between class struggles and nonclass struggles). The essential point in the present context is that such processes can be represented in this kind of formal model, and such formal models can facilitate systematic empirical investigation.

After a long period of systematic exclusion from American social science, Marxism is reemerging as a serious theoretical force. In the 1960s this resurgence of Marxist perspectives largely took the form of radical critiques of the accepted wisdom in various academic disciplines. Increasingly in the course of the 1970s these critiques have been transformed into the elaboration of a positive alternative. New questions are being asked, new answers are being formulated, and new research agendas are being generated.

This book is an initial result of that process. It has attempted to take the theoretical challenge and insights of the emerging Marxist social science onto an empirical terrain familiar to non-Marxist social scientists. If nothing else, I hope that this effort has established the empirical significance of various issues raised within contemporary Marxist theory.

Regression equations, of course, cannot establish the general validity of a conceptual paradigm. Especially where disagreements over theoretical issues are reinforced by powerful ideological commitments, empirical results alone can never be decisive. Nevertheless, the kind of empirical investigation pursued in this study may help to demonstrate the scientific side of Marxism, and in so doing expand the possibilities for sustained theoretical debate.

APPENDIX A

Data Sources

Four different data sources will be used to investigate the various hypotheses discussed in chapter 4.

THE PANEL STUDY OF INCOME DYNAMICS (PSID)

Since 1968, the University of Michigan Institute for Social Research (ISR) has been conducting a panel study of households across the United States. Each year the heads of these households have been reinterviewed on a wide variety of topics concerned with income, standards of living, work, and similar issues. When individuals leave a family and begin a new household, or when families split apart, both old and new households are included in the study. As a result, the original sample of 5000 households has grown to nearly 6000 households, in spite of a certain amount of attrition due to deaths and nonresponses.

The original sample was stratified in order to oversample blacks and poor people. As a result, an elaborate system of weights has been developed to make the sample approximately a national random sample. These weights were readjusted after the fifth wave of the study (1972) in order to compensate for differential nonresponses during the first five years of the study. In every year since 1970 the response rate has been well over 95%, and thus no additional changes in weights were felt to be necessary (see IRS, 1972, pp. 9–46, for a further discussion of the sample and weights).[1]

Until the 1975 interview year, the study did not contain the necessary questions to allow for a class analysis of the data. In the 1975 questionnaire, several items which I had submitted to the project director were included (along with a number of others) which make at least a rough class typology possible. The data we will use, therefore, come primarily from this single year of the study.

Since most of the questions in the Panel Study are asked only of heads of households, whom the researchers in the project generally assumed to be the husband in married families, the PSID data can really only be used to investigate the relations of class and income among men. Nevertheless, because of the large sample size and the richness of the questions asked on the survey, we will use the Panel Study as the central data source for most of the hypotheses.

THE SURVEY OF WORKING CONDITIONS, 1969 (SWC)

A second source of data is the ISR Survey of Working Conditions. This survey consists of a national random sample of some 1500 adults

[1]A word needs to be said about the randomness of the sample in the Panel Study. Although the response rates have been very high in the third through eighth years of the study, they were not terribly high during the first two years (76% and 90%, respectively). By the eighth wave of the survey, then, when the questions on class position were added, the net response rate was only slightly above 50%. Even with the elaborate system of weightings used to compensate for differential nonresponses, the sample we are using cannot really be considered a true random sample of the American population.

If we were especially interested in studying frequencies, this nonrandomness would pose a serious problem. For example, 19.3% of the men in the weighted sample who are active in the labor force are professionals, compared to only 14% in the population as a whole; only 8.7% of the men in the sample are laborers or service workers, compared to some 18% of the total population. Given this departure from expected frequencies for the occupational distribution, we cannot have a great deal of confidence in the frequencies for class distribution in the Panel Study. Our main statistical technique, however, will be linear regression analysis, which is much less sensitive to nonrandomness. As long as nonresponses are more or less random with respect to the residuals in a regression equation, the estimated coefficients will remain unbiased estimates of the true population coefficients. We will assume that this is the case throughout this analysis.

active in the labor force. The major focus of the study was job satisfaction and working conditions, but enough questions were asked about employment status, position within authority hierarchies, income, and other variables relevant to our concerns to enable us to use the data to test a number of our hypotheses. Since the Survey includes women, these data will be especially relevant for hypotheses 10.1 and 10.2 about the interactions of class and sex.

THE QUALITY OF EMPLOYMENT SURVEY, 1973 (QES)

This survey is basically a replication of the 1969 Survey of Working Conditions. A number of questions on the earlier survey were dropped, however, and this made the 1973 replication somewhat less useful for our purposes. In particular, self-employed respondents were not asked whether or not they had employees in the 1973 survey, and thus it is difficult for us to distinguish the petty bourgeoisie from small employers. We will therefore rely more heavily on the 1969 Survey in the hypotheses concerning sex differences.

