during which the individual's mother worked are also aggregated over this three year period, yielding the variables EARLY_HEAD_DISABLED, EARLY_SEPARATIONS, and EARLY_MOM_WORKED. Therefore, when matching on these background variables for experimental groups one and two, some post-treatment information is included in the matching. This can unfortunately not be avoided and represents a shortcoming in the analysis.

The only available information about the number of siblings each individual has and the number of years during which he/she lived in the south is aggregated over the *entire ten year period.* The variable YRS_SOUTH gives the number of years that each individual lived in the south. More than 95% of the sample individuals have a value of zero or ten for YRS_SOUTH (1635 out of 1705, to be exact) though, so this absence of yearly information is not as problematic as it might initially appear to be. If an individual lived in the south for all ten years, then it logically follows that he/she must have lived in the south for the first, second, and third years.

The same cannot be said, though, of the number of siblings which a sample individual has. Because the variable AVG_NUM_SIBS represents the average number of siblings an individual had during the ten years of interest, matching on this variable may present problems. For example, if a person had two siblings throughout the ten years, he/she would have the same value for this variable as one who had one sibling for the first four years, two siblings for the next three years, three siblings for the next two years, and four siblings for the final year. To suggest that these two individuals were similar with respect to the number of siblings they had seems preposterous. Nevertheless, the size of an individual's family is an important determinant of the type of environment in which he/she grew up. So, despite the obvious problems with the AVG_NUM_SIBS variable, it is the best *available* measure of the size of an individual's family and is included as a matching variable.

A complete list of matching variables is provided in the following table. Unweighted averages for the control group and the three experimental groups are included for all of the relevant matching variables. Two new variables, DAD_ED and MOM_ED are included. The first/second variable takes on a value of one if the individual's father/mother graduated from high school and zero otherwise. These variables were introduced because very few of the experimental or control individuals had parents who had attended college. The control group includes the 79 PSID children whose household incomes were below the poverty line for all ten years of interest.

Matching Variable	Control Avg. 79 individuals	Group One Avg. 40 individuals	Group Two Avg. 21 individuals	Group Three Avg. 13 individuals
DAD_ED	.063	.250	.143	.154
MOM_ED	.127	.450	.429	.462
ONE_PARENT_68	.620	.175	.333	.539
NO_PARENT_68	.101	.000	.048	.000
FIRSTBORN	.114	.250	.191	.077
FEMALE	.570	.450	.619	.462
NONWHITE	.962	.475	.524	.539
YR_SIX_INC_NEEDS	.505	.705	.763	.615
YR_SEVEN_INC_NEEDS	.550	—	.722	.778
YR_EIGHT_INC_NEEDS	.576	—	—	.752
WEIGHT	5.92	14.70	12.52	13.54
AVG_NUM_SIBS	4.20	2.79	2.73	2.62
EARLY_SEPS	.042	.033	.032	.026
EARLY_MOM_WORKED	.414	.658	.603	.333
EARLY_HEAD_DISABLED	.262	.133	.206	.282
EARLY_WITH_ONE	.806	.217	.508	.692
YRS_SOUTH	7.81	7.00	6.19	3.85

A quick comparison between the means of the background variables for the control and experimental groups reveals that there exist substantial differences between the groups. For example, the parents of the individuals in the three experimental groups tend to be more well-educated than those in the control group. Also, it appears that a much larger fraction of control individuals than experimental individuals are nonwhite, and that the control individuals tend to have more siblings than those individuals in the experimental groups. The average years of education and high school graduation rates for the four groups are provided in the following table.

	Control Avg.	Group One Avg.	Group Two Avg.	Group Three Avg.
YEARS_ED	11.71	12.00	12.10	12.31
HS_GRAD_RATE	63.3%	72.5%	81.0%	76.9%

In order to make reasonable estimates regarding the effect of the rise in income which the experimental individuals experienced, one must first adjust for the systematic differences between the groups.

5.3 Loosening the Constraints on the Control Group

One background variable on which the experimental groups appear to be very different from the control group is NONWHITE. Approximately half of the individuals in each of the experimental groups are nonwhite, whereas nearly all of the control individuals are nonwhite. In fact, only three of the seventy-nine members of the control group are white, which means that exact matches on race will be impossible with any of the three experimental groups. Therefore, even after matching between the control and experimental groups has taken place, substantial bias will remain on this background variable.

One possible way to deal with this problem is to include more individuals in the control group. In order to do this, the criteria for inclusion in the control group are relaxed. Previously, each individual in the control group had a household income which was below the poverty line for *all* ten years of interest. This constraint has the desirable property of constructing a control group which is very "tight", in the sense that none of the individuals had risen out of poverty while they were between six and fifteen years old. But, if one is willing to relax this constraint, better matches for the experimental individuals on the background variables will be found.

Of the 1705 sample individuals in the PSID sample, 314 have an *average* income-to-needs ratio which is below one. Though we could construct a control group which includes all 314 of these individuals, many of them were not exposed to poverty in the early years, a characteristic which all of the individuals in experimental groups one, two, and three share.

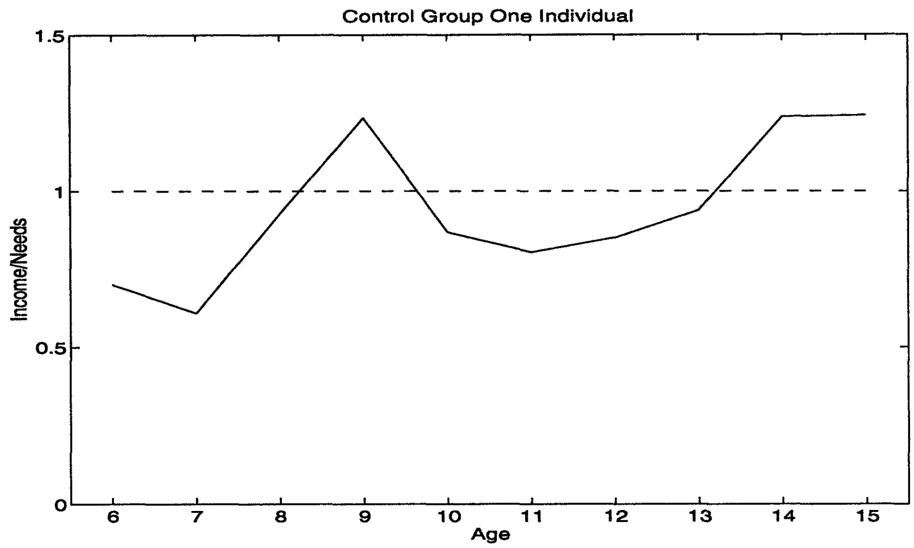

Figure 5-2: Income/Needs Data for a New Control Individual

A more appropriate control group includes those individuals who were exposed to poverty during the early years, and whose household income was *on average* below the poverty line during the ten years of interest. Therefore, all of the individuals considered in the matched comparisons would share the characteristic of *early childhood exposure to poverty.*

A control group for experimental group one is composed of every individual whose household income was below the poverty line when he/she was six years old, and whose household income was *on average* below the poverty line during the ten years of interest. There are 264 individuals whose longitudinal income patterns satisfy these criteria. An example of an income pattern which satisfies these criteria but did not satisfy the original ones is shown in figure 5-2. Similarly, the control group for the second experimental group contains those sample individuals who are exposed to poverty when they are six *and* seven years old, and also have an average household income which falls below the poverty line. There are 222 individuals who satisfy these criteria, all of whom were also included in the first control group. Finally, the third control group is composed of the 198 individuals who are exposed to poverty during their sixth, seventh, and eighth years and whose households' average income-to-needs ratios are below one. Unweighted averages for the three experimental-control group pairs are provided in the following table. By relaxing the constraints for inclusion in the control groups, the number of white control individuals with whom to match the white experimental individuals has markedly increased. There are

25, 18, and 14 white individuals in control groups one, two, and three. The corresponding number in the experimental groups are 21, 10, and 6. If one wishes to match exactly on race though, then the matches for the white experimental individuals may not be particularly good, mainly because there are relatively few candidate control individuals with whom to match each white experimental individual.

