IRP Discussion Papers
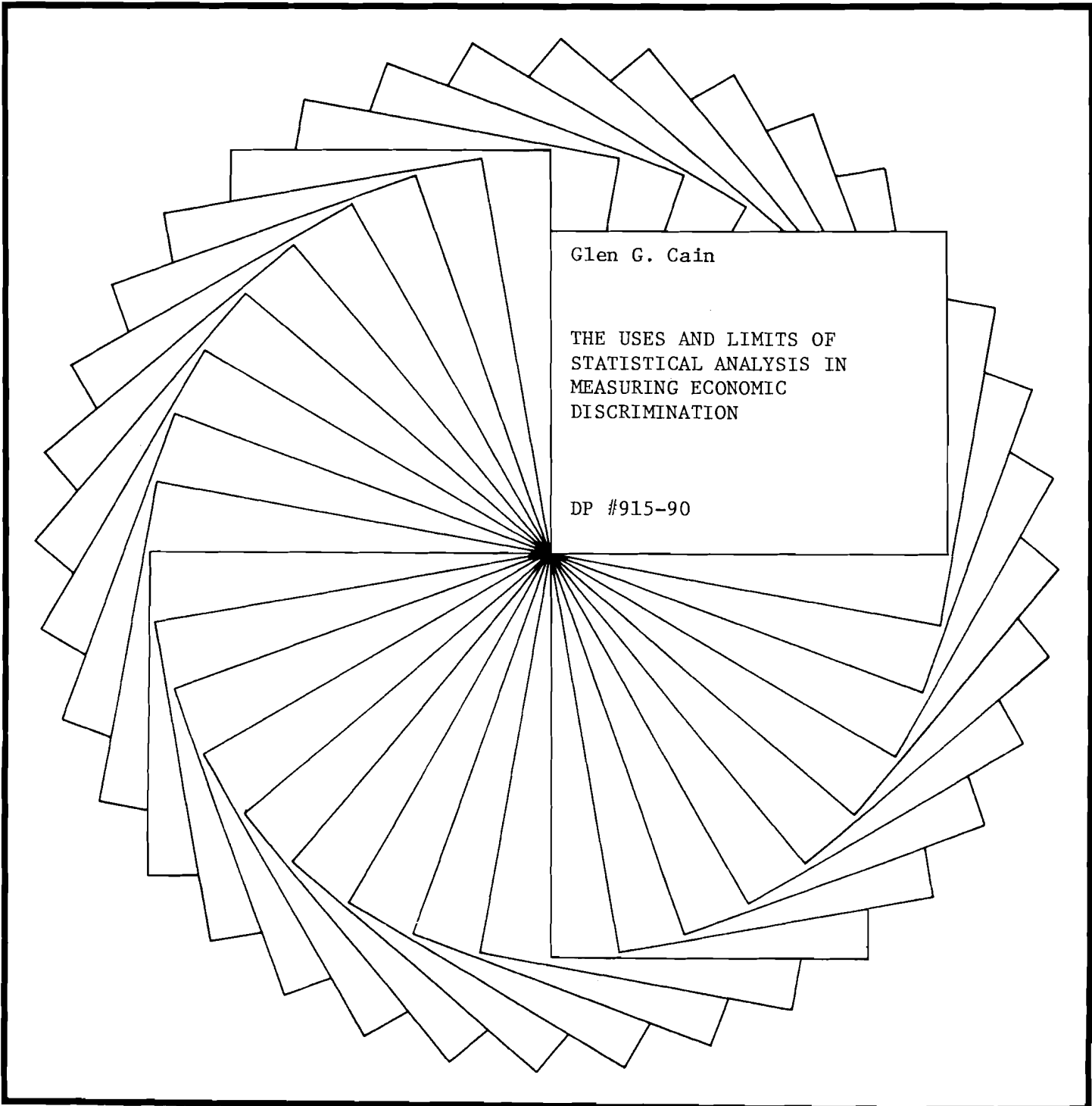
Glen G. Cain

THE USES AND LIMITS OF
STATISTICAL ANALYSIS IN
MEASURING ECONOMIC
DISCRIMINATION

DP #915-90

The Uses and Limits of Statistical Analysis in Measuring Economic Discrimination

Glen G. Cain
Department of Economics

and

Institute for Research on Poverty
University of Wisconsin-Madison

July 1990

Abstract

To measure economic discrimination, one must be able to measure the
productive capacity of people, the opportunity structure they face, and
outcomes in the form of income and earnings. The first two are
difficult to measure but crucial to any assertion that discrimination
exists. To illustrate the uses and limits of statistical analysis to
overcome these difficulties, two case studies of ethnic discrimination
in American history--Irish-Americans in 1900 and Japanese-Americans in
1940--are examined.

A persistent problem in using statistical models of discrimination
is distinguishing between an opportunity structure that is imposed on a
worker from one that reflects the choices or preferences of that worker.
For example, although a low wage relative to others with the same
productive capacity may be the result of discrimination, it may also
reflect the choice by the worker to work at a part-time job at more
convenient hours. Such a choice may explain why some married women fail
to hold jobs that fully use their productive capacity. Measuring
discrimination against women--who tend to have dual careers--is
especially difficult. Their wages may be low because they have spent
less time in the paid labor market, but it is also plausible that the
reason they have devoted more time to housework and less to market work
is that they have faced restricted opportunities in the marketplace.
The causation can run either way.

Statistical models of economic discrimination measure associations
but do not establish causation. Therefore, evidence based on
statistical analysis is necessary but not sufficient for establishing

the presence of economic discrimination. It must be supported by

historical and institutional evidence and interpreted within a

theoretical framework.

The Uses and Limits of Statistical Analysis
in Measuring Economic Discrimination

This paper begins with a brief discussion of the basic concepts of

economic discrimination. In the second part, two case studies of ethnic

discrimination from American history illustrate the uses and limits of

statistical analysis. (The detailed evidence of discrimination against

these groups is presented in Appendix A.) The next part deals with

current analyses of economic discrimination, and a summary concludes

that evidence based on statistical analysis is necessary but not

sufficient for establishing the presence of economic discrimination. It

must be supported by historical and institutional evidence and

interpreted within a theoretical framework.

I. BACKGROUND IDEAS

Economic discrimination is rooted in the practical matter of an

inequity; usually an outcome in which equally productive people receive,

on average, different rewards for their efforts. It is useful to

distinguish economic discrimination, when referring to the economy as a

whole, from labor market discrimination. Discrimination in the labor

market may be represented by the example of equally productive groups

receiving unequal earnings. In fact, the inequality in earnings is

usually a result of unequal access to the better-paying jobs and will

show up as inequality in occupational attainment.