THE HIERARCHY IN ORGANIZATIONS STUDY

These data were gathered by Arnold Tannenbaum and associates (Tannenbaum et al., 1974) as part of a cross-national comparison of hierarchical structures within economic organizations in five countries: the United States, Austria, Italy, Yugoslavia, and Israel. In each country five large and five small plants were selected for study, and within each plant an assortment of people from different levels of the authority structure was interviewed. While the data do not represent random samples within each country, they have the advantage of containing extremely detailed information about each respondent's position within the authority hierarchy of the plant. We will, therefore, use these data to explore the various hypotheses which concern divisions within the managerial hierarchy.

APPENDIX B

Variables

The discussion of variables in this appendix will focus primarily on the definitions and operationalizations in the PSID data. The variables in the SWC and QES data will be mentioned only when they differ in significant ways from the PSID data. The variables used in the Hierarchy study are discussed in chapter 7.

OPERATIONALIZATION OF CLASS

None of the surveys available for this study enables us to operationalize precisely the class typology presented in Table 2.3. In particular, none of them contains any information on the extent to which individuals actually control their own labor process. Thus it is impossible to define the contradictory class position between the petty bourgeoisie and the proletariat (semiautonomous employees), and as a result throughout the present research such positions will be merged with the working class itself. Since most "semiautonomous employees" are probably located fairly close to the working class anyway, this should not seriously distort the results.

All of the surveys also lacked any information on the relations of real economic ownership (i.e., actual control over investments and accumulation) as discussed in chapter 2. In defining capitalists, therefore, we must rely on the formal legal criterion of being self-employed and employing others. This means that a top executive in a monopoly corporation would not be placed in the bourgeoisie. Since it is unlikely that any such individuals appear in these surveys, this is unlikely to affect the analysis significantly. Because of these limitations, the operationalization of class which will be used throughout this research should be viewed as a first approximation, clearly in need of further refinement.

Table B.1 presents the basic operationalization of class we will use with the PSID data. The questionnaire items used to construct this typology appear in Table B.2. Several comments on the typology are necessary.

First, approximately 2.8% of the respondents stated that they were both self-employed and worked for others. Of these ambiguous cases, nearly two thirds (1.7% of the sample) said that they had employees. Since it seemed reasonable to assume in such cases that the individual was probably mainly an employer, these individuals have been included in the employer category. The most important results of the study were run separately for employers who said that they were self-

TABLE B.1
Criteria for Class Position in the Panel Study of Income Dynamics

Class Category	Self-employed (D5)[a]	Employ Others (D12,D8)	Supervise Others (D14)	Say in Pay or Promotions (D16)	Employed (D5)
Employers	Yes	Yes	Yes[b]	Yes	No
Managers	No	No	Yes	Yes	Yes
Supervisors	No	No	Yes	No	Yes
Workers	No	No	No	No	Yes
Petty bourgeoisie	Yes	No	No	No	No

[a]These numbers represent the question number in the questionnaire administered; see Table B.2.

[b]Self-employed people did not actually answer questions D14 and D16, but these responses are implicit in their response to question D12.

TABLE B.2
Items in the PSID Questionnaire Used to Operationalize Class Position

SECTION D: EMPLOYMENT

D1. We would like to know about your (HEAD'S) present job—are you (HEAD) working now, looking for work, retired, a housewife, or what?

- 1. Working now, or *only* temporarily laid off
- 2. Looking for work, unemployed (Turn to page 12, E1)
- 3. Retired
- 3. Permanently disabled
- 4. Housewife
- 5. Student

(Turn to page 14, F1)

- 6. Other ________ (Specify) (Go to D2 if has job, otherwise turn to page 14, F1)

D2. What is your main occupation? (What sort of work do you do?)

________________________________ OCC ☐ IND ☐☐

(IF NOT CLEAR) → D3. Tell me a little more about what you do.

____

D4. What kind of business is that in? ________________________________

D5. Do you work for someone else, yourself, or what?

1. Someone else | 2. Both someone else and self | 3. Self only

[1. Someone else]

D6. Do you work for the Federal, State or local Government?

1. Yes | 5. No

(Turn to page 8, D14)

[2. Both someone else and self]

D7. Is your own business incorporated?

1. Yes | 5. No | 8. DK

D8. When you work for yourself, do you employ other people?

1. Yes | 5. No (Go to D10)

D9. How many?________

D10. When you work for others, do you work for the Federal, State or local Government?

1. Yes | 5. No

(Turn to page 8, D14)

[3. Self only]

D11. Is your business incorporated?