Matching Variable	Control One	Exper. One	Control Two	Exper. Two	Control Three	Exper. Three
DAD_ED	.053	.250	.041	.143	.046	.154
MOM_ED	.159	.450	.162	.429	.141	.462
ONE_PARENT_68	.470	.175	.487	.333	.490	.539
NO_PARENT_68	.068	.000	.068	.048	.071	.000
FIRSTBORN	.129	.250	.135	.191	.126	.077
FEMALE	.557	.450	.568	.619	.586	.462
NONWHITE	.905	.475	.919	.524	.929	.539
YR_SIX_INC_NEEDS	.587	.705	.571	.763	.565	.615
YR_SEVEN_INC_NEEDS	—-	—-	.592	.722	.582	.778
YR_EIGHT_INC_NEEDS	—-	—-	—-	—-	.628	.752
WEIGHT	7.33	14.70	6.85	12.52	6.32	13.54
AVG_NUM_SIBS	3.69	2.79	3.69	2.73	3.76	2.62
EARLY_SEPS	.039	.033	.036	.032	.037	.026
EARLY_MOM_WORKED	.428	.658	.411	.603	.409	.333
EARLY_HEAD_DISABLED	.354	.133	.345	.206	.347	.282
EARLY_WITH_ONE	.605	.217	.613	.508	.635	.692
YRS_SOUTH	7.27	7.00	7.39	6.19	7.28	3.85
Number of Individuals	264	40	222	21	198	13
YRS_ED	11.96	12.00	11.96	12.10	11.92	12.31
HS_GRAD_RATE	70.5%	72.5%	71.6%	81.0%	70.7%	76.9%

5.4 The Three Experimental Groups

Given that the PSID sample contains information about 1705 individuals, the total number of children in experimental groups one, two, and three is relatively small at 74. Nevertheless, because their yearly income during the ten year period follows precisely the trajectory whose effect we hope to estimate, it is reasonable to focus on these 74 individuals in a first-pass analysis. The number of individuals in the third experimental group is, at thirteen, particularly low. Finding statistically significant results for such a small number of individuals will require a substantial difference in the outcomes for the matched treated and control individuals. Because the sample size of this experimental group is so small, this group of individuals will not be considered first.

Though the first experimental group has more individuals than either of the other two groups, it has an important disadvantage which must be considered. The children in the first experimental group were only exposed to poverty for the first of the ten years of interest, so it may be the case that the child's sixth year was a peculiarly bad one for his/her family. In other words, this year may simply represent an aberration in the individual's childhood, for his/her family's income may only have temporarily dipped below the poverty line. If this is indeed the case for a particular individual, then this person was quite unlike the individuals with whom he is being matched in the five years before yearly income data is available. To get an idea of the number of individuals who experience one-year dips in household income like this, one can examine how many of the other PSID individuals dipped below the poverty line for *exactly one* of the ten years of interest. This information is provided in the following table.

The *One* Year Below the Poverty Line	Number of Individuals
Six	40
Seven	18
Eight	13
Nine	19
Ten	12
Eleven	14
Twelve	9
Thirteen	11
Fourteen	11
Fifteen	19

The average number of individuals who experience one-year dips for the other nine years is fourteen, which represents approximately one-third of the forty in the first experimental group. So, many of the individuals in this experimental group may really be from non-poor families who only temporarily fell into poverty. Because a substantial fraction of this group may not have been consistently poor during the first several years of life, the first experimental group is not ideal in its characteristics, and will not be the first group to be considered in the empirical analysis.

By the process of elimination, only one of the three experimental groups now remains. The second experimental group has a sample size which is above twenty and has individuals who are more likely to have been poor in the years before yearly income data is available. The number of individuals who temporarily dip below the poverty line for *any other two consecutive* years (and are above the poverty line for all other eight years) is, on average, three. Whereas for experimental group one, the fraction of "false poor" (people who only temporarily dropped into poverty) was estimated to be greater than one-third (or 14/40), for group two the best estimate is approximately one-seventh (or 3/21). Since the second group does not possess undesirable qualities to the extent which the other two groups do, I will consider it first in the upcoming empirical analysis.

5.5 Propensity Score Matching With The Second Groups

With nearest available propensity score matching, the individuals in the experimental group are randomly ordered. Then, the first individual in the experimental group is paired with the control individual with the closest *LogOdds* value. Both of these individuals are eliminated from their respective lists. Then, the second experimental individual is paired with the control individual *still in the list* with the closest *LogOdds* value, and so on. The distance between two individuals is defined in terms of the *LogOdds* instead of the propensity score because the distribution of *LogOdds* is much more nearly normally distributed than is the propensity score (whose values are compressed between zero and one).

Using the relevant background variables which were described earlier in this chapter, logistic regression was performed on the second treated and control groups. Maximum likelihood estimates for the true coefficients of the background variables are included in the following table, along with the corresponding t-statistics. Also included in this table is the standardized difference between the treated and control groups before the matching takes place, and between the *matched* treated and control groups.

Variable	StdDiffBefore PropScoreMatch	StdDiffAfter PropScoreMatch	Beta	t-stat
DadEd	35.3	14.4	1.9900	1.68
MomEd	60.1	19.2	0.5431	0.78
OneParent	-31.1	0.0	-1.9730	-1.86
NoParent	-8.5	0.0	-0.5889	-0.42
FirstBorn	14.8	-22.0	0.2286	0.29
Female	10.4	-9.7	0.9981	1.51
NonWhite	-96.3	0.0	-4.1011	-3.36
YearSixIncNeeds	88.6	10.8	2.9655	1.74
YearSevenIncNeeds	60.7	18.7	2.4092	1.48
Weight	47.1	-12.1	-0.0952	-2.06
AvgNumSiblings	-52.9	-16.7	-0.3900	-1.83
EarlySeps	-4.1	18.1	-2.3882	-0.72
EarlyMomWorked	46.8	32.7	2.4481	2.85
EarlyDisabled	-36.1	-25.4	0.1730	0.20
EarlyWithOnePar	-21.9	33.2	0.7019	0.72
YrsInSouth	-25.8	-12.1	-0.0806	-1.00
LogOdds	163.6	48.8		
Constant			-1.9601	-0.73
AverageStdDiff	47.3	17.3		
ExperimentalYrsEd	12.10	12.10		
ControlYrsEd	11.94	12.15		
ExperimentalHSGradRate	81.0	81.0		
ControlHSGradRate	71.6	81.0		

For a particular background variable X, the standardized difference between the two groups is calculated as follows:

$$StandardizedDifference = 100.0 * \frac{\bar{X}_{experimental} - \bar{X}_{control}}{\sqrt{(\sigma^2_{treated} + \sigma^2_{control})/2.0}} \% \tag{5.2}$$

With nearest available propensity score matching, the average standardized difference on the background variables (including *LogOdds*) has dropped from 47.3% to 17.3%. A substantial fraction of the mean difference along *LogOdds*, perhaps the most important of all of the matching variables, has been eliminated. Particularly large reductions in initial bias have taken place along the *Nonwhite* and *YearSixIncNeeds* variables. However, the biases for other variables, including *EarlyWithOnePar* and *EarlySeps*, have actually increased. The residual differences on a number of the variables are quite high and make some additional adjustments for these variables desirable. The high school graduations rates are identical for the matched treated and control groups, each of which has 21 individuals. The small difference in the years of education between the two groups is not statistically significant.

5.6 Ignoring "Unmatchable" Experimental Individuals

Because the coefficient for the *Nonwhite* variable is particularly large, a white and black individual would have to be very different on other background variables in order to be matched. Therefore, exact matches on the *Nonwhite* variable were performed in the previous matching, with the hope that such different individuals would not be paired with one another. The large residual difference along the *LogOdds* variable is mainly due to the relatively small number of control individuals with whom to pair the white experimental individuals. Four stem and leaf plots for the *LogOdds* variable are provided in the figure 5-3 to show the extent of this problem.

	WhiteExper.	*WhiteControl*	*NonwhiteExper.*	*NonwhiteControl*
−9				4
−8				5332
−7				998443221
−6				86665555444443332111100
−5				9988888777766666555533332221111000000
−4				99999888877777666655544333222222111111000000
−3		9540	5300	9998887776666655544433333222111110000
−2		9730	77	999887777665555544221110
−1	3	95411	82	987766544433322111000
−0	4	753	82	9877
0	234488	28	3	
1	4			
2				
3				
4				
5				
6				
7	1			

Figure 5-3: Stem-and-Leaf Plots for Second Experimental and Control Groups

These plots show how very dissimilar the distributions of the *LogOdds* variable are for white and nonwhite individuals. As evidenced by the two plots to the left of the figure, some of the matches between the control and experimental groups for white individuals cannot be very close on their *LogOdds* values. This situation is quite different from the one for the experimental nonwhite individuals. There appear to be many possible matches for all but two of these individuals.

The one white experimental individual with a *LogOdds* value of 7.1 appears to be very different from *any* other individual, and thus will not be included in a direct comparison between treated and control groups. Similarly, the *LogOdds* value of 1.4 for an experimental individual is substantially above that of any control individual. After eliminating these two experimental individuals, there are eight white experimental individuals remaining for comparison with control individuals. Though the matches on the *LogOdds* variable will not be perfect, these eight can be paired with the eight white control individuals with the highest *LogOdds* values. If one also ignores the two nonwhite experimental individuals with the highest *LogOdds* values, then there exist many "close" possible controls for each of the nine who will remain.

After eliminating these four individuals from the matched comparisons, the residual difference along the *LogOdds* variable drops by a substantial amount; the standardized difference is less than half of what it was with the four individuals included. Surprisingly, though, the *average* residual difference on the background variables increases by more than 5%. The following table illustrates this.