The focus on monetary rewards is indisputably narrow and leads to

certain anomalies. For example, a finding of equal pay leads to a

verdict of no labor market discrimination, but the finding may, and probably does, coexist with plenty of labor market discrimination in the ordinary use of the term. This seeming paradox can be resolved with a simple example. Assume that 10 percent of the workers are a minority group, that no laws bar discrimination, and that 50 percent of the employers discriminate against the minority by paying them less than majority members. All that is needed for the economic verdict of "no discrimination" is a sufficient number of employers who will hire the minority workers on a nondiscriminatory basis, and here the proportion of nondiscriminating employers is relatively large. Examples of legal discrimination could be exposed if minority workers applied for jobs from discriminating employers, but the minority applicants would quickly shift to the employers who do not discriminate and end up receiving equal pay.[1]

The economist's conventional criterion of equal pay even has the embarrassing property of constituting evidence for no discrimination under conditions of complete segregation. The Supreme Court has rejected the doctrine of "separate but equal," but the customary analysis in economics accepts the doctrine. Despite this awkward property, I will focus on income and pay differences between groups as the indicator of economic discrimination. In defense, all I will say is that the money measure is important in its own right, and it does tend to be associated with other measures of discrimination such as segregation by residence or schooling or occupation.

Problems in measuring economic discrimination are revealed by considering its three essential components: (1) the productive capacity

of people, which is difficult to measure; (2) an opportunity structure facing the people, which is even more difficult to measure; and (3) the outcomes in the form of income and earnings, which are relatively well measured. The economist's method is to observe the outcomes, compare these for people who are assumed to be of the same productive capacity, and then infer whether the opportunity structure reveals discrimination.

Some examples will illustrate when the worker's opportunity structure implies or does not imply discrimination. Recent immigrants who are unable to speak English and who are less productive for that reason will face restricted employment opportunities. Language facility may well be a justifiable requirement for hiring, but what about the requirement that the worker be a citizen? For long periods in our history, occupational licensing in some states required citizenship for dentists, engineers, master plumbers, and barbers.[2] A barrier based on citizenship is arguably not only discrimination but tends to injure some ethnic groups more than others. In this century Asian immigrants to the United States were for a generation barred from citizenship in some states, notably California.

For another example, assume that workers in a particular ethnic group are untrained and have lower earnings. Their lower earnings suggest, but only suggest, that they are being discriminated against, but just where the barrier to training is located is often difficult to determine. Does the barrier rest with employers, with trade unions, with the educational system, with the workers' upbringing as children? Thus, these workers appear to face some type of discrimination in their economic opportunities, but it may not be labor market discrimination.

The above discussion involves wages and individual workers. The worker is the unit of analysis. I now turn to economic discrimination involving inequality in family income. The family is the unit of analysis. Family income is a fundamental measure of economic well-being and is more comprehensive than wage earnings. Moreover, examining income and the family may help us understand the opportunities and productive capacities of the person. Again, let me illustrate these points with examples.

A mother of young children may be working at a part-time job at a low wage, but her husband's earnings may be high enough to offer her the option of working less in the market than someone who must rely only on her own earnings and who works full time at a higher wage. The mother's low wage is not compelling evidence for labor market discrimination, because she may have high wage opportunities that match her productive capacity, but she does not choose these job opportunities because they require a full-time commitment. Another possible reason for the mother's lower wage is that her housework tasks may reduce her productive capacity in the paid labor market, because she does not have as much time or energy for market work as someone with less housework.

In this example the mother's family context is assumed to be causal to her job and wage outcomes. Now consider an example when causation runs in the other direction. Assume that discrimination in the labor market imposes low earnings on a young person or on a mother with young children, and these persons have to move in or stay with a larger household of other family members to take advantage of the pooling of resources and the economies of scale in living arrangements.

Comparisons of household or family incomes are likely to be more equal
than individual incomes, and the famiily comparisons can hide situations
of economic discrimination.

To illustrate, young black men are less likely than young white men
to live apart from their immediate family.[3] One reason they continue
to live at home is that fewer young black men have good paying jobs.
Therefore, they are less able to afford either to live alone or to marry
and set up their own households. Table 1 shows, with a hypothetical
example, how these living arrangements can understate the true gap
between black and white family incomes.

The original family unit is assumed to consist of three persons--
the parents and a 22-year-old daughter or son. The black-to-white ratio
of the family income of these original units is .75, derived from an
average family income of $36,000 for blacks and $48,000 for whites. We
do not observe these original family units, however; instead we observe
living arrangements in which more 22-year-old whites live apart from
their parents than do the 22-year-old blacks. To sharpen the
comparison, assume that 50 percent of the young whites and 25 percent of
the young blacks live apart from their parents. The observed family
incomes show .90 as the black-to-white ratio of family incomes--$28,800
for black families and $32,000 for whites.

The simple point of this table is that the gap between black and
white income is understated by the comparison of family incomes. A
second and more profound point is that the family unit itself, which is
a basic unit for income comparisons, is affected by the labor market
outcomes and, therefore, affected by discrimination. Discrimination is

6

Table 1

Family Income Comparisons When Living Arrangements Depend on Labor Market Opportunities (Hypothetical Example)

Original Families	White	Black	B/W Ratio
Number	100	100	
Income			
Family	$48,000	$36,000	.75
(Parents)	($30,000)	($24,000)	(.80)
(Daughter or son, age 22)	($18,000)	($12,000)	(.67)

	White			Black			B/W Ratio
		Separate			Separate		
New Living Arrangements	Together	Parents	Child	Together	Parents	Child	
Number	50	50	50	75	25	25	
Family (household) income	$48,000	$30,000	$18,000	$36,000	$24,000	$12,000	
Average family (household) income	$32,000			$28,800			.90

a cause of the living arrangements of the families. This contrasts with the example of mothers of young children, where the family unit had a causal effect on the labor market outcomes. Causation can run both ways, and one's measure of economic discrimination can be either over- or understated if this is not taken into account.

Two practical conclusions follow from the above examples. (1) The wage of the married woman with young children overstated labor market discrimination for her. Recall, also, that her household income showed no disadvantaged outcome. (2) Labor market discrimination against black youth was correctly measured by their lower wages, but economic discrimination was understated by the comparison of black and white household incomes. These conclusions depend on the particular circumstances of my examples. Later I discuss how the opposite conclusions might emerge from further analyses of these cases.

II. MEASURING DISCRIMINATION BETWEEN GROUPS IN A HISTORICAL SETTING

Assume that we adopt the following statistical procedure to measure economic discrimination, if any, against a given ethnic group, using all other native-born whites as the comparison group. First, we classify the families according to the age and rural/urban residence of the primary earner; then we compare the average family income and average per capita income of the two ethnic groups across these classifications. Disparities in income are initial evidence of economic discrimination. To keep matters relatively simple, let us assume that the number of earners per family and family size are approximately the same in the two groups and that those in the ethnic group are all born in the United

States. These steps are illustrated with hypothetical numbers in panel
A of Table 2.