1. Yes | 5. No | 8. DK

D12. Do you employ other people?

1. Yes | 5. No (Turn to page 8, D18)

D13. How many?________

(Turn to page 8, D18)

(continued)

TABLE B.2 (*continued*)

D14. Do you supervise the work of others, or tell other employees what work to do?

[1. Yes] [5. No] (Go to D17)

D15. About how many people are you responsible for? ______

D16. Do you have any say about their pay or promotion?

[1. Yes] [5. No]

(Go to D17)

D17. Does your boss have a supervisor over him?

[1. Yes] [5. No]

D18. How long have you had this job? ______

(If *one year or more,* turn to page 9, D24)
(If *less than one year*)

D19. What month did you start this job? ______

D20. What happened to the job you had before—did the company fold, were you laid off, or what? ______

D21. Does your present job pay more than the one you had before?

[1. Yes, more] [5. No, same or less]

D22. On the whole, would you say your present job is better or worse than the one you had before?

[1. Better] [5. Worse] [3. Same] → (Turn to page 9, D24)

D23. Why is that? ______

(Turn to page 9, D24)

employed only, and in all instances the results were essentially identical to those of the enlarged employer category. In the case of respondents who stated that they were both self-employed and worked for others but did not employ anyone, it seemed unreasonable to assume that the individual was either mainly petty bourgeois or mainly a worker. These respondents, 1.1% of the sample, have thus been excluded from the analysis.

Second, the designation "employer" is being used rather than "capitalist" since the large majority of individuals in this category are

small businesspeople: 77% of the employers employed fewer than ten workers, and only 7% employed 50 workers or more. The distribution of number of employees of the employers in the sample is given in Table B.3.

Third, many of the people classified as "supervisors" probably belong in the working class proper. To be classified as a supervisor one need only say "yes" to the question: "Do you supervise the work of others or tell other employees what work to do?" The head of a work team who really has no capacity to invoke sanctions on other workers would probably be classified as a supervisor by this criterion. The "manager" category, on the other hand, is unambiguously in the contradictory position between the working class and the bourgeoisie since to be in this category one must have some say in the pay or promotions of subordinates. Throughout our analysis of the manager/supervisor class position, therefore, we will examine the results both for the combined category (managers and supervisors) and for managers taken separately.

The operationalization of class in the 1969 Survey of Working Conditions is essentially the same as in the Panel Study. Only two differences are worth mentioning. First, the question about having a say in pay and promotions of subordinates was not asked in the Survey, and thus it is impossible to make the distinction between supervisors and managers. Secondly, the supervision question itself was posed in a somewhat ambiguous fashion in the Survey: "Do you supervise people on your job?" In the Panel Study the question explicitly mentions supervision of other *employees* on the job. The result is that in the

TABLE B.3
Number of Employees and Expected Income

Number of Employees	Mean Annual Taxable Income	Employers in Each Category N	%
1–2	$18,811	90	36.3
3–5	29,752	72	29.0
6–9	21,598	28	11.3
10–19	26,247	24	9.7
20–49	36,694	16	6.5
50–99	43,106	10	4.0
100 or more	36,092	8	3.2
All employers	$25,898	248	100%

TABLE B.4
Characteristics of the Data Sets for Basic Class Operationalizations

Questions Available for Operationalizing Class	Panel Study of Income Dynamics	1969 Survey of Working Conditions	1973 Quality of Employment Survey
Do you supervise anyone on the job?	No	Yes	Yes
Do you supervise other employees?	Yes	No	No
Do you have say in pay and promotions of subordinates?	Yes	No	No
Are you self-employed?	Yes	Yes	Yes
Do you employ others?	Yes	Yes	No
Special assumptions used in operationalizing class	People who say they are both employers and work for others are probably mainly employers.	Teachers are generally nonsupervisors even if they say they supervise people on their jobs.	
Restrictions on the sample used in this study	Males only	Active participants in labor force only (work 20 hours a week or more)	
Effective N for this study	3,205	1,533	1,496

Survey well over half of all elementary and secondary school teachers reported that they "supervise" people on the job. In the Panel Study less than a quarter of all teachers said that they were supervisors. Presumably the difference is that the SWC teachers assumed that supervision of students counted as supervision. In the SWC data, therefore, teachers have all been classified as workers rather than managers/supervisors. While this undoubtedly results in a misclassification of some of the teachers, this is a less serious error than classifying most teachers as managers.

The 1973 QES lacked the question on having employees. When we use this survey we will assume that all self-employed people who say that they supervise people on the job are employers. In other respects the operationalization of class is the same as in the 1969 Survey.

The operationalization of class in the Hierarchy study is discussed in chapter 7.