Variable	StdDiffBefore PropScoreMatch	StdDiffAfter PropScoreMatch	StdDiffAfter Eliminating Four
DadEd	35.3	14.4	0.0
MomEd	60.1	19.2	23.4
OneParent	-31.1	0.0	24.1
NoParent	-8.5	0.0	0.0
FirstBorn	14.8	-22.0	-39.6
Female	10.4	-9.7	-23.4
NonWhite	-96.3	0.0	0.0
YearSixIncNeeds	88.6	10.8	-12.2
YearSevenIncNeeds	60.7	18.7	23.0
Weight	47.1	-12.1	-4.1
AvgNumSiblings	-52.9	-16.7	-5.2
EarlySeps	-4.1	18.1	50.1
EarlyMomWorked	46.8	32.7	30.3
EarlyDisabled	-36.1	-25.4	-22.0
EarlyWithOnePar	-21.9	33.2	72.1
YrsInSouth	-25.8	-12.1	-28.1
LogOdds	163.6	48.8	22.9
AverageStdDiff	47.3	17.3	22.4
ExperimentalYrsEd	12.10	12.10	12.18
ControlYrsEd	11.94	12.15	12.18
ExperimentalHSGradRate	81.0	81.0	82.4
ControlHSGradRate	71.6	81.0	82.4

Despite the large reduction in the residual difference along what is arguably the most important matching variable, the increased standardized differences on some of the other variables, most notably *EarlyWithOnePar* and *EarlySeps*, are bothersome. These large systematic differences make an alternative approach desirable.

5.7 Mahalanobis-Metric Matching With the Second Groups

By constructing matched treated and control groups with mahalanobis-metric matching, one places greater emphasis on the values of *all* of the important background variables when assessing the appropriateness of a particular match. Thus, one might intuitively expect the residual bias on variables such as *EarlySeps* and *EarlyWithOnePar* to decrease when using this method. Because the *LogOdds* variables is perhaps the most important matching variable, it will be used in determining which control individuals are "candidates" for a particular experimental child. More specifically, one can define a set of potential controls whose *LogOdds* values are close to a particular experimental individual's, and then select from these the one whose Mahalanobis distance from the experimental child is smallest. The steps of this procedure are as follows:

1. Randomly order the experimental individuals.
2. Define candidate controls for the first treated child by *caliper matching* on the *LogOdds* variable. Find all available control individuals whose *LogOdds* values are within some absolute distance C of the experimental individual's *LogOdds* value. If there are no such control individuals, then simply select the control individual with the closest *LogOdds* value.
3. Mahalanobis metric matching within calipers: From the candidate controls, select as the match the individual whose Mahalanobis distance from the experimental child is smallest.
4. Remove the experimental child and his/her match from the list, and then go to step two for the next treated child.

Different caliper widths were analyzed by Cochran and Rubin (1973) to assess the reductions in bias for each. Let σ_E and σ_C equal the standard deviations of the *LogOdds* variable in the experimental and control groups, respectively, and let $\sigma = \sqrt{(\sigma_E^2 + \sigma_C^2)/2.0}$. In the hope of removing at least 90% of the bias on the background variables, they suggest a caliper width less than or equal to $c = .25\sigma$.

The mahalanobis distance between two individuals, which was defined in the previous chapter, will depend on how the two differ with respect to the *seventeen* background variables. These background variables include *LogOdds* and the sixteen for which coefficients were estimated in the logistic regression. Nearest available mahalanobis-metric matching within calipers defined by the *LogOdds* variable (and with a caliper width of $c = .15\sigma$) substantially reduced the average residual difference along the seventeen background variables for the matched treated and control groups. The following table illustrates this.

Variable	StdDiffBefore PropScoreMatch	StdDiffAfter PropScoreMatch	StdDiffAfter Eliminating Four	StDiffAfter Mahalanobis
DadEd	35.3	14.4	0.0	20.2
MomEd	60.1	19.2	23.4	11.5
OneParent	-31.1	0.0	24.1	0.0
NoParent	-8.5	0.0	0.0	0.0
FirstBorn	14.8	-22.0	-39.6	0.0
Female	10.4	-9.7	-23.4	-11.5
NonWhite	-96.3	0.0	0.0	0.0
YearSixIncNeeds	88.6	10.8	-12.2	21.1
YearSevenIncNeeds	60.7	18.7	23.0	-12.4
Weight	47.1	-12.1	-4.1	0.4
AvgNumSiblings	-52.9	-16.7	-5.2	-4.6
EarlySeps	-4.1	18.1	50.1	0.0
EarlyMomWorked	46.8	32.7	30.3	25.2
EarlyDisabled	-36.1	-25.4	-22.0	11.7
EarlyWithOnePar	-21.9	33.2	72.1	16.4
YrsInSouth	-25.8	-12.1	-28.1	-41.4
LogOdds	163.6	48.8	22.9	43.7
AverageStdDiff	47.3	17.3	22.4	12.9
MaximumStdDiff	163.6	48.8	72.1	43.7
ExperimentalYrsEd	12.10	12.10	12.18	12.18
ControlYrsEd	11.94	12.15	12.18	11.94
ExperimentalHSGradRate	81.0	81.0	82.4	82.4
ControlHSGradRate	71.6	81.0	82.4	82.4

As suggested by the results listed above, the third method seems superior to the other two in reducing bias along the sixteen background variables. For example, the standardized difference has been completely eliminated for several of the covariates, and the average standardized difference is approximately 75% less than it originally was. Unfortunately, the difference along the propensity score remains quite high, and thus the matches are still not as good as we would like them to be.

The "best" of the three matching methods, in terms of the average standardized difference on important background variables, appears to be Mahalanobis-metric matching.

Though the high school graduation rates are identical for the two groups, one can use a difference-of-means test to see if the difference on the *YearsEd* variable is statistically significant. When two populations are not normally distributed and the sample sizes from the two populations are sufficiently large (fifteen to twenty for each group is a reasonable lower bound), one can invoke the Central Limit Theorem and assume that the difference in sample means is normally distributed. Assume that, for the "populations" of experimental and control individuals, the actual average years of education are μ_e and μ_c and that the corresponding variances in the population are σ_e and σ_c. With the sample means for a variable X defined as $\bar{X}_e$ and $\bar{X}_c$, one then can calculate the Z-statistic and test for the presence of a statistically significant difference of means.

$$Z = \frac{(\bar{X}_c - \bar{X}_e) - (\mu_c - \mu_e)}{\sqrt{\frac{\sigma_c^2}{n_c} + \frac{\sigma_e^2}{n_e}}} \tag{5.3}$$

For a one-sided difference of means test with unknown population variances, one would test the null hypothesis H_0 that μ_c is greater than or equal to μ_e by setting the $(\mu_c - \mu_e)$ term equal to zero and by estimating the population variances from the available data. This yields the following formula for calculting the Z-statistic.

$$Z = \frac{(\bar{X}_c - \bar{X}_e)}{\sqrt{\frac{s_c^2}{n_c} + \frac{s_e^2}{n_e}}} \tag{5.4}$$

If the value of Z is less than -1.645, then the null hypothesis is rejected. For example, a result of Z = -2.1 would provide statistically significant evidence that the true population mean for the experimental group is greater than that for the control group. For the groups matched using the Mahalanobis distance, the Z-statistic is only -0.52. Therefore, the difference in sample means is not statistically significant. If there were a higher ratio of potential controls to experimental individuals *and* if there were more experimental individuals, then better matches would be possible and statistically significant results would not require such a large difference in outcomes between the two groups.

It is important to note that, if a few of the experimental individuals are only temporarily visiting poverty during their sixth and seventh years (in other words, they are "false poor" individuals), then it is likely that the estimate of educational benefit from a rise in income is *biased upwards*. The average years of education for individuals who are, on average, non-poor in the PSID sample is 12.74. In a previous calculation, we estimated the fraction of "false poor" individuals in experimental group two to be approximately one-seventh. If one assumes that, consistent with this earlier estimate, three of the remaining group two individuals are actually non-poor and received the 12.74 years of education, then the average years of education for the remaining fourteen is 12.06. Thus, the estimated treatment effect for the groups matched using the Mahalanobis distance has dropped by 50%, from 0.24 years to 0.12 years. This first-pass adjustment suggests that the estimated treatment effects may be misleading if some of the experimental individuals are not actually poor before their apparent climb out of poverty.

5.8 Empirical Analyses With The First Experimental Group

The set of matching variables for the first experimental group is the same as those for the second group except the *YrSevenIncNeeds* variable is not included. As before, logistic regression is performed on the treated and control groups, and maximum-likelihood estimates for the coefficients of the background variables are obtained. Stem-and-leaf plots for the distribution of the *LogOdds* variable in the experimental and control groups are provided in figure 5-4. Once again, these are broken up into distributions for both white and nonwhite individuals. There appear to be a number of experimental individuals for whom there are no good matches. This is particularly true for the white experimental group. If one ignores for the moment the one control individual with a *LogOdds* value of 3.1, there exist eight control individuals with *LogOdds* values which are greater than any of the control individuals have. Therefore, after matching all forty individuals using nearest-available propensity score matching, these eight experimental individuals are ignored for the subsequent analyses. The one white control individual with a *LogOdds* value of 3.1 appears to be so much different from the rest of the control individuals that he/she will also be ignored.