Another step in the investigation is to subdivide the family groups
according to the educational attainment of the primary earner, and then
compare incomes within each educational category. If the same income
disparities persist, we may view this as stronger evidence of economic
discrimination in general and of labor market discrimination in
particular. If, however, this statistical control for educational
attainment, along with the controls for age and rural/urban residence,
do not show income differences, then we have two conventional
interpretations. (Note that panel B of Table 2 shows almost no
difference in the family incomes of the ethnic group relative to the
comparison group of native whites.) One interpretation is that the
ethnic group's lower educational attainment is the source for their
lower incomes, and their lower education reflects social disadvantages
that accrued to the ethnic group before they reached adulthood. This
view can be interpreted as exonerating the labor market, and employers
in particular, from blame for the discrimination.

A second and contrary interpretation is that labor market
discrimination is a cause of less education among members of the ethnic
group because the market does not sufficiently reward them for
additional schooling. Which of these two interpretations is correct is
a challenging question. In the case studies that follow, data
limitations prevent me from including education in my tabulations.

I would like to carry out the procedure outlined above for two
ethnic groups, Irish-Americans and Japanese-Americans, for two years,

Table 2

Family Income Comparisons by Ethnic Group
(Hypothetical Example)

Age	Ethnic Group X		Whites	
	Urban	Rural	Urban	Rural
Panel A				
.				
.				
.				
40-49	$20,000	$18,000	$25,000	$20,000
.				
.				
.				
Panel B				
Age by education				
.				
.				
.				
40-49				
< High school	...		...	
High school	$22,000		$23,000	
Some college	...		...	
College grad	$32,000		$32,000	
.				
.				
.				

1900 for the Irish and 1940 for the Japanese, using the decennial census. If the historians I have read are correct in their descriptions of the economic discrimination that these groups suffered in American history, then the economic statistics from these particular periods should show this.

The 1900 decennial census offers no data on income and earnings, and the 1940 census does not give earnings data for Japanese-Americans. These two censuses do, however, provide limited information about occupational attainment. I use these occupational data to support the proposition that a comparison of incomes in 1900 and in 1940 would show economic discrimination against, respectively, Irish-Americans and Japanese-Americans.

In 1900 the number of descendants of Irish immigrants was large enough to permit reliable statistics on the occupational attainments of those who were born in the United States. Also, within-group marriage was so prevalent that defining who was Irish is relatively unambiguous. Finally, as noted above in referring to the research of historians, it is likely that the Irish-Americans faced discrimination in the labor market during the period around 1900.

The occupational attainment of second generation Irish-American men in 1900 is compared to the attainments of all white men and of second-generation white men.[4] The group of all whites is composed of three ancestry subgroups: (1) 23 percent who were foreign born; (2) 20 percent who were second-generation Americans, that is, sons of foreign-born parents; and (3) 57 percent who were third- or higher-generation Americans, called native whites.[5] As a comparison group for evaluating

the economic attainment of the second-generation Irish, all white men probably represent a relatively low standard of economic attainment. Consider that the 23 percent who are foreign born include many who could not speak English and some who were illiterate. Even the Irish-Americans who were born in Ireland were all English-speaking, and virtually all were literate.[6]

My source for the occupational statistics from the 1900 census is E. P. Hutchinson, cited in footnote 5. The occupational categories are not well suited for measuring economic status, but I select certain occupations that should be unambiguous indicators of high or low attainment. For one comparison, independent farmers along with five white-collar occupations (mainly professionals, agents, bookkeepers, merchants, and salesmen) designate high occupational attainment. This assumes that the farmer occupation, which excludes farm laborers, represented higher economic status than most urban jobs in 1900. Three laborer occupations (mainly servants and waiters, other laborers in the service industry, and workers in the low-paying textile industry) designate low occupational attainment.

The 1900 census reported that 33 percent of all white men in the labor force were farmers or were in the five white-collar occupations, compared to 24 percent among the second-generation Irish-Americans. Restricting the comparison to the urban labor force, 18 percent of all white men were in the five white-collar occupations compared to 15 percent of the second-generation Irish. The three laborer occupations contained 18 percent of all white men and 23 percent of the second generation Irish.

The comparison between second-generation Irish-American men and all second-generation white men gives similar but weaker findings, based mainly on the sharp contrast of representation in the low occupations. Only 14 percent of the second-generation whites were in the three low occupations compared to 23 percent of the second-generation Irish. The two second-generation groups had similar percentages in farming and in the five white-collar occupations.

The statistics in the above two paragraphs show a lag in the occupational attainment, and presumably in the incomes, of the second-generation Irish compared with other white Americans. (Further details and citations for these comparisons are given in the Appendix A.)

The article on the Irish in the Harvard Encyclopedia of American Ethnic Groups gives a more pessimistic picture of the occupational attainments of Irish-Americans. Patrick J. Blessing states that the Irish "were the only immigrant group whose occupational mobility during the late 19th century appeared almost as small as that of American blacks . . . . Their . . . record of movement up the occupational scale [was] dismal."[7] Although the occupational statistics from the 1900 census do not seem to me to justify this judgment, which may have referred to the Irish experience in Boston, I conclude that the statistical analysis supports the verdict of discrimination against Irish-Americans at this period in our history.

In 1940 nearly all the Japanese-Americans in the United States lived in the four western states of Arizona, California, Oregon, and Washington, with about 85 percent in California. The first generation had immigrated to the United States between 1890 and 1920. By 1940

almost two-thirds of the Japanese-Americans were born in the United States. There is abundant evidence that they faced severe social and legal discrimination, especially in California. Again, however, statistical evidence on how this discrimination affected their incomes and earnings in 1940 is not directly available. Instead, I compare the occupations of second-generation Japanese-American men in the four western states with the occupations of all white men in California.

The occupations in 1940 are defined in modern terms, permitting a rough ranking by skill, although there are the customary distinctions between the white-collar occupation "clerical and sales" and the blue-collar occupation "craftsmen" where a ranking according to skill is uncertain. To measure the occupational attainment for the two groups, white men and second-generation Japanese-American men, I calculate average earnings for each of the two groups as follows. First, the median wage and salary earnings of each occupation is assigned to that occupation as a cardinal measure of its rank. There are eleven occupational categories, with professionals ranked highest and domestic servants ranked lowest according to the earnings measure. Second, the percentage of each ethnic group in the occupation is multiplied by the occupation's median earnings, and the sum of these products gives an average earnings for each ethnic group, although for the Japanese-Americans the average uses the occupational earnings of all the workers, who are 95 percent white.

The ratio of this dollar-valued occupational attainment of Japanese-Americans to that of all whites is .92, implying a moderate lag in the attainments of Japanese-Americans. The ratio is probably

upwardly biased. The method assumes that the earnings of Japanese-Americans and whites are the same within an occupation. Another source of bias is that second-generation Japanese-Americans had more education than whites in 1940, indicating that a control for educational attainment would show a lower ratio than .92. (The details of the above calculations and arguments are given in Appendix A.)

Let us assume, then, that a comparison of incomes in 1940 would reflect these findings about occupational differences between whites and second-generation Japanese-Americans. We may conclude, therefore, that the statistical analysis again supports a verdict of discrimination.

If we carry out income and occupational comparisons for the same two ethnic groups in 1970 and 1980, however, we find that the Irish- and Japanese-Americans have higher income, occupational, and educational attainment than other white Americans. This claim is stronger and clearer for Japanese-Americans, who have sharply higher family or per capita incomes than comparable native-white Americans.