Table B.4 summarizes the questions available for the operationalization of class in the Panel Study and two surveys, along with other characteristics of the samples.

INCOME

The Panel Study provides an extraordinarily rich source of data on income. Each year the heads of the households in the sample were asked detailed questions on wages, asset income, transfer income, business income, and much more. Unfortunately, in a number of instances the data on the computer tape were bracketed into ranges rather than presented in raw form. It was necessary in such cases to estimate the actual dollar amount of income by interpolating between categories and extrapolating in the open-ended category. The method for this estimation procedure is presented in Wright (1976b, pp. 162–64).

On the basis of these diverse income data, five income variables were constructed: total annual taxable income in 1974, "permanent" income, nonwage income in 1974, wage income in 1974 (earnings), and imputed hourly wage. In each case these variables refer to the head of household's income before taxes. Since the results using these alternative income measures were virtually identical in nearly all cases, the analysis reported in this study is limited to the first income variable. Formal definitions of the other measures and the regression results using them can be found in Wright (1976b).

Oddly enough, the PSID data tape did not actually contain a variable for the head's total annual income. It was necessary, therefore, to construct this variable from several others—total family taxable income (FI), head's labor income (HLI), wife's labor income (WLI), and wife's asset income (WAI). No composite head's asset income variable was available. "Labor income" represents wage and salary earnings plus a part of business income.[1] Of these, only WAI had to be estimated from a bracketed variable. Since in general the wife's asset income only amounted to a few thousand dollars at most, the error introduced by this estimation should not be terribly significant. In principle, the head's total annual income is equal to: FI − WLI − WAI. The difficulty is that the raw score variables are truncated at $99,999. As a result it is possible for the actual total annual income of the head to be much higher than $99,999 and yet for the expression FI − WLI − WAI to be much less than this. Thus, the following convention was adopted: if the head's labor income is greater than FI − WLI − WAI, then the head's total income is set equal to the head's labor income; otherwise it is set equal to FI − WLI − WAI. Since in fact very few individuals in the survey have incomes in excess of $99,999, little distortion is introduced by this procedure.[2]

Total annual taxable income thus includes the head's income from rents, interest, assets, wages, salaries, bonuses, businesses, and most other taxable sources. Taxable income does not include transfer payments such as social security, welfare, or scholarships, which are not taxed. The result is that the total income of the poor will be somewhat understated. In terms of our analysis this will probably tend to *increase* the returns to education among workers, and thus, if anything, would tend to reduce the hypothesized differences between classes.[3]

[1]The Panel Study apportioned business income into an imputed-earnings component and a return-on-investment component in the following way. The total number of hours worked in a year by the businessman was calculated; he/she was then arbitrarily "credited" with $1 for each of these hours. This figure was then subtracted from total business income. The remainder was then divided in two parts: half was added to the $1 × number-of-hours-worked to produce a total "labor part of business income," and the other half was considered the asset part of business income. Similar procedures were used to apportion rent, farm income, income from roomers, and other "mixed" incomes (see ISR, 1972, pp. 307–8 for a more detailed discussion). As far as I am aware, there is no real theoretical rationale for any of these procedures.

[2]The alternative to this procedure for estimating the head's total income would have been to add income from rent, interest, bonuses, businesses, etc., to the head's labor income. Since each of these variables except for HLI was bracketed, this would have involved considerably more error than the procedure adopted.

[3]The omission of untaxed transfer income from the income variable will probably increase the income returns to education among workers since it would be expected that such transfer income would be most significant among less educated sections of the working class.

The total taxable income variable also does not systematically include realized capital gains. A question was included in the study on windfall income ("Did you get any other money in 1974—like a big settlement from an insurance company or an inheritance?") and people were asked more generally if they had income from any unspecified source in 1974. Although some realized that capital gains may have been included in the responses to these questions, it is unlikely that such income is generally included in the taxable income variable. Unrealized capital gains are totally excluded. As a result, the income of employers and perhaps managers will be somewhat understated by this variable.

Income in the SWC and QES Data

Neither of these data sets contained as elaborate data on income as the Panel Study. The income variable is simply measured as total personal income of the respondent in raw dollars.

A Note on the Form of the Income Variable

Throughout this study, the income variables will be used in raw, unlogged dollars. The coefficients in equations using unlogged dollars are much more easily interpretable than in equations using logarithmic transformations of income. In the unlogged equation, the coefficient for the education variables tells you how much additional income would be expected for an increase of one unit of education. In the logged equation, the coefficient tells you approximately what the percentage increase would be. Unless the theory on which the equations are based is couched in terms of proportional increases (i.e., rates of return rather than absolute returns), the unlogged equations are more immediately understandable. Since the hypotheses developed in chapter 4 are all based on an analysis of absolute returns to education, the unlogged form of the equation is more appropriate for the task at hand.[4]

[4]Within a human capital framework, where education is viewed as an "investment" in human capital quite analogous to investment in physical capital, it makes some sense to estimate a parameter which can be viewed as the human capital analogue to interest rates (i.e., rates of return to education). In our analysis, however, the education coefficients using unlogged income are much more immediately interpretable in terms of the hypotheses we are testing, and thus raw income will be used.