	WhiteExper.	*WhiteControl*	*NonwhiteExper.*	*NonwhiteControl*
−7				640
−6				886543221100000
−5				8887665544333222111100000
−4		21		9999888777776666555555554444444443333333222221100000000000
−3		955	3	999998888887777777777777666666666655555554443333333333322222111110000
−2		8610	5521	999999887777766666555544443333333333322111100000
−1	96	87211	9775552	99987765555444333333
−0	641	64322	44	88611
0	01456789	2236	0237	0026
1	344489	2	2	
2				
3	1	1		
4	0			

Figure 5-4: Stem-and-Leaf Plots for First Experimental and Control Groups

Because none of the individuals in the experimental group have a value of one for the variable *NoParent*, any person who does have a value of one is obviously from the control group. Thus, a maximum likelihood estimate for the true value of the coefficient of *NoParent* converges, on successive iterations, to negative infinity. Exact matches are therefore required on this variable, and its coefficient is not estimated in the final logistic regression. The table below provides the coefficients and the corresponding t-statistics for the remaining fourteen matching variables. In addition to this, five sets of standardized differences in background variables are included. The first represents the initial standardized difference between the treated and control groups. Then, results are given for the three matching methods (with exact matches on race) which were described for the second experimental group in the previous chapter. The fifth set of standardized differences is provided because, by exact matching on gender instead of race with mahalanobis-metric matching, the average standardized difference is reduced and only one of the thirty-two experimental individuals is not exactly matched on race.

Variable	Beta	t-stat	StdDiffBefore PropScoreMatch	StdDiffAfter PropScoreMatch	StdDiffAfter Eliminating 8	StDiffAfter Mahal RACE	StDiffAfter Mahal FEMALE
DadEd	0.6552	1.01	56.5	32.0	8.9	8.9	18.7
MomEd	0.4626	0.92	66.0	36.6	13.0	26.8	13.0
OneParent	-0.6817	-0.88	-66.1	-23.8	-7.7	0.0	-7.7
NoParent	——	—	-38.2	0.0	0.0	0.0	0.0
FirstBorn	0.0037	0.01	31.0	0.0	-14.2	7.7	0.0
Female	-0.3090	-0.68	-21.3	-9.9	-24.9	-24.9	0.0
NonWhite	-2.3988	-3.26	-104.1	0.0	0.0	0.0	-6.3
YearSixIncNeeds	1.1947	1.72	34.8	-5.7	-18.9	-10.0	-10.4
Weight	-0.0456	-1.61	58.2	-0.7	7.4	0.7	12.3
AvgNumSiblings	-0.2268	-1.43	-50.9	-21.9	-6.1	-0.4	-3.1
EarlySeps	-0.6101	-0.25	-5.5	31.0	44.8	0.0	0.0
EarlyMomWorked	2.1570	3.46	55.9	51.2	36.8	10.2	2.6
EarlyDisabled	-1.5421	-2.07	-58.9	-5.5	14.8	-10.3	3.7
EarlyWithOnePar	-2.2445	-2.74	-91.0	-12.3	7.6	2.6	2.6
YrsInSouth	-0.1531	-2.47	-6.0	4.4	4.7	-12.7	-27.2
Constant — LogOdds	1.6999	1.22	197.5	56.6	13.2	22.7	16.6
AverageStdDiff			58.9	17.2	13.9	8.6	7.8
MaximumStdDiff			197.5	56.6	44.8	26.8	27.2
ExperimentalYrsEd			12.00	12.00	12.00	12.00	12.00
ControlYrsEd			11.96	11.80	11.84	12.03	11.78
Exper HSGradRate			72.5	72.5	71.9	71.9	71.9
Cont HSGradRate			70.5	75.0	75.0	75.0	71.9

Once again, the Mahalanobis-metric matching seems to do a much better job of adjusting for residual differences on the background variables. After eliminating the eight individuals for whom there were no good matches, mahalanobis-metric matching (with an exact match on the *Female* variable) reduces the average standardized difference on the background variables by 87% (from 58.9% to 7.8%). This represents a substantial improvement over nearest-available propensity score matching with the subsample of 32 experimental individuals. For this method, the average standardized difference is 13.9%.

In fact, it appears that all three matching methods have, for this first experimental group, yielded larger reductions in the standardized differences along background variables than they did for the second experimental group. One possible reason for this is that, because there are more matches for this group (forty and thirty-two as opposed to twenty-one and seventeen), one or two particularly poor matches in the background variables is less likely to have a substantial effect. As for inferences about the effect of the income infusion on the academic achievement of the individuals in the experimental group, none of the differences in high school graduation rates or years of education are statistically significant. For the Mahalanobis-metric matched groups with an average standardized difference of 7.8%, a one-sided difference of means test can be conducted. The Z-statistic is calculated to be -0.63, which is not sufficiently negative to provide statistically significant evidence of an actual difference in population means.

Once again, it is important to bear in mind that many of the experimental children may not actually be poor. Assuming that 11 of the 32 are non-poor (approximately one-third, which is the estimate from above) and that each of these children received the 12.74 years of education which the average non-poor PSID child received, then the average years of education for the remaining twenty one is 11.61. Therefore, the estimated educational effect of the income infusion has fallen from 0.22 years to -0.17 years (neither of which are statistically significant). Though this represents only a first-pass adjustment, it reveals how sensitive the estimated treatment effects may be to the presence of "false poor" individuals.

5.9 Empirical Analyses With The Third Experimental Group

The last of the three experimental groups is composed of those individuals whose household incomes were below the poverty line for the first three years and above for the next seven. Though the small sample size of this group means that statistically significant results will be difficult to come by, the ratio of control to experimental individuals is greater for this group than it is for the other two. Thus, one might expect that poor matches are less likely for this experimental group.

Maximum-likelihood estimates for the coefficients of the background variables are obtained using logistic regression. Because there are no experimental individuals with a value of one for the variable *NoParent*, it is excluded from the logistic regression and exact matches are required for it. The four stem-and-leaf plots for the distribution of the *LogOdds* variable are provided in figure 5-5.

	WhiteExper.	*WhiteControl*	*NonwhiteExper.*	*NonwhiteControl*
−9				530
−8				8743221100
−7				98776665544443321100
−6				9988887776665444433211000O
−5				99987777766655544333322221100
−4				9998777776655554333322110O
−3		5311	62	999887777766666655544444333222111100
−2		63	95	99998888887666444300
−1	4	765431	74	8888776643200
−0	1	30	2	752
0	0			
1	17			
2	3			

Figure 5-5: Stem-and-Leaf Plots for Third Experimental and Control Groups

Again we see that there are no good matches along the propensity score for some of the white experimental individuals. It appears that all seven of the nonwhite members of the experimental group have at least one nonwhite individual in the control group who is close on the propensity score. After nearest-available propensity score matching with all thirteen members of the experimental group, the three experimental children with the highest *LogOdds* values are ignored in the subsequent matched comparisons. Estimates for the coefficients, the corresponding t-statistics, the initial standardized differences between the two groups, and the standardized differences after using all three matching methods for the background variables are included in the table below. A few variables were successively eliminated from the logistic regression because their t-statistics were particularly low and because of the small number of individuals in the experimental group.

Variable	Beta	t-stat	StdDiffBefore PropScoreMatch	StdDiffAfter PropScoreMatch	StdDiffAfter Eliminating 3	StDiffAfter Mahal
DadEd	0.7910	0.51	35.7	57.9	0.0	0.0
MomEd	1.5639	1.93	72.4	48.0	20.0	0.0
OneParent	0.7532	0.88	9.5	-15.0	0.0	0.0
NoParent	——	—-	-38.9	0.0	0.0	0.0
FirstBorn	-1.9735	-1.57	-16.1	-41.9	-26.8	-26.8
Female	-0.8531	-1.11	-24.6	-14.8	-19.2	0.0
NonWhite	-2.4428	-2.80	-95.5	0.0	0.0	0.0
YearSixIncNeeds	-3.0810	-1.65	19.7	3.5	-10.1	7.8
YearSevenIncNeeds	6.2583	2.79	105.6	24.4	-4.0	0.0
YearEightIncNeeds	——	—-	67.3	24.1	-11.2	52.9
Weight	——	—-	65.3	-17.1	-6.4	18.1
AvgNumSiblings	-0.6849	-2.34	-79.6	35.6	52.8	18.8
EarlySeps	——	—-	-11.2	39.2	44.7	44.7
EarlyMomWorked	-0.8940	-0.83	-18.6	-18.8	-31.3	-8.6
EarlyDisabled	-0.5455	-0.57	-15.6	35.0	60.5	-7.2
EarlyWithOnePar	——	—-	12.1	0.0	0.0	20.0
YrsInSouth	——	—-	-72.8	-58.4	-35.9	-35.9
Constant — LogOdds	-1.0462	-0.46	185.4	45.2	1.9	1.9
AverageStdDiff			52.6	26.6	18.0	13.5
MaximumStdDiff			185.4	58.4	60.5	52.9
ExperimentalYrsEd			12.31	12.31	12.00	12.00
ControlYrsEd			11.92	11.46	11.60	10.90
Exper HSGradRate			76.9	76.9	70.0	70.0
Cont HSGradRate			70.7	69.2	70.0	50.0

For the third experimental group, Mahalanobis-metric matching does the best job of reducing the standardized differences on the background variables of interest, and there remains almost no difference along the *LogOdds* variable for this method. The percentage reduction in the average standardized difference is, at 74% (from 52.6% to 13.5%), not as large as it was for the previous group. This may be due to the relatively small number of experimental individuals which remain after the "unmatchable" individuals are eliminated. Though the number of matched comparisons is too low to reasonably conduct a difference of means test, the outcomes do suggest a modest academic benefit from the rise in income after the child's eighth year.