In Table 3 two comparisons are shown for family income.[8] For married-couple families the income of Japanese-Americans is 47 percent higher than that of non-Hispanic whites; 26 percent higher if we look at just the primary earner's income. Not shown is a more technical analysis in which the incomes are standardized for age, education, marital status, and region of residency of the primary earner. With these characteristics taken into account there remains an 8 percent advantage in favor of Japanese-Americans. The last part of the table shows higher occupational attainments of Japanese-Americans.

Table 3

Income and Occupational Attainment of American-Born
Japanese and Non-Hispanic White Americans, 1980

	Japanese-Americans	Non-Hispanic White	Ratio, J/W
Average income of married-couple families	$41,700	$28,300	1.47
Average income of married-couple families excluding income of family members other than the primary earner	$26,900	$21,300	1.26
Occupations: Percentage in (a) Professional, Technical Sales and Administrative	54%	44%	1.23
(b) Operators, Fabricators, Laborers, Unemployed	13%	25%	.52

Source: Harriet Orcutt Duleep, "The Economic Status of Americans of Asian Descent: An Exploratory Investigation," U.S. Commission on Civil Rights, Clearinghouse Publication 95, October 1988, pp. 35, 70, 73.

The evidence for higher incomes of Irish-Americans is not so strong, and the amount of their advantage is not so large. By 1970 the family incomes of men who reported Irish ancestry were about 5 percent above that of other white families.[9] Also, 48 percent of the men were in white-collar occupations, compared to 43 percent of other white men. One problem in measurement here is that Irish ancestry is defined by both parents being Irish. But by 1970 intermarriage between Irish-Americans and persons of other ancestry was common, so there is a selection of those who are recorded as Irish that lends an unknown bias to the income and occupation comparisons.

One reason why the economic success of Irish-Americans in overcoming discrimination is understated during this century is that discrimination was mainly against the Catholic Irish, and today the incomes and educational attainments of Irish Catholics are greater than for the Irish as a whole.[10]

We now have a puzzle, or maybe the right word is dilemma. If we accepted the 1900 and 1940 statistical evidence as supporting a verdict of discrimination against these two ethnic groups, then how should we interpret the 1970 and 1980 evidence? Does the latter imply discrimination in favor of Irish- and Japanese-Americans in contemporary America? Was there reverse discrimination in 1970 and 1980? If we dismiss this interpretation, should we then call into question our verdicts of discrimination in 1900 and 1940? Our investigatory procedures should be consistent. I believe, and this is my principal theme, that consistency is achieved by supplementing the statistical evidence with institutional and historical evidence and with theoretical

analysis. Indeed, my preference is to consider statistics as supplementary to the institutional-theoretical analysis.

The limitations of space and of my knowledge about the history of ethnic groups in America prevent an extensive discussion about how we should interpret the potentially inconsistent evidence presented above. The case of Japanese-Americans, however, appears to offer a straightforward reconciliation in the light of several well-known historical facts.

I claim that the statistical evidence supports the verdict of discrimination in 1940 and a verdict of no "reverse discrimination" in 1980. The reasons for the conclusion about 1940 are easiest. The record in the United States of legal and social persecution of Japanese immigrants and their offspring is astounding and appalling, culminating in the effective robbery of much of their land and wealth in 1942 when almost all Japanese-Americans were imprisoned in concentration camps.[11] Their lower incomes in 1940 reflect, as we see with subsequent evidence, discriminatory barriers rather than voluntary choices or any innate incapacity for economic success.[12]

The explanation for their economic success since the end of World War II, shown by the income statistics for 1980, is not so easy. They have benefited by living in the most prosperous part of the United States, the West Coast, especially California. Also, there were, apparently, two strong selective forces at work on the Japanese-American population. One is that the Japanese government selected the immigrants who went to the United States, beginning around 1890 and until 1920 or so.[13]

A second selective factor is that about 50 percent of the immigrants returned to Japan.[14] A natural inference is that the most successful stayed in the United States, and supporting testimony is given by Dorothy Swaine Thomas: "At the time of the Immigration Commission's investigations in 1908-1909, a majority of all classes [of the Japanese-Americans] interviewed either expected to return to Japan or were undecided about their future place of residence. The proportions indicating that they had decided to become permanent residents of the United States were, however, highest among those who had 'succeeded in rising from the ranks of the laboring classes.'"[15]

We could, of course, appeal to various theories about the special work ethic and efficiency of the Japanese people to explain their economic achievements. In the case of the Japanese-Americans, however, these theories are not needed.

Fortunately, today we have data and tools of analysis that permit us to use both statistical and institutional-theoretical methods to analyze discrimination. Unfortunately, both methods are inherently imprecise, and reasonable people can disagree about the uses and interpretations of the evidence. I turn next to the current use of statistical evidence to measure labor market discrimination.