It should be noted that in a number of cases, the use of natural logarithms does make a difference in our results. For example, in the comparisons of blacks and whites, when unlogged income is used, white males as a whole have significantly larger education coefficients than black males. When logged income is used, white males have signifi-

If the relationship between education and income were in fact highly nonlinear, then this use of unlogged income could indeed generate some serious distortions in the results. The linear approximation of a nonlinear relationship could seriously underestimate the returns to education for more highly educated individuals, and since there is a relationship between education and class, this could distort some of our class comparisons. As is indicated in chapter 6, however, the relationship between the education variable and unlogged income is quite linear, and thus this should not be a serious problem.

EDUCATION

Education is measured by a quasi-credential scale rather than by years of education, in the following manner:

0 = no schooling or illiterate
1 = some elementary school
2 = completed elementary school
3 = some high school
4 = high school
5 = high school plus nonacademic training
6 = some college
7 = college degree
8 = graduate training

Each step in this scale represents a socially recognized level of education. In practice, it is highly unlikely that any of the results of this study would have been different if years of education had been used instead, but sociologically a credential is a more meaningful unit of education than a year.[5]

cantly smaller coefficients than black males. Given that black males have lower absolute incomes than white males at every level of education, this difference in results is hardly surprising. Since the education coefficient in an equation predicting logged income roughly indicates the proportional increase in income for a unit increase in education, a smaller absolute increase among blacks in the raw income equation can easily produce a greater *rate* of increase in the logged income equation. The choice of which form of the income variable to use should therefore be based on the substantive theoretical problems being investigated. Again, since the hypotheses in chapter 3 are structured around raw income, we will use this form of the variable throughout our analysis.

[5]There is no general consensus in research on income inequality whether years or quasi-credentials of education constitute a more appropriate variable. In fact, the question is almost never even raised. From a theoretical point of view, if one adopts a human capital perspective on the relationship of education to income, in which income is seen

Education will be used in regression equations in two forms, first as the single, nine-level scale presented above, and second, as a series of six dummy variables. When the dummy variables are used, the left-out category will be levels 0–2 in the above scale, i.e., elementary school or less education.

The QES and SWC data use essentially the same scale, except that no distinction is made between high school and high school with additional nonacademic training.

STATUS

Occupational status is measured by the standard Duncan status scale adjusted for the 1970 census occupational categories (see Featherman et al., 1975, appendix B). In the Panel Study, the three-digit occupation codes were recorded only for the 1974 interview year. Since this is the year during which the income reported in the 1975 interview was earned, there is no particular disadvantage in not having the three-digit codes for the 1975 interview year. In effect, we have the occupational status for each individual at the beginning of the year in which we have income data, and the class position at the end of the year (i.e., beginning of the following year).

Approximately 8% of the sample were split-offs in the 1975 interview year (i.e., heads of new households), and thus three-digit occupation codes from 1974 were unavailable. In these cases, only one-digit occupational codes were available for the 1975 interview year. For these individuals the mean status scores for the one-digit occupational code were used. In the SWC data, only decile status scores were available.

JOB SENIORITY

Job seniority or job tenure is a simple measure of the number of years worked on the current job. For employed people, this is the number of years working for current employer; for self-employed people, it is the number of years in the present business.

as a return to real increments of skill level, expertise, etc., then years of education is probably the better variable. If, on the other hand, one views education primarily as a certification process used in job recruitment screening, then the quasi-credential version of the variable is more appropriate. At any rate, both versions have been extensively used in the literature.

This variable has a very different theoretical meaning in the perspective underlying our analysis than in standard economic treatments of job tenure. Specifically, in human capital theory, seniority would be viewed as a proxy for the experience gained within a particular job setting, and thus as a measure of productivity. In our analysis, on the other hand, job tenure is more central to the social control logic of income, since seniority pay increases constitute one of the basic devices for structuring inducements into jobs.

AGE

The variable "age" has a double meaning in income determination models. On the one hand, age is a fairly good measure of labor market experience, especially for the analysis of income among men (in the PSID data, the correlation between age and actual number of years worked is .94). Within the general human capital framework such experience would be considered another ingredient in productivity. But age is also a proxy for history, for the specific historical conditions that have shaped the economic lives of successive cohorts. When sociologists control for age in studies of income inequality and status attainment, the rationale is usually a concern for cohort effects rather than a desire to control for labor market experience as such.