5.10 Initial Inferences

None of the ten matched comparisons (three for the second groups, four for the first groups, three for the third groups) yield statistically significant differences in academic achievement as measured by the number of years of education or the high school graduation rates. However, by taking a quick glance at the matched comparison from each group with the lowest average standardized difference on *all* relevant background variables, one sees a trend which suggests a modest benefit from the rise in income. One must bear in mind that, if some of the experimental children are actually non-poor before their rise in income (except for a temporary drop below the poverty line), then these estimated effects are biased upwards. The table below summarizes these results.

Groups	AverageStdDiff	ExperimentalYrsEd	ControlYrsEd	ExperHSGradRate	ContHSGradRate
One	7.8%	12.00	11.78	71.9	71.9
Two	12.9%	12.18	11.94	82.4	82.4
Three	13.5%	12.00	10.90	70.0	50.0

Because the number of individuals in each of the three experimental groups is so low, getting statistically significant results from the matched comparisons requires a large difference in the outcome variables. Though none of the differences in the years of education outcome are sufficiently large, the fact that all three point in the same direction suggests that there may well be a small educational benefit from the increase in income. A more accurate assessment of the true treatment effect could perhaps be obtained by adjusting for residual bias on the background variables. Of course, a larger set of data would also aid in establishing a more reliable estimate of the true effect of this rise in income.

5.11 Regression Adjustment to Reduce Residual Bias

For all three pairs of groups, the standardized difference for a few of the variables remain quite high. As a result, the estimated "treatment effects" may well be biased. Regression adjustment can help to control for the bias which remains after matching has taken place.[38, p. 185]. Because the matched individuals are relatively "close" to one another on the distribution of background variables, assumptions about the linear dependence of some outcome variable on a number of background variables is not as unreasonable as it would be for a much more diverse data set.

For the second pair of groups, the standardized difference on the variables *DadEd, YrSixIncNds, EarlyMomWorked, YrsSouth,* and *LogOdds* is greater than 20%. Though there are other background variables which contain residual bias, with only 17 matched pairs not all of the variables can be included in the regression adjustment. When performing least-squares regression to adjust for residual bias the hope is that, by adjusting for those variables with the highest standardized differences, the standardized difference for other variables will also be reduced. Let $(x_{1e}, x_{2e}, x_{3e}, x_{4e}, x_{5e})$ equal the vector of background variables for an experimental individual and y_e be the corresponding years of education for him/her. Similarly, allow $(x_{1c}, x_{2c}, x_{3c}, x_{4c}, x_{5c})$ and y_c to equal a control individual's covariate vector and outcome variable, respectively. By making least squares estimates of the coefficients in the following formula, one can gain an estimate of the average *treatment effect* β_0.

$$(y_e - y_c) = \beta_0 + \beta_1 * (x_{1e} - x_{1c}) + \beta_2 * (x_{2e} - x_{2c}) + \beta_3 * (x_{3e} - x_{3c}) + \beta_4 * (x_{4e} - x_{4c}) + \beta_5 * (x_{5e} - x_{5c}) \quad (5.5)$$

Multiple regression for seventeen matched pairs yields the following equation, which suggests a treatment effect which is positive but which is not statistically significant (t-statistics for each of the coefficients are included in parentheses).

$$\begin{array}{llll} \Delta YearsEd = & 0.471 & -3.241 * \Delta DadEd & -0.529 * \Delta YrSixIncNds \\ & (0.63) & (-1.37) & (-0.16) \\ & -0.670 * \Delta EarlyMomWork & -0.01394 * \Delta YrsSouth & +0.0253 * \Delta LogOdds \\ & (-0.49) & (-0.15) & (0.03) \end{array}$$

Though the estimated effect of a sustained income infusion above the poverty line after a child's seventh year is not statistically significant, the least-squares estimate of it, after adjusting for those variables with the greatest residual bias between the two groups, has a value of 0.471 (nearly a half of a year of education). This estimate is actually greater than the previous one for these two groups, which yielded an estimated treatment effect of .235 years of education.

Similar regression adjustments can be performed for the first and third experimental groups. For the first matched groups, the largest standardized differences can be found on the variables *DadEd, MomEd, Weight, YrsSouth,* and *LogOdds.* Proceeding as before, least-squares estimates of the corresponding coefficients are obtained, and the results are provided in the following equation.

$$
\begin{array}{llll}
\Delta YearsEd = & 0.033 & +0.473 * \Delta DadEd & +0.714 * \Delta MomEd \\
 & (0.08) & (0.51) & (1.15) \\
 & +0.0164 * \Delta Weight & -0.0426 * \Delta YrsSouth & +0.158 * \Delta LogOdds \\
 & (0.75) & (-0.49) & (0.24)
\end{array}
$$

Once again, the estimated effect on the years of education which a child receives is not statistically significant. Though the sign is consistent with the result before regression adjustment took place, the least-squares estimate of β_0 is quite small.

As the number of matched pairs decreases, the number of least-squares coefficients which can reasonably be estimated also goes down. Thus it is the case that, for the third experimental and control groups, only three variables will be included in the least-squares regression. The standardized differences for *YrEightIncNds, EarlySeps,* and *YrsSouth* are highest, and thus these variables will be used in the least squares regression analysis.

$$
\begin{array}{lllll}
\Delta YearsEd = & 0.750 & +12.60\Delta YrEightIncNds & -2.87\Delta EarlySep & +0.302\Delta YrsSouth \\
 & (0.83) & (1.61) & (-0.30) & (1.49)
\end{array}
$$

Here, the least-squares estimate of the treatment effect is, as before, not statistically significant. Nevertheless, the sign of the effect is positive, which is consistent with the result obtained before regression adjustment took place.

The table below summarizes the results both before and after least-squares regression adjustment for all three pairs of groups. Only the *YearsEd* outcome variable is provided in this comparison between the two sets of results. Also included is the z-statistic for the difference of means in the *YrsEd* variable before regression adjustment is performed. For the z-statistic, a value less than -1.645 corresponds to statistically significant evidence that the actual mean for *YearsEd* in the control population is *less* than the corresponding mean in the experimental population (the group whose income rose and remained above the poverty line). A t-statistic whose absolute value is greater than 1.96 provides statistically significant evidence that the *true* coefficient β_0 is not zero.

Groups	Matched Pairs	AvgStdDiffBefore	ExperYrsEdBefore	ControlYrsEd	z-statistic	β_0	t-statistic
One	32	7.8%	12.00	11.78	-0.63	0.033	0.08
Two	17	12.9%	12.18	11.94	-0.52	0.471	0.63
Three	10	13.5%	12.00	10.90	—	0.750	0.83

Though *none* of the six estimates of the effect of an income infusion on an individual's eventual years of education yield statistically significant results, all of them suggest that the rise in income does have a positive effect. Because of the small number of individuals whose income undergoes the trajectory whose effect we wish to estimate, obtaining statistically significant results requires, in the first case, large differences in means or, in the second case, relatively high estimates of the true value of the coefficient β_0.

A careful look at *both* sets of results reveals the peculiar outcome that the infusion seems to have a greater benefit after the age of eight than it does after the age of six or seven. Similarly, it appears that the infusion after the age of seven has a greater benefit than it does after the age of six. One must bear in mind that, because of the paucity of relevant data, this may simply be due to chance fluctuations around some true values.

Chapter 6

Two Alternative Methods

Though the matching methods described in the previous chapter have the desirable property that they eliminate a substantial amount of the *pre-treatment* bias along various background variables, they force the analyst to ignore all information about most of the individuals in the PSID sample, some of which may contain valuable information for the question which we are trying to answer. In the next couple of sections, two other ways of examining the effect of childhood exposure to poverty on an individual's educational attainment are briefly described, both of which consider more of the PSID sample individuals.