III. STATISTICAL MODELS OF WAGE DISCRIMINATION

One of the earliest econometric studies to measure labor market discrimination against women was by Henry Sanborn and based on data from the 1950 census. Sanborn concluded that the wages of women were only moderately lower than those of men of the same region of residence, age,
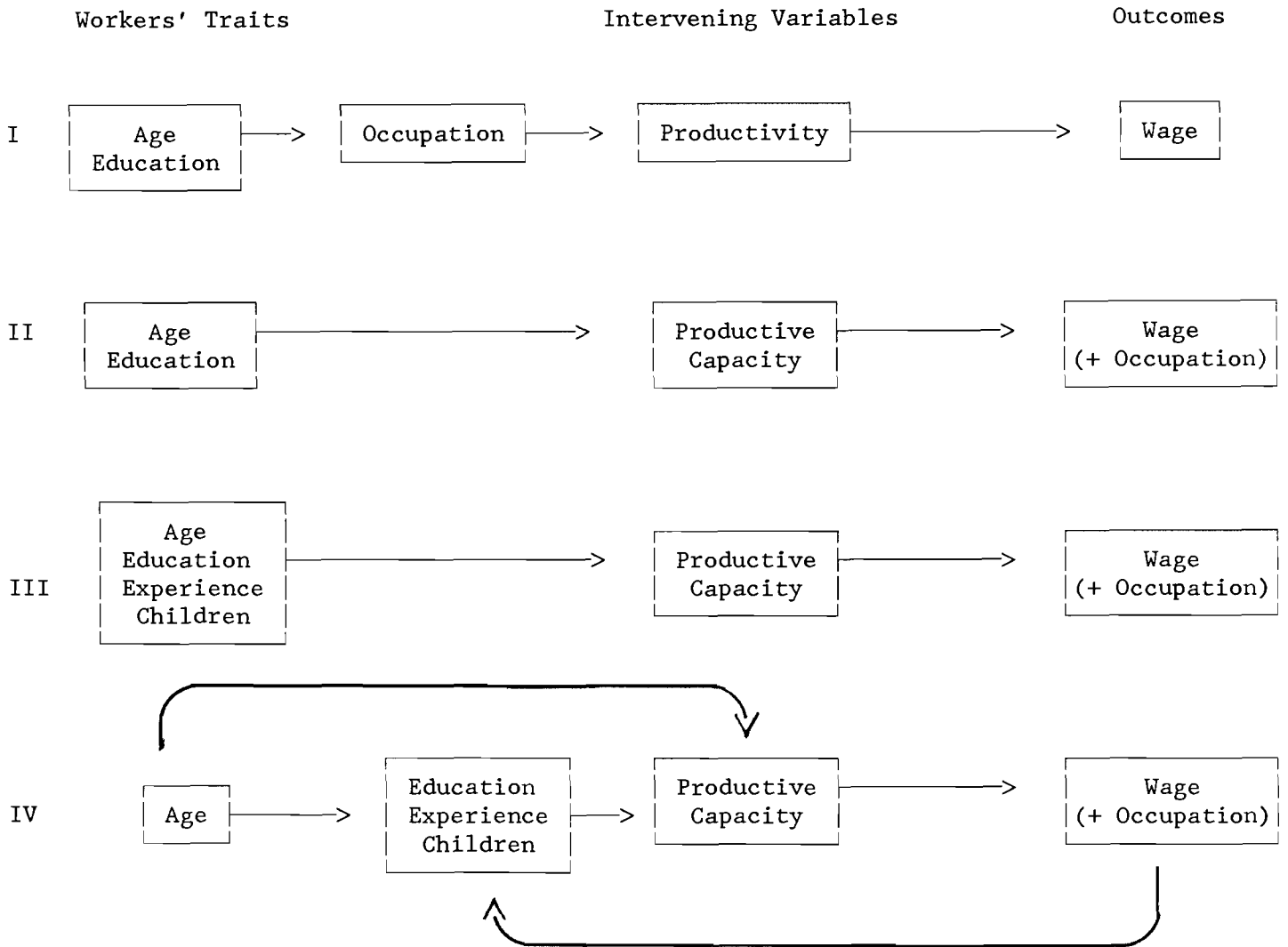
education, and occupation.[16]   In 1973 a study by James D. Gwartney and

Richard Stroup based on the 1960 and 1970 censuses found a sizable

gender wage gap and extensive discrimination against women workers.[17]

The different conclusions had nothing to do with the different

years for the data sources.   Instead, there are reasons for believing

that the first study understated, and the second study overstated, labor

market discrimination--defined as unequal pay for the same productive

capacities.   The early study by Sanborn compared the wages of men and

women in the same occupation, but because barriers to occupational entry

are the most important form of labor market discrimination, this study

surely understated it.   A diagram of this model is shown in the first

row of Figure 1.   By controlling for occupation in the statistical

analysis, the researcher may be said to have "overcontrolled."

The models shown in Figure 1 are deliberately simplified and are

intended to serve pedagogic purposes.   They should not be viewed as

representing the precise models used in the research literature I cite.

With this qualification in mind, let us interpret Model I.   Arrows

denote causal paths.   The right-side column headed "Outcomes" is self-

explanatory, except for the important distinction between Model I, where

only the wage is listed, and the other models, which treat occupation as

another outcome that measures labor market success or failure.   The

conventional economic proposition that workers' wages are determined by

their productivity is represented in column 3 (which is the second of

two columns under the heading "Intervening Variables").   A distinction

is made between "productivity" in Model I, which is associated with the

occupational skills of worker, and "productive capacity," which

Figure 1

Analytic Models of Labor Market Discrimination: Women
(Arrows denote causal paths)

determines, among other things, the worker's ability to enter various occupations. Once workers are in a particular occupation their productivity may be enhanced by on-the-job training specific to that occupation.

The first column lists variables or traits of workers that may sometimes be assumed to be exogenous to the workings of the labor market. In the case of education the assumption I adopt for convenience in Models I-III is that education is determined before the person enters the labor market and is unaffected by the operation of the labor market.

Model II in Figure 1 describes the assumptions behind the 1973 study by Gwartney and Stroup that showed a large gender wage gap. The 1973 study, however, did not allow for the difference between men and women in their years of experience in the labor market--a gap which is related to the two careers, home and market, of many women. Without controlling for labor market experience, the observed lower average wage for women probably overstates labor market discrimination against women.

Later studies of the gender gap in wages allowed for the effects of the woman's previous employment experience either indirectly, by including a fertility variable, or with a variable that directly measured the woman's previous years of labor market experience.[18] (See Model III in Figure 1.) The gender gap in wages was much smaller in these studies. In my view, however, accounting for women's dual careers by statistical methods is an intractable problem in measuring labor market discrimination. Let me try to explain my view.

Recall the two sources of complexity caused by women's dual careers: first, that their choice of a less demanding job may be

voluntary, perhaps as a consequence of their total family income;
second, that the housework demands on their time and energy reduce their
productive capacity in the paid labor market. These interpretations
rationalize statistical strategies that include the following sorts of
control variables when comparing women's and men's wages: full- or part-
time work, years of experience in the paid labor market, numbers and
ages of children, and marital status.

These control variables are usually interpreted to be explanations
for the lag in women's wages that was otherwise attributed to labor
market discrimination. This conclusion, however, may be challenged with
an alternative interpretation that is based on the mutual causation
between women's wages and the variables measuring their productivity.
In this model (IV in Figure 1), the restricted opportunities for women
in the labor market that are attributable to discrimination are causal
to their concentration on housework at the expense of market work.
Statistical analysis can measure associations, but it cannot determine
causation; at least, not without a supporting theoretical
rationalization.

This theoretical point about mutual causation may be explained
without mentioning labor market discrimination. During recent decades
employment opportunities in the labor market have improved for women,
and women have postponed their age of first marriage, had fewer
children, and increased their rates of marital breakup.
But what is cause, and what is effect? Was the decision to have fewer
children independent of what was occurring in the labor market, and did
increased market work by women then follow the decline in the birth

rate? Or were the improved wage and job opportunities in the labor market causal to the decision to have fewer children? Economists have not been successful in identifying and quantifying the causal forces in these events. It is even more difficult to determine the causal role of discrimination in the labor market in these outcomes.

IV. SUMMARY

I am pessimistic about the power of statistical methods to determine and measure economic discrimination. We are not sure what variables should be used as control variables in the analysis. If relevant causal variables are omitted, the model "undercontrols" and, therefore, overstates economic discrimination. If, however, the control variables are themselves effects of labor market discrimination, then they should not be included in the model; this is the problem of "overcontrolling," which leads to an understatement of economic discrimination.

These statements are expressions of humility but not of despair. We can take a lesson from the historical studies, such as those alluded to above concerning Irish-Americans and Japanese-Americans. We need a theoretical framework supported by knowledge about the institutional and historical factors. In the modern context of assessing discrimination by employers and labor unions, we need to study the histories of the firms and unions to try to determine motivations and intentions. Statistical methods are indispensable in the task of measuring labor market discrimination, but they are still only one component of the analysis.