In the present research, it does not particularly matter which of these interpretations of age is used. We are less interested in studying age effects in their own right than in showing that the class effects are not simply artifacts of age distributions within classes.

FAMILY BACKGROUND

The Panel Study contained three family background questions which we will use: father's education, father's occupation, and "average" parental economic situation. The first of these is measured on the same scale as respondent's education. Father's occupation appears only as a one-digit occupation code. This code was converted to a status score by attributing the mean occupational status to each occupational category. This is clearly a mediocre measure of occupational status, but it is the best available in the present data. Finally, average parental economic situation during the respondent's childhood consists of a three-point scale:

1 = parents were generally poor
2 = parents were generally about average
3 = parents were generally well off

As in the case of age and status, we will not be particularly concerned with class differences in these background variables, but will mainly use them as controls.

Neither the QES nor the SWC data contained any background information at all.

ANNUAL HOURS WORKED

When total annual income is a dependent variable one obvious source of variation is the total number of hours worked during the year. This variable was calculated by asking respondents how many weeks they worked in the previous year and how many hours, on the average, they worked per week. Although this variable undoubtedly is subject to considerable measurement error, it will enable us to see whether or not any of the class differences are simply consequences of the amount of time spent earning the income.

NUMBER OF EMPLOYEES

Small employers were asked how many people they employed. This variable was scaled:

0 = none
1 = 1–2 employees
2 = 3–5 employees
3 = 6–9 employees
4 = 10–19 employees
5 = 20–49 employees
6 = 50–99 employees
7 = 100–499 employees
8 = 500 or more

No subject in the study had 500 or more employees, so the practical maximum of the scale is level 7.

APPENDIX C

Statistical Procedures

STATISTICAL TESTS

Most of the hypotheses presented in chapter 3 constitute hypotheses about interactions between class and various other variables. The standard statistical approach for dealing with such problems is the analysis of covariance. I compared the various class categories and the sex and race categories within classes in terms of a series of regression equations to see: (a) if they differed significantly in the slopes of the independent variables in the equations, especially education; and (b) if a significant "gap" in income existed between the compared groups when the independent variables were held constant. Figure C.1 illustrates these two kinds of comparisons for the simple regression of income on education for managers and workers.

The test for the significance of slope differences simply involves testing whether the slope of the manager minus worker education coefficient is significantly different from zero. The conventional procedure for such a test is to pool the groups being compared into a single sample, construct a dummy variable that is equal to 1 for one of the groups and zero for the other, and then estimate a regression equation which contains as independent variables all of the original independent variables, all of the original independent variables multiplied by the dummy variable, and the dummy variable itself. It is quite simple to show that in such an equation the coefficients of the dummy variable interaction terms constitute the difference between the coefficients of the independent variable for the two groups (see Kmenta, 1971, pp. 419–23). Thus a simple t-test on these dummy variable interaction terms constitutes a test of the slope differences between the groups in question.

This is a rather cumbersome procedure, since it involves constructing large correlation matrices consisting of all the independent variables and dummy interactions. Furthermore, separate matrices must be calculated for every comparison being made. A much simpler procedure is simply to treat the coefficients of the independent variables in the regressions for the two groups being compared as normally distributed variables with expected values equal to the coefficient, and then to perform a direct t-test of their difference. Suppose we estimate the simple regression of income on education for two groups:

Group 1: Income $= a_1 + b_1$ Education;
Group 2: Income $= a_2 + b_2$ Education.

The direct t-test is then

$$t = \frac{b_1 - b_2}{(\mathrm{se}^2_{b_1} + \mathrm{se}^2_{b_2})^{1/2}}$$

where se_{b1} and se_{b2} are the standard errors of the coefficients b_1 and b_2, respectively. The conventional dummy-variable interaction strategy uses the following t-test:

$$t = \frac{b_1 - b_2}{[(v_1 \mathrm{se}^2_{b_1} + v_2 \mathrm{se}^2_{b_2})/(v_1 + v_2)]^{1/2}}$$

where v_1 and v_2 are the degrees of freedom for groups 1 and 2, respectively. This t-value is always greater than or equal to the simpler way of calculating t, and thus if anything the simpler t-test is a more conservative procedure.[1] We will use the simpler t-test throughout the analysis of the PSID data.

[1]A discussion of the relationship between the t-test we will use and the t-test proposed by Mallinvaud (1966, pp. 304–6), can be found in Wright (1976b, pp. 165–67).

The analysis of the QES and SWC data was completed before I was aware of the advantages of this simpler t-test. Since in general the two tests produce the same results, and the dummy variable interaction approach is the usual method anyway, it did not seem necessary to redo the entire analysis using the direct t-test.