6.1 An Econometric Analysis of all 1705 Individuals

Despite the advantages of restricting the analysis to individuals who satisfy particular income pattern criteria, constructing a model which considers all 1705 sample individuals is not necessarily superfluous. By proceeding in this manner, one can investigate the extent to which the results obtained using matched sampling methods differ from those estimated with traditional econometric techniques. For example, one can conduct a linear regression which has as its dependent variable the number of years of education which a child receives and as its independent variables the background information which was adjusted for above. For the following model, the income measure, known as *AvgEarlyInc*, will simply equal the average income-to-needs ratio during the first three years (when child is six, seven, and eight). The first equation provides all variables included in the initial model. The second equation was obtained by eliminating variables one at a time until no independent variable with a t-statistic below 2.0 remained.

$$
\begin{array}{lcccc}
YearsEd = & \underset{(62.0)}{11.645} & \underset{(6.49)}{+0.637 * DadEd} & \underset{(5.90)}{+0.507 * MomEd} & \underset{(1.26)}{+0.189 * OneParent} \\
 & \underset{(1.53)}{+0.3398 * NoParent} & \underset{(0.28)}{+0.023 * Firstborn} & \underset{(3.80)}{+0.274 * Female} & \underset{(3.60)}{+0.415 * NonWhite} \\
 & \underset{(1.08)}{+0.00389 * Weight} & \underset{(-2.81)}{-0.074 * AvNumSibs} & \underset{(-1.17)}{-0.510 * EarlySeps} & \underset{(0.78)}{+0.070 * EarlyMomWorked} \\
 & \underset{(-2.63)}{-0.3113 * EarlyDis} & \underset{(-2.44)}{-0.3455 * EarlyWithOnePar} & \underset{(-1.2)}{-0.011 * YrsSouth} & \underset{(4.68)}{+0.154 * AvgEarlyInc}
\end{array}
$$

$$
\begin{array}{lcccc}
YearsEd = & \underset{(63.8)}{11.680} & \underset{(6.97)}{+0.649 * DadEd} & \underset{(6.13)}{+0.516 * MomEd} & \underset{(3.41)}{+0.300 * NonWhite} \\
 & \underset{(-2.84)}{-0.331 * EarlyDis} & \underset{(-2.98)}{-0.075 * AvNumSibs} & \underset{(3.95)}{+0.285 * Female} & \underset{(5.69)}{+0.178 * AvgEarlyInc} \\
 & \underset{(-2.37)}{-0.933 * EarlySeps} & & &
\end{array}
$$

From both the first and second equation, one can see that the average income which a child receives during the early years of his/her life seems to have a statistically significant relationship with the number of years of education he/she receives. One problem with a model such as this one is that it assumes a linear relationship between the dependent variable and all of the independent variables. Nevertheless, the result suggests that income does indeed matter, which is consistent with our previous results (which were *not* statistically significant).

An alternative income variable for a model such as the previous one could measure the number of years (out of the first three) during which a given child is exposed to poverty. A variable *EarlyExposed* is constructed which can take on one of four possible values (0, 1, 2, or 3). This variable replaces *AvgEarlyInc* from above in the least-squares regression analysis, which yields the following two equations.

$$
\begin{array}{lcccc}
YearsEd = & 11.96 & +0.751 * DadEd & +0.550 * MomEd & +0.198 * OneParent \\
 & (67.4) & (7.81) & (6.40) & (1.38) \\
 & +0.407 * NoParent & +0.0289 * Firstborn & +0.289 * Female & +0.400 * Nonwhite \\
 & (1.77) & (0.31) & (3.97) & (3.38) \\
 & +0.0063 * Weight & -0.107 * AvNumSibs & -0.552 * EarlySeps & +0.101 * EarlyMomWorked \\
 & (1.71) & (-4.06) & (-1.27) & (1.13) \\
 & -0.358 * EarlyDis & -0.441 * EarlyWithOnePar & -0.016 * YrsSouth & +0.0032 * EarlyExposed \\
 & (-2.93) & (-2.98) & (-1.82) & (.07)
\end{array}
$$

$$
\begin{array}{lcccc}
YearsEd = & 11.93 \quad [t]cl & +0.732 * DadEd \quad [t]cl & +0.581 * MomEd \quad [t]cl & -0.113 * AvNumSibs \quad [t]cl \\
 & (66.8) & (7.89) & (6.90) & (-4.65) \\
 & -0.361 * EarlyDis & -0.769 * EarlySeps & +0.416 * NonWhite & -0.246 * EarlyWithOnePar \\
 & (-3.05) & (-2.01) & (3.57) & (-2.43) \\
 & +0.0079 * Weight & +0.290 * Female & & \\
 & (2.21) & (3.99) & &
\end{array}
$$

Though the previous model showed a statistically significant relationship between income and educational attainment, this one does not show one between the years of education and the number of years (of the first three) a person is exposed to poverty. Because the sign of the *EarlyExposed* variable is positive, this model does *not* suggest that early childhood exposure to poverty has an adverse effect on an individual's eventual educational attainment. This contrasts with the results obtained using the matched sampling techniques, all of which suggested that children who rise out of poverty tend to stay in school for a longer period of time than do similar individuals who remain poor.

The previous econometric models include independent variables such as *EarlyWithOnePar, EarlyExposed,* and *AvgEarlyInc*, which provide information about the first three of the ten years of interest. One could also include information about the final seven years with variables such as *LateWithOnePar, LateExposed*, and *AvgLateInc.* These variables, in addition to *LateDis, LateMomWorked,* and *LateSeps*, contain information about the seven years during which the children's ages were between nine and fifteen. Performing a stepwise linear regression with the appropriate additional independent variables for each of the previous models, one finds that, in both cases, the late income variable is statistically significant whereas the early one is not. The results of both stepwise linear regressions are provided below.

$$
\begin{array}{lcccc}
YearsEd = & 11.82 \quad \text{[t]cl} & +0.589 * DadEd \quad \text{[t]cl} & +0.498 * MomEd \quad \text{[t]cl} & -0.081 * AvNumSibs \quad \text{[t]cl} \\
 & (63.4) & (6.36) & (5.88) & (-3.28) \\
 & +0.145 * AvgLateInc & -1.370 * LateSeps & +0.366 * NonWhite & -0.282 * EarlyWithOnePar \\
 & (6.03) & (-2.10) & (4.09) & (-2.84) \\
 & -0.420 * LateDis & +0.276 * Female & & \\
 & (-3.32) & (3.84) & &
\end{array}
$$

$$
\begin{array}{lcccc}
YearsEd = & 11.97 \quad \text{[t]cl} & +0.718 * DadEd \quad \text{[t]cl} & +0.527 * MomEd \quad \text{[t]cl} & -0.091 * AvNumSibs \quad \text{[t]cl} \\
 & (64.1) & (7.90) & (6.21) & (-3.64) \\
 & -0.066 * LateExposed & -0.980 * EarlySeps & +0.448 * NonWhite & -0.149 * LateSeps \\
 & (-3.30) & (-2.49) & (3.83) & (-2.29) \\
 & -0.410 * LateDis & +0.0081 * Weight & +0.307 * Female & \\
 & (-3.09) & (2.29) & (4.24) &
\end{array}
$$

For both of these econometric models, the income variables concerning the latter seven years are statistically significant, whereas the two concerning the first three years are not. This result is surprising, especially when one considers the belief held by most educationists that *early* childhood exposure to poverty has the most pernicious effect on an individual, whereas inadequate income later in one's childhood is not so harmful. Both models suggest that childhood exposure to poverty adversely affects an individual's educational attainment.

The ability of various econometric methods to accurately determine the effect of employment and training programs on future earnings has been investigated by Lalonde.[25] It appears that, when compared with the results from randomized experiments, many econometric procedures *do not* yield accurate assessments of the effect of a treatment (in this case, a training program) on a particular outcome variable. Because the 1705 PSID sample individuals are very different from one another with respect to observed background characteristics, the assumption that there exists a linear relationship between the dependent variable and the several independent variables should not be readily accepted. Thus, these results are perhaps less defensible than those obtained using multivariate matched sampling techniques.

6.2 Matched Sampling on both Pre- and Post-Treatment Information

When using matched sampling methods in the previous chapter, individuals whose families rose out of poverty were compared with others whose families did not. One may also be interested in comparing those individuals who were temporarily impoverished with others whose household incomes were *consistently above* the poverty line. Then, one may gain a first-pass estimate of the educational "cost" of being exposed to poverty at a relatively early age.

The second experimental group is composed of children who were exposed to poverty at the ages of six and seven and were then unexposed for the next eight years. One can then compare these individuals with others who are *similar with respect to all important background variables*, but who were never exposed to poverty during the ten years of interest. After rising out of poverty, all of the twenty-one experimental individuals have average household income-to-needs ratios (during the final eight years) between one and four. Thus, a control group is constructed of all individuals who are not exposed to poverty and whose average income-to-needs ratio is between one and four. There exist 696 sample individuals whose income patterns satisfy these criteria. As before, logistic regression is used to estimate the coefficients in the propensity score calculation for background variables which are deemed important. Though the following methodology runs counter to the matched sampling method, which emphasized matching only on *pre-treatment* information, it represents a different way to investigate a complicated problem. By matching appropriately, one can compare the educational attainments of individuals who differ only on their poverty status during the first two observed years.