Appendix A

The Evidence for Economic Discrimination Against Irish-
and Japanese-Americans in 1900 and 1940

In the text of this paper I assumed the following hypothetical

proposition. If we had for 1900 and 1940 the type of economic data that

we have had since, say, 1960, we would be able to show economic

discrimination against Irish-Americans in 1900 and against Japanese-

Americans in 1940 by conventional methods of economic research. The

evidence mentioned in the text to support this proposition is admittedly

meager. I doubt that the necessary data exist from these earlier years

for applying conventional methods. This appendix discusses in more

detail the evidence that is available. First, a theoretical point.

The Economic Theory of Labor Market Discrimination

The point was made in the text that merely documenting examples of

employer discrimination against a minority group is not sufficient to

indicate that the incomes or earnings of the minority group are lower

than those of the majority group for workers of the same productive

capacity. The documentation is necessary for a verdict of economic

discrimination, because it provides the historical and institutional

evidence for how discrimination occurred. Statistical evidence showing

lower earnings of the minority group is also necessary, however.

In the case of Irish-Americans at the turn of this century, for

example, it is not enough to refer to the legacy of the infamous signs,

"No Irish Need Apply." Knowing that some, even many, employers

discriminated against the Irish does not tell us that the employers who

did not were too few to enable Irish workers to earn a wage equal to that of other workers with their productive capacity.

In the hypothetical example of the text, where minority workers were 10 percent of the work force, and half of the employers did and half did not discriminate against the minority group, the claim was that competitive forces would secure pay for the minority workers in accordance with their productive capacity. The reason is that the same competitive forces that tend to equate the wage and productivity of majority workers would accomplish this for the minority workers. In other words, we should expect that the labor demand for minority workers by the 50 percent of employers who do not discriminate will provide enough bidding to ensure the same competitive solution for the minority group.

Interestingly, and importantly, a large percentage of employers relative to the percentage that the minority group constitutes is not necessary to achieve a nondiscriminatory outcome. The argument runs as follows. Assume that only a few employers do not discriminate against the minority group. They will hire the minority workers because, unlike the discriminating employers, they suffer no nonpecuniary disutility from employing minority workers. At the outset the nondiscriminating employers have no incentive to pay the minority worker any more than their market wage, which, by assumption, is less than the prevailing wage paid to workers in the majority group. By employing the equally productive but lower-paid minority workers, the nondiscriminating employers will make extra profits, and, to repeat a point, they suffer no disutility from employing the minority workers. The

nondiscriminating employers will expand production in response to their extra profits, and this increases the demand for and wages of minority workers.

The discriminating employers, who are losing their share of the market output and losing their minority workers, will be forced to raise the wages of the minority workers if they want to keep them. In any case, the bidding for the labor of minority workers by nondiscriminating employers will continue to raise the wages of the minority workers until their wage equals that of the majority workers. Only at this point will excess profits disappear, permitting an equilibrium allocation of the share of output among the employers.

The point of this argument is that a comparison of earnings of minority and majority workers is needed to tell us whether the demand for labor by employers and consumers has produced a discriminatory outcome by the economist's definition. Recall, however, that some employers may be discriminating in the legal sense even though the market shows equality in earnings of the two groups.

Available Data on Irish-Americans

In view of the lack of data on incomes, how might we measure the economic status of Irish-Americans around the turn of the century? Several historical works have examined the occupational attainments of Irish-Americans in specific cities. Stephan Thernstrom, for example, reports relatively slow progress in occupational attainment by the Irish in Boston in the period around 1900.[19] Andrew M. Greeley argues, however, that the historical record in Boston understates the

occupational achievements of the Irish in the United States as a whole.[20]

Nationwide data on the occupations of the labor force from the Bureau of the Census seem to me to be our best source. From 1870 to 1900 the decennial censuses reported the occupations of American workers classified by nativity and country of origin. This information is summarized in the book, cited in footnote 5, by E. P. Hutchinson. Unfortunately, the occupational classifications used for those censuses were based mainly on the worker's industry. Hutchinson commented that "The primarily industrial basis of classification largely conceals whatever characteristics of skill or occupational status the . . . [worker] may have had."[21] For example, in the 1900 census the occupation of "iron and steel workers" includes unskilled and skilled workers.

Despite this problem, I am forced to use the 1900 occupational classifications instead of the those in the 1910 census, which do correspond to our current occupational classifications, because the 1910 census did not report the country of origin of the workers. In fact, the 1950 census is the next one to provide occupational data for workers according to their country of origin.

Although the 1900 occupational classifications do not, in general, permit a ranking by skill, there are several "high" and "low" occupations that clearly correspond with economic status. (Unless otherwise noted, the statistics that follow are from Hutchinson, pp. 172-175.) From 34 listed occupations from the 1900 census, I select the following five high occupations: (1) professional service, (2)

agents, (3) bookkeepers and accountants, (4) merchants and dealers, and (5) salesmen. Occupations 2-5 are all contained in the "trade and transportation" industry. A large and relatively high occupation in 1900 is that of "farmers, planters, and overseers." Agricultural laborers are excluded, so this farm occupation probably has a higher status than the occupations of most urban workers. The low occupations that I selected are (1) servants and waiters and (2) other laborers, both from the "domestic and personal service" industry and (3) "textile mill operatives," selected because the textile industry was the lowest-paying manufacturing industry.[22]

The occupational data from the 1900 census report whether the worker is born in Ireland, designating a first-generation Irish-American, or has foreign-born parents, designating a second-generation Irish-American. The Irish-American male workers are compared to all white male workers and to all second-generation white male workers. Table A-1 summarizes the occupational statistics for these four groups. As noted in the text, in 1900 the all-worker group was composed of native workers (57 percent), foreign-born workers (23 percent), and second generation foreign workers (20 percent).

The first evaluation of the occupational attainments of second-generation Irish-American males compares them with all white males. Row 3 of Table A-1 shows that 24 percent of the second-generation Irish are in the farmer occupation and the five high (white-collar) occupations, and the corresponding figure for all whites is 33 percent. In the urban sector 15 percent of the second-generation Irish are in the five high occupations and 23 percent are in the three low occupations. They

Table A.1

Occupational Distribution of Selected Occupations for
Male Irish-Americans and All White Males, by Nativity, 1900

Occupational Groups	Ethnic Group			
	All Whites	2nd Generation Whites	2nd Generation Irish-American	1st Generation Irish-American[a]
In total labor force:				
1. Farmers, planters and overseers[b]	22%	13%	12%	12%
2. Five high occupations[c]	11	12	12	8
3. Sum of 1, 2	33	25	24	20
In urban labor force:				
4. Five high occupations[c]	18	16	15	10
5. Three low occupations[d]	18	14	23	42

Notes:

[a]1st generation is the same as foreign born.

[b]Does not include agricultural laborers.

[c]Professionals; agents; bookkeepers and accountants; merchants and dealers; salesmen. (The latter four are in the trade and transportation industry.)

[d]Laborers; servants and waiters (both groups from the domestic and personal service industry); textile mill operatives.