The analysis of "gaps" in income, controlling for values on the independent variables, poses rather different problems from the analysis of slope differences. The difficulty is that the magnitude of this gap is strictly dependent upon the level of the independent variables at which it is evaluated. Thus, in Figure C.1, if managers and workers were compared at zero education (the usual constant term in regression equations), the gap would be slightly negative; if they were compared at the workers' mean education ($\bar{E}_w$), the gap would be positive but relatively small; if they were compared at the managers' mean education ($\bar{E}_m$), the gap would be positive and large.

Probably the most common convention in such exercises is to evaluate the gap at the level of the independent variables equal to the means for the hypothesized privileged group. Thus, when Duncan (1969) examines how much of the total difference in mean income between blacks and whites can be attributed to differences in the values of the independent variables for the two races, he substitutes the white mean values into the black equation and examines what the expected income would be. This procedure serves the polemical point of showing what the incomes of a disadvantaged group would be if they were not disadvantaged in their personal characteristics (education, age, experience, IQ, etc.) but were still disadvantaged in their ability to transform those personal characteristics into outcomes (money, occupations, etc.).

We are less interested in revealing what a worker's income would be if he or she had all the personal characteristics of a manager than in examining the differences between the working class and the managerial category as structural positions within class relations. This means that the income gap should be evaluated at an average level of the independent variables rather than at the means of the privileged category. This will be accomplished by assessing the gap in incomes at the average of the means of the two groups being compared. The income gap at this point represents the expected difference in income between a worker and a manager with identical education equal to the average of the mean worker and mean manager education. We will refer to the comparison of expected incomes at this point as the analysis of the "average gap" in income between the groups being compared.

If the constant term in the regression equations is shifted to the point at which we want to assess the income gap, then the statistical

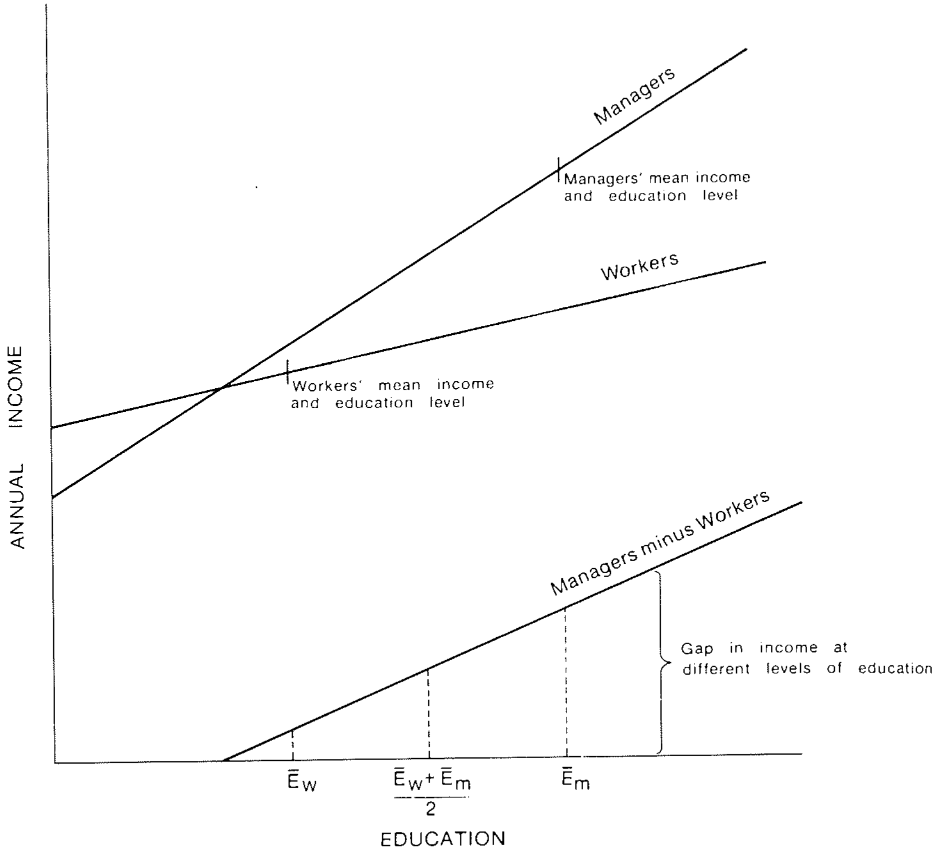

Figure C.1. Illustration of method of structural comparisons. $\bar{E}w$ = mean worker education; $\bar{E}$ m = mean manager education; $(\bar{E}_w + \bar{E}_m)/2$ = level of education at which the gap in income will be evaluated.