For this analysis, matching variables will include information about the sample children during all ten years of interest. As was previously explained, the variables *EarlySeps, EarlyMomWorked, EarlyDisabled,* and *EarlyWithOnePar* provided information about the three years during which a child's age was between six and eight. Corresponding information is provided for the next seven years in the variables *LateSeps, LateMomWorked, LateDisabled,* and *LateWithOnePar.* Because matching on the income during *all* eight years would result in too many background variables for only twenty-one matched comparisons, two income measures to describe an individual's household income pattern during the eight years are constructed. The first, *LateAvgIncNds*, is simply the average houshold income-to-needs ratio during the final eight years whereas the second, *LateStDevIncNds* is the standard deviation of the household income-to-needs ratio during this time. The table below provides initial standardized differences between the experimental and control groups on twenty one variables. In addition to this, the maximum likelihood estimates for the coefficients of the background variables, with the corresponding t-statistics, are provided. Finally, the standardized differences after employing all three matching methods are given. For the second and third matched comparisons, five of the twenty-one experimental individuals were ignored because there existed no good matches for them.

Variable	StdDiffBefore	StdDiffAfter PropScoreMatch	PrScoreMatch Eliminated 5	Mahalanobis Eliminated 5	Beta	p-value
DadEd	-96.4	-12.5	0.0	0.0	-0.9049	.262
MomEd	-38.7	29.4	26.4	-12.4	-0.2661	.710
OneParent	72.2	44.8	0.0	16.7	0.0000	—
NoParent	17.9	0.0	0.0	0.0	1.3249	.371
FirstBorn	-14.9	12.5	0.0	0.0	0.2886	.691
Female	26.7	0.0	0.0	0.0	0.6814	.256
NonWhite	44.4	0.0	0.0	0.0	-1.4559	.151
EarlySeps	9.4	0.0	35.4	0.0	-5.0793	.154
MidEndSeps	-26.5	0.0	-35.4	0.0	-4.8391	.678
EarlyMomWorked	14.5	-4.1	-5.4	4.9	-0.7392	.405
MidEndMomWorked	29.7	9.3	-0.0	-3.1	2.1475	.066
EarlyDisabled	50.5	-8.8	-11.8	-11.2	2.2049	.131
MidEndDisabled	41.7	-10.3	-2.8	8.8	-1.3542	.393
EarlyWithOnePar	106.2	17.1	-4.4	4.6	6.3183	.000
MidEndWithOnePar	24.8	-14.7	-18.2	-4.2	-5.1253	.000
Weight	-78.8	-5.5	0.9	-7.4	-0.0535	.147
AvgNumSiblings	29.5	16.0	25.2	29.0	0.1964	.343
YrsInSouth	51.1	-27.6	-28.2	-13.5	0.1213	.080
LateAvgIncNds	-81.6	27.7	3.3	-1.3	-1.3433	.032
LateStDevIncNds	1.3	5.9	7.4	4.0	1.3812	.137
LogOdds — Constant	192.3	29.4	3.4	10.7	-2.4637	.226
AverageStdDiff	50.0	13.1	9.9	6.3		
InfusionYrsEd	12.10	12.10	11.94	11.94		
ControlYrsEd	12.82	12.62	12.50	12.44		
Z-statistic		-1.14	-1.04	-1.04		
Inf HSGradRate	81.0	81.0	75.0	75.0		
Cont HSGradRate	89.2	81.0	81.3	87.5		

Though none of the differences in the years of education between the treated and control groups are statistically significant, the consistency between the three matched comparisons (differences of .52, .56, and .50 years of education) suggests that, with a larger number of pairwise comparisons, a statistically significant result might be found. The third matched comparison, with an average standardized difference of only 6.3%, is arguably the most reliable of the three.

Twelve of the individuals in the second experimental group have matches from both types of control groups (i.e. poor and non-poor). Therefore, one can compare twelve consistently poor individuals with twelve consistently non-poor individuals who are similar with respect to the important background variables. The non-poor control subgroup averages 12.25 years of education, whereas the poor control subgroup remains in school for an average of 11.92 years. This difference is not statistically significant, though it does suggest that poor children tend to achieve less academically than do similar individuals who are not poor.

For these matched comparisons, the requirement that one define a "treatment" and match individuals who are similar before that treatment takes place has been relaxed. Nevertheless, the outcomes suggest that a child who has been temporarily exposed to poverty early in his/her life tends to remain in school for a shorter period of time than one who is similar but is not exposed to poverty. A similar set of matched comparisons can be conducted for the first experimental group with the same control group. The results are summarized in the following table.

Variable	StdDiffBefore	StdDiffAfter PropScoreMatch	PrScoreMatch Eliminated 3	Mahalanobis Eliminated 3	Beta	p-value
DadEd	-66.2	24.9	26.1	12.5	-0.5447	.244
MomEd	-34.5	-5.0	-5.3	5.4	-0.2843	.484
OneParent	36.4	0.0	-7.0	0.0	0.5255	.397
NoParent	-17.9	0.0	0.0	0.0	——	—-
FirstBorn	-0.7	0.0	-12.5	0.0	0.3113	.496
Female	-7.1	0.0	-10.7	10.9	-0.2779	.454
NonWhite	34.6	0.0	0.0	0.0	-1.3930	.021
EarlySeps	11.0	18.8	10.6	-9.1	-1.1774	.588
MidEndSeps	7.3	0.0	-8.2	9.1	4.4389	.277
EarlyMomWorked	27.8	0.0	-6.5	-4.3	0.5404	.370
MidEndMomWorked	13.6	-9.9	-8.4	3.0	0.3395	.609
EarlyDisabled	26.0	28.3	31.6	14.4	3.0692	.005
MidEndDisabled	16.2	28.1	29.4	7.5	-3.0610	.016
EarlyWithOnePar	36.9	6.6	-2.4	2.5	1.6834	.098
MidEndWithOnePar	20.4	-9.3	-12.7	-5.8	-1.1613	.239
Weight	-62.0	-2.2	-3.5	-11.1	-0.0453	.035
AvgNumSiblings	34.4	-6.7	6.9	-15.4	0.3018	.023
YrsInSouth	70.2	7.6	8.0	-16.1	0.1154	.006
LateAvgIncNds	-60.8	24.8	12.5	9.1	-1.5486	.001
LateStDevIncNds	10.6	28.2	14.4	18.8	1.6238	.002
LogOdds — Constant	138.5	7.1	1.4	0.0	-0.5059	.688
AverageStdDiff	34.9	9.9	10.4	7.4		
InfusionYrsEd	12.00	12.00	11.89	11.89		
ControlYrsEd	12.82	12.25	12.24	12.30		
Z-statistic		-0.70	-1.06	-1.24		
Inf HSGradRate	72.5	72.5	70.3	70.3		
Cont HSGradRate	89.2	85.0	83.8	73.0		

These results, though not statistically significant, also suggest that children who are temporarily exposed to poverty incur an educational cost. The estimates of the "cost" of this one year of poverty are .25, .35, and .41 years of education. The residual bias on some of the background variables remains relatively high, but these matched comparisons allow a first-pass estimate of the cost of temporary exposure to poverty on a child's eventual educational attainment.

Twenty nine of the children in the first experimental group have matches from both types of control groups. The non-poor control subgroup averages 12.24 years of education, while the corresponding poor control subgroup has an average of 11.83 years of education. As before, this (statistically insignificant) result suggests that non-poor individuals tend to achieve more academically than do similar children who are poor.

Chapter 7

Conclusions

Many policymakers believe that providing the families of poor children with additional income is a desirable way to ameliorate childhood poverty and improve the opportunites open to poor children. In this document, I have investigated the extent to which childhood exposure to poverty affects an individual's eventual educational attainment. Using a filtered portion of the *Panel Study on Income Dynamics*, I estimate the potential educational benefits of markedly improving the economic circumstances of poor children and the educational costs which children who are exposed to poverty for the first years of their lives incur.

For the purposes of this study, I have employed multivariate matched sampling methods and linear regression techniques to isolate the effect of poverty on an individual's educational attainment. For the matched comparisons, individuals who rose out of poverty at an early age were compared with similar individuals who remained poor throughout their childhood. These comparisons were conducted to make a first-pass estimate of the potential benefit which lifting a poor child's household income above the poverty line could have on his/her educational attainment. Those children who rose out of poverty were also compared with similar individuals who were never exposed to poverty to determine any educational cost which they may have incurred. Finally, multiple linear regression models were estimated to determine the effect which household income during an individual's early years has on his/her eventual educational attainment.

The results of the first set of matched comparisons suggest that additional income *may* tend to improve the educational attainment of poor children. The estimated benefits are small and not statistically significant. There exists no income information about the PSID children for the first six years of their lives, and some data suggest that a substantial fraction of the children who seem to have risen out of poverty were only exposed for one or two years. If this is true, then our estimates of the educational benefits of improving a poor child's economic circumstances may be artificially inflated. In any case, the first set of matched comparisons does not convincingly demonstrate that giving consistently poor families additional income will improve the academic achievement of the children from these families. The PSID children (the members of the three experimental groups) whose families have, *on their own*, risen out of poverty may differ from poor children in ways which have not been considered. Hence, if the poor families were to receive a permanent income subsidy from the government, it is not entirely clear that their children would behave in the same way that the "experimental" children have.