Source: E. P. Hutchinson, Immigrants and Their Children, 1850-1950, New York: John Wiley and Sons, 1956, pp. 159, 172-174.

compare unfavorably to the urban occupational attainments of all whites--18 percent in the five high occupations and 18 percent in the three low occupations.

Another comparison is between the second-generation Irish and the second generation of all whites. The two groups had nearly the same percentage, 24 and 25, respectively, in the combined farmer and five high occupations. In the urban sector, the occupations of the second-generation whites show somewhat higher attainments: 16 percent are in the five high occupations and only 14 percent are in the three low occupations, compared with 15 and 23 percent, respectively, for the second-generation Irish. The occupations of the first-generation Irish are considerably lower in all comparisons.

Available Data on Japanese-Americans

The 1940 census publications included a special report that gives demographic, educational, and occupational information, but no wage and salary data, for Japanese-Americans by country of birth.[23] The data cover the four western states, Arizona, California, Oregon, and Washington, where nearly all Japanese-Americans lived. The occupations are limited to eleven broad categories. Another problem is the small population of Japanese-Americans in 1940. Of the total of 127,000, 63 percent (80,000) were born in the United States, but 50,000 of the latter were less than 20 years old. As a result there were only 15,000 second-generation Japanese-American workers in 1940. By contrast, there were over one million second generation Irish-Americans in the labor force in 1900.

Adding to the difficulty of using the occupational data to evaluate

the economic attainments of Japanese-Americans are the following

problems:

(1) About 14 percent of the second-generation Japanese-American

men and 5 percent of white men worked in agriculture. In 1940,

unlike 1900, we cannot assume that farmers have higher incomes

and economic status than most urban employees, so I exclude

farmers from the occupational comparisons. Farm laborers are

included because the low wages in this occupation are a clear

indication of its low rank.

(2) Second-generation Japanese-American workers were, on average,

much younger than white workers. Because occupational

attainment depends on age, the occupational comparison needs

to take account of the age differences in the two populations.

(3) Two occupations, "proprietors, managers, and officials

(excluding farmers)" and "clerical, sales, and kindred

workers" are likely to consist of jobs with lower status and

lower workers' incomes among the second-generation Japanese-

Americans than among white workers. Dorothy Swaine Thomas

reports that, among the Japanese-Americans, "the personnel of

so many enterprises consisted merely of an Issei [first-

generation] 'entrepreneur' and his cooperating

relatives. . . . Thus, the concentration of Nisei [second-

generation] males at 'white-collar' levels . . . may be

interpreted to mean that many of those classified in the

proprietor-managerial group were in fact working in secondary

positions in the larger, Issei-controlled enterprises, and
that most of those classified in the clerical-sales category
were, similarly, working for other Japanese. The few who had
progressed beyond the limits of the Japanese business
community were, for the most part, on low rungs of the urban
occupational ladder and held the less desirable jobs."[24]

Each of the problems listed above is addressed below in the
comparisons of occupational attainments of Japanese-American and white
workers.

The 1940 census included, for the first time, wage and salary
information for the labor force, and this permits a quantitative
economic scale for the occupations. (No data on occupational earnings
are available for Japanese-Americans, however.) Given the concentration
of Japanese-Americans in California (about 85 percent lived there), I
use the occupations and wages of male workers in this state, 95 percent
of whom were white, to construct a dollar value for the occupational
attainments of the white and Japanese-American workers. Farmers are
excluded, as noted above, and the median wage and salary earned by
"professional workers" is assigned to the "proprietors, managers, and
officials" because income data from self-employment were not
collected.[25]

The occupational distributions and median earnings are shown in
Table A-2. An average earnings of white workers, excluding farmers, is
calculated by multiplying the median earnings of each occupation by the
percentage of white workers in the occupation and then summing the
products. (To be precise, the median earnings are for workers who

Table A.2

Occupational Distribution of White Males and Second Generation
Japanese-American Males, Evaluated by Median Earnings of the
Occupations, 1940

Occupation	Median Earnings of Occupations in California[b] (1)	Percentage in Occupation		
		All White Males California[b] (2)	Japanese-Americans, 2nd Generation	
			Reported (3)	Adjusted[c] (4)
Professionals and semi-professionals	$2263	7.9 [8.3]	2.5 [2.9]	[4.2]
Farmers and farm managers	--	5.0	13.8	--
Proprietors, managers, and officials (exc. farms)	2263	13.0 [13.7]	7.3 [8.5]	[14.0]
Clerical, sales, and kindred	1680	16.8 [17.7]	16.0 [18.6]	[30.6]
Craftsmen, foremen, and kindred	1730	17.6 [18.5]	2.4 [2.8]	[4.3]
Operatives and kindred	1543	16.8 [17.7]	8.9 [10.3]	[16.4]
Service workers, except domestic	1176	8.7 [9.1]	3.6 [4.2]	} [6.5]
Domestic service	811	0.2 [0.2]	3.0 [3.5]	
Laborers, except farm	1276	6.1 [6.4]	33.0 [38.3]	} [24.0]
Farm laborers	820	7.8 [8.2]	9.5 [11.0]	
Earnings, averaged over occupations, excluding farmers (weighted by occupation percentage)		$1646	$1435	$1601

(Continued)

Table A.2, Continued

Sources: Columns 1-2: U.S. Bureau of the Census, <u>16th Census of the United States, 1940,</u> <u>Population, vol. III. The Labor Force</u>, Part 2, Reports by States, California (Washington, D.C.: U.S. GPO, 1943) pp. 232-233, 278-281.

Column 3: U.S. Bureau of the Census, <u>16th Census of the United States, 1940,</u> <u>Population</u>, "Characteristics of the Nonwhite Population by Race" (Washingotn, D.C.: U.S. GPO, 1943) pp. 107-109.

Column 4: Dorothy Swaine Thomas, <u>The Salvage</u> (Berkeley, Calif.: University of California Press, 1952), p. 600.

[a]The bracketed percentages exclude the occupation of "farmers and farm managers."

[b]Because about 85 percent of the second-generation Japanese-Americans lived in California, the occupational distribution and median earnings-by-occupation of California males are used for the comparisons. About 95 percent of the Californian male work force was white. The median earnings are derived from the "wage and salary" earnings for workers in these occupations who worked 12 months in 1939. The "farmer" and "proprietors, etc." occupations have no meaningful earnings reported. For "proprietors, etc.," the median earnings of the "professional" occupation are assigned. See appendix text for the justification.

[c]The second-generation Japanese-Americans are, on average, younger than all males. Using the all-male age distribution of the states of Washington and California, Dorothy Swaine Thomas calculated an age-adjusted occupational distribution of the second-generation Japanese-Americans. The age adjustment assumes the same age-specific occupational percentages of the Japanese-Americans but applies to those percentages the all-male age distribution. See appendix text for further explanation.

worked 12 months in 1939.) This average, shown in Table A.2 to be $1646, is a dollar measure of the occupational attainments of white workers.

The same method is used for evaluating the occupational attainments of second-generation Japanese-American workers. Their occupation percentages are multiplied by the white median earnings in each occupation. Thus, the dollar measure of the occupational attainments of Japanese-Americans assumes that Japanese-American and white workers in the same occupation received the same earnings. As shown in Table A.2, the ratio of the dollar value of the Japanese-American occupational attainments to that of white workers is .87 (= 1435/1646).

A second ratio of the two groups' occupational attainments uses the same earnings figures, but the occupational distribution of the Japanese-Americans who lived in the states of California and Washington in 1940 is adjusted to match the age distribution of the entire population of workers in these states. (About 90 percent of all Japanese-Americans lived in these two states in 1940, and about 95 percent of the entire population in the two states was white.) The age-adjusted occupational distribution of Japanese-Americans is taken from Thomas.[26] The method may be explained with a simple example.

Assume that 5 percent of the Japanese-Americans over 35 years of age, and 2 percent of those under 35 years of age, were professionals, and that 80 percent of the Japanese-American population of workers were under 35. Thus, their overall percentage professional is 2.6 percent (=.8 X .02 + .2 X .05). Assume the white population has 50 percent of its labor force over 35. Then if the Japanese workers had the same age

distribution as whites, 3.5 percent (=.5 X .02 + .5 X .05) would be in the professional occupation. The age-adjusted percentage of professionals for the Japanese-Americans is, therefore, 3.5 percent.

The ratio of the dollar value of the Japanese-American occupations to that of the white workers, adjusting for age, is .97 (=1601/1646). The ratio without the age adjustment, .87, is biased down because of the younger ages of the second-generation Japanese-Americans. The ratio of .97 is biased up because of Thomas's point about the lower status and earnings of the large age-adjusted percentage (44.6) of Japanese-Americans in the proprietors and clerical/sales occupations, relative to the status and earnings of whites (28.6 percent) in these occupations. Both ratios are probably biased up because the calculation assumes the same median earnings of Japanese-Americans and whites within each occupation. Impressionistic evidence suggests that the earnings of Japanese-Americans were lower than those of whites. I simply averaged the two ratios and used .92 in the text to measure the lag in occupational attainment of Japanese-Americans.

Another reason why the occupational measure may understate discrimination against second-generation Japanese-Americans in 1940 is that it does not allow for their higher educational attainment relative to that of whites. Thomas reports an age-adjusted distribution of years of school completed for the second-generation Japanese-Americans who were 25 years of age or older.[27] Her data show that the median years of school completed in 1940 was 11.2 for the Japanese-American males and 10.6 for all males.

The educational advantage of the Japanese-Americans, therefore, did not translate into an occupational advantage. This leads to the observation by Thomas that among the Nisei, "Engineers, accountants, teachers, and social workers found it almost impossible to practice the skills they learned."[28] Supporting this point are the following amazing statistics reported in another study: among male Japanese-American college graduates who entered the labor force before 1942, only 10 percent went into professional jobs. In the 1950s, by contrast, about 70 percent of Japanese-American college graduates entered professional occupations.[29] All this supports the conclusion that second generation Japanese-Americans faced considerable economic discrimination in the period around 1940.

Notes

[1]See Appendix A for a fuller explanation of how employer discrimination can coexist with the economist's finding of no labor market discrimination.

[2]Reed Ueda, "Naturalization and Citizenship," in Stephan Thernstrom, ed., Harvard Encyclopedia of American Ethnic Groups (Cambridge: Harvard University Press, 1980), pp. 740-742.

[3]Glen G. Cain, "Black-White Differences in Employment of Young People: An Analysis of 1980 Census Data," Institute for Research on Poverty Discussion Paper no. 844-87, University of Wisconsin-Madison, 1983.

[4]Appendix A describes the use of census and other data from 1900 and 1940 to measure economic discrimination against Irish-Americans and Japanese-Americans.

[5]E. P. Hutchinson, Immigrants and Their Children, 1850-1950 (New York: John Wiley, 1956), p.159.

[6]Robert Higgs cites statistics from the 1910 report of the Immigration Commission of the U.S. Congress that show that the immigrant (foreign-born) Irish-Americans who were surveyed were 100 percent English-speaking and 96 percent literate. See Higgs, "Race, Skills, and Earnings: American Immigrants in 1909," Journal of Economic History, June 1971, pp. 420-428.

[7]Patrick J. Blessing, "Irish," in Harvard Encyclopedia of American Ethnic Groups, p. 531.

[8]Table 3 and the statistics comparing Japanese-Americans and white

Americans in this paragraph are from Harriet Orcutt Duleep, "The

Economic Status of Americans of Asian Descent: An Exploratory

Investigation," U.S. Commission on Civil Rights, Clearinghouse

Publication 95, October 1988, pp. 35, 70, 73.

[9]The statistics in this paragraph are from U.S. Bureau of the

Census, Current Population Reports, Series P-20, No. 249,

Characteristics of the Population by Ethnic Origin: March 1972 and 1971

(Washington, D.C.: U.S. GPO, 1973).

[10]This is a point forcefully made in two books by Andrew M.

Greeley: That Distressful Nation (New York: Quadrangle Books, 1972);

and The Irish Americans (New York: Harper and Row, 1981).

[11]Although the official justification for the confinement of

Japanese-American citizens who lived on the west coast was that they

were a military threat, there was no evidence then or later to support

this allegation. Persuasive reasons for their imprisonment are (1) the

history of racial hostility toward Japanese-Americans, especially by

white Californians; (2) the opportunity for whites to acquire assets of

Japanese-Americans when they were rushed to the concentration camps and

had to leave their belongings behind or sell at severe losses; and (3)

the economic gains of various white producer groups such as farmers or

small businessmen, who competed with the Japanese-Americans in these

businesses. No German-Americans nor Italian-Americans were put in

concentration camps, so it is hard to deny that the basic source of this

treatment of Japanese-Americans was racial discrimination.

[12]A study by the Carnegie Corporation during the 1930s reached the

same conclusion: ". . . the Nisei [second-generation Japanese-

Americans] were as 'bright' as Caucasians as measured by I.Q., their rates of crime and delinquency were low, and their records of achievement in school were extraordinary. The study found that race prejudice, not low achievement, was the fundamental cause of the Nisei's employment problems" (Harry H. L. Kitano, "Japanese," in the Harvard Encyclopedia of American Ethnic Groups, p. 565).

[13]Positive selectivity, with respect to productive skills, among the Japanese immigrants is stressed by Thomas Sowell, Ethnic America: A History (New York: Basic Books, 1981).

[14]Charles A. Price, "Methods of Estimating the Size of Groups," in Harvard Encyclopedia of American Ethnic Groups, pp. 1036-1039.

[15]Dorothy Swaine Thomas, The Salvage (Berkeley: University of California Press, 1952), p. 45.

[16]Henry Sanborn, "Pay Differences between Men and Women," Industrial and Labor Relations Review, 17 (1964), 534-550.

[17]James D. Gwartney and Richard Stroup, "Measurement of