test for the significance of the income gap is simply a t-test of the difference in constant terms for the groups being compared. This t-test can be calculated in the same manner as in the slope comparisons.[2]

Since the average income gap is assessed at a different level of the

[2]In computer packages like SPSS where the correlation matrix and the accompanying means and standard deviations of the variables can easily be output on cards, it becomes very simple to perform a large number of different comparisons among a series of groups. All that is necessary is to change the means on the independent variables from their true values to an adjusted value corresponding to the point at which one wants to assess the income gap. When this is done, the constant term in regression equations calculated from the correlation matrix using these adjusted means becomes the adjusted constant discussed above.

independent variables in each comparison, it is not possible to directly compare average income gaps across different comparisons. This is particularly a problem when different groups have very different slopes (e.g., workers and managers), since in this case the income gap increases as you evaluate it at higher levels of the independent variables. In order to compare gaps across comparisons it is thus necessary to assess the gap at the same point on all comparisons. We will, therefore, also calculate a "standardized income gap" by assessing the gap at the level of the independent variables of employers, the most "privileged" category in our analysis.

In the original analysis of the QES and SWC data in a paper by Wright and Perrone (1975), the income gaps were assessed at the overall sample means for each of the independent variables rather than at the average of the means for the groups being compared. Since, in practice, it makes little difference which of these two conventions is adopted, the income gaps for the SWC data will not be recalculated using the procedure described above.

Since all of the hypotheses spelled out in chapter 4 posit a specific direction for the differences being tested, one-tailed t-tests will be used throughout the analysis. In general the differences being examined are large, and few if any of the conclusions would have been different if a more conservative two-tailed test had been adopted.

A NOTE ON WEIGHTING AND MISSING DATA

All of the regression equations in the Panel Study will be estimated using the weighted sample. In order to avoid inflating the sample size in a given regression (and thus decreasing the standard errors of the regression coefficients), the weights for the subjects included in a particular regression will be divided by the mean weight for those subjects. The N for a regression equation will thus be very close to the true number of cases included in the regression.

Missing data in the regressions were handled through pairwise rather than casewise deletion (i.e., the correlation coefficients on which the regressions were based were each calculated for all cases for which data were available on both variables in the correlation). The usual procedure when using pairwise deletion is to take the lowest N in any of the correlations as the N for the regression equation as a whole. However, since we are substantively interested mainly in the coefficient for the education variable and in the adjusted constant term, the N in the correlation of education and income will be used.

POSSIBLE BIASES DUE TO RESTRICTIONS OF THE SAMPLES

All of the data used in this study are restricted to active, working participants in the labor force. This undoubtedly introduces certain biases. Most obviously, since most unemployed people belong in the working class, our estimates of the class distribution will somewhat understate the size of the working class.

This restriction of the sample may also tend to bias some of the results for returns to education. Under most circumstances, the rate of unemployment is especially high among less educated people. The exclusion of the unemployed from the analysis will thus tend to bias the returns to education *downward* (if the unemployed had been included, the expected income of poorly educated individuals would be less, and thus the returns to education greater). Conceivably, this might affect some of our class comparisons, since this bias is undoubtedly greater in the working class than in other classes. On the other hand, as already indicated in the discussion of the income variables, the exclusion of transfer income probably tends to increase the returns to education among workers. While there is no reason to assume that these two biases in fact cancel each other out, it does seem unlikely that the net bias in returns to education is terribly large because of the restrictions on the sample.

PRESENTATION OF RESULTS

Rather than encumber the text with endless tables and figures, I have presented only data that are directly discussed in the analysis, and have, therefore, excluded the tables of correlation matrices upon which the regression equations are based. These matrices, along with complete results for all the different income-dependent variables and complete results for the SWC and QES data, can be found in Wright (1976b, pp. 289–375).

References

Alexander, K. L., Eckland, B. K., and Griffin, L. J. 1975. The Wisconsin model of socioeconomic achievement: A replication. *American Journal of Sociology*, 81(2), 324–342.

Althusser, L. 1970a. *For Marx*. Trans. by B. Brewster. New York: Vintage Press.

———. 1970b. From "Capital" to Marx's Philosophy. In Althusser and Balibar, 1970. Pp. 11–70.

———, and Balibar, E. 1970. *Reading Capital*. Trans. by B. Brewster. London: New Left Books.

Balibar, E. 1970. The basic concepts of historical materialism. In Althusser and Balibar, 1970. Pp. 199–308.

Barber, B. 1957. *Social stratification*. New York: Harcourt, Brace and World.

Baudelot, C., Establet, R., and Malemort, J. 1974. *La petite bourgeoisie en France*. Paris: Maspero.

Becker, G. S. 1971. *The economics of discrimination*. 2nd ed. Chicago: University of Chicago Press.