The second set of matched comparisons, which compared individuals who were similar in all measurable respects except for the children's poverty status during their early years, were conducted to estimate the cost of being exposed to poverty early in life. The results of these matched comparisons suggest that those exposed to poverty early in life tend to achieve less academically than do similar children who are never exposed. The estimated cost is small, and the results are not statistically significant. For these matched comparisons, though, the existence of "false poor" individuals in the experimental groups would tend to downwards-bias the estimate of the cost of early childhood exposure to poverty. Thus, the children exposed to poverty early in life may well experience an educational cost which is greater than the best estimates suggest.

Finally, linear regression models were constructed to estimate the effect which household income and exposure to poverty during an individual's childhood has on his/her educational attainment. For the first set of econometric models, the *longitudinal* independent variables contain information about the first three years of interest (when children's ages were between six and eight). These models suggest that household income during these three years is a statistically significant determinant of an individual's educational attainment and that the number of years exposed to poverty is not. The more elaborate models, which also include information about the last seven years of interest, yield the surprising result that income in later years (i.e. 9-15) is a more important determinant of one's educational attainment than it is in the early (6-8) years. These results go against the belief, which is held by many educationists, that income in the early years of a child's life are the most critical to his/her eventual academic performance.

The results from the matched comparison methods are consistent with those obtained by Haveman and Wolfe in their analysis of the data. As the results from their econometric models suggested, if the government were to make a concerted effort to ameliorate poverty, the educational attainment of poor children would only slightly improve. Because these two methods are quite dissimilar, and yet both suggest only a modest benefit, the combined results of these two analyses are more credible than either would be on its own. One must bear in mind, though, that both analyses use the same data set, which is an imperfect set of information for the question we are trying to answer. The presence of several confounding factors is disturbing, and it is impossible to know to what extent other important information has been excluded. More recent empirical research has focused on the effects of changes in state tax laws on the behavior of impoverished families. Rigorous analyses of these data may provide more illuminating answers than the PSID data set can.

If the aggregate effect of the confounding factors is small, then the matched sampling methods used in this analysis suggest two things. First, it appears that children who are exposed to poverty for the first several years of their lives, and are then permanently lifted out of poverty, do incur an educational cost later on. First-pass estimates suggest that these children lose approximately a half of a year of education from this early exposure to poverty. Second, it is unlikely that improving the economic circumstance of children who have been exposed to poverty for the first several years of their lives will substantially improve their educational attainment. If the parents of a poor child do not use an income subsidy in ways which directly benefit the child, then the money may have no effect on the quality of the child's home environment or on his/her academic performance. Compensatory education programs or increased funding for the public schools in poor communities may prove to be much more effective ways to improve the educational prospects for poor children.

Bibliography

[1] P.S. Albin. Poverty, Education, and Unbalanced Economic Growth. *Quarterly Journal of Economics*, pages 70–84, June 1970.

[2] W.J. Baumol. Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. *American Economic Review*, pages 415–426, June 1967.

[3] G.S. Becker. *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education.* National Bureau of Economic Research, 1964.

[4] F. Black. The Trouble with Econometric Models. *Financial Analysts Journal*, pages 29–37, March-April 1982.

[5] M. Blaug. An Economic Interpretation of the Private Demand for Education. *Economica*, pages 166–182, May 1966.

[6] G.E.P. Box. Use and Abuse of Regression. *Technometrics*, 8(4):625–629, November 1966.

[7] R. Carter. The Sustaining Effects Study of Compensatory and Elementary Education. *Educational Researcher*, 13(7):4–13, 1984.

[8] W.G. Cochran and D.B. Rubin. Controlling Bias in Observational Studies: A Review. *Sankhya*, 35:415–446, 1973.

[9] J.S. Coleman. *Equality of Educational Opportunity.* U.S. Government Printing Office, 1966.

[10] D.R. Cox. *The Analysis of Binary Data.* Methuen, 1970.

[11] S. Danzinger and D. Weinberg. *Fighting Poverty: What Works and What Doesn't.* Harvard University Press, 1986.

[12] S. Danzinger and J.F. Witte. *State Policy Choices: The Wisconsin Experience.* The University of Wisconsin Press, 1988.

[13] S.H. Danzinger and K.E. Portney. *The Distributional Impact of Public Policies.* St. Martin's Press, 1988.

[14] G.J. Duncan. *Years of Poverty, Years of Plenty.* Institute for Social Research, The University of Michigan, 1984.

[15] Generation X-onomics. *The Economist*, 19th March 1994.

[16] R. Becketti et al. *Attrition from the PSID.* Unicon Research Corporation, 1983.

[17] R.A. Fisher. *Design of Experiments.* Oliver and Boyd, 1937.

[18] J. Hart. Poverty Takes Dramatic Toll on Education, Study Says. *The Boston Globe*, 6th March 1994.

[19] R. Haveman. *Starting Even.* The Twentieth Century Fund, 1988.

[20] M. Hill. *The Panel Study of Income Dynamics: A User's Guide.* Sage Publications, 1989.

[21] P.W. Holland. Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396):945–960, December 1986.

[22] C. Jencks. *Inequality.* Basic Books, Inc., 1972.

[23] T. Smeeding J.L. Palmer and B.B. Torrey. *The Vulnerable.* The Urban Institute Press, 1988.

[24] W. Knowlton and R. Zeckhauser. *American Society: Public and Private Responsibilities.* Ballinger Publishing Company, 1986.

[25] R.J. LaLonde. Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review*, pages 604–620, September 1986.

[26] I. Lazar and R. Darlington. *Lasting Effects After Preschool: A Report of the Consortium for Longitudinal Studies.* U.S. Department of Health, Education, and Welfare, 1978.

[27] A. Leibowitz. Home Investments in Children. *Journal of Political Economy*, pages S111 – S131, March-April 1974.

[28] R.A. Maynard. The Effects of the Rural Income Maintenance Experiment on the School Performance of Children. *American Economic Review*, pages 370–375, February 1977.

[29] R.A. Maynard and R.J. Murnane. The Effects of a Negative Income Tax on School Performance: Results of an Experiment. *The Journal of Human Resources*, pages 463–476, Fall 1979.

[30] F. Mosteller and J. Tukey. *Data Analysis and Regression.* Addison-Wesley Publishing Co., 1977.

[31] C.D. Petersen. Can JOBS Help the Underclass Break the Cycle of Poverty? *Journal of Economics Issues*, pages 243–254, March 1992.

[32] J.M. Quigley and D.L. Rubinfeld. *America's Domestic Priorities: An Economic Appraisal.* University of California Press, 1985.

[33] T.I. Ribich. *Education and Poverty.* The Brookings Institution, 1968.

[34] P.R. Rosenbaum and D.B. Rubin. Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome. *The Journal of the Royal Statistical Society*, 45(2):212–218, 1983.

[35] P.R. Rosenbaum and D.B. Rubin. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1):41–55, 1983.

[36] P.R. Rosenbaum and D.B. Rubin. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, 39(1):33–38, February 1985.

[37] D.B. Rubin. Matching to Remove Bias in Observational Studies. *Biometrics*, pages 159–183, March 1973.

[38] D.B. Rubin. The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies. *Biometrics*, 29(1):185–203, March 1973.

[39] D.B. Rubin. Estimating the Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5):688–701, 1974.

[40] D.B. Rubin. Multivariate Matching Methods That Are Equal Percent Bias Reducing, I: Some Examples. *Biometrics*, 32(1):109–120, march 1976.

[41] D.B. Rubin. Multivariate Matching Methods That Are Equal Percent Bias Reducing, II: Maximums on Bias Reduction for Fixed Sample Sizes. *Biometrics*, 32(1):121–132, march 1976.

[42] D.B. Rubin. Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*, 2(1):1–26, Spring 1977.

[43] D.B. Rubin. Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics*, 6(1):34–58, 1978.

[44] D.B. Rubin. Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association*, 74(366):318–328, June 1979.

[45] D.B. Rubin. Bias Reduction Using Mahalanobis-Metric Matching. *Biometrics*, 36(2):293–298, June 1980.

[46] D.B. Rubin. Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, 79(387):516–524, September 1984.

[47] I.V. Sawhill. Poverty in the U.S.: Why Is It So Persistent? *Journal of Economic Literature*, 26:1073–1119, September 1988.

[48] Steven Shulman. The Causes of Black Poverty: Evidence and Interpretation. *Journal of Economic Issues*, 24(4):995–1016, December 1990.

[49] H. Watts. Have Our Measures of Poverty Become Poorer? *Focus*, Summer 1986.

[50] R.D. Weiss. The Effect of Education on the Earnings of Blacks and Whites. *The Review of Economics and Statistics*, pages 150–159, May 1970.

[51] J.G. Williamson. *Inequality, Poverty, and History.* Basil Blackwell, Inc., 1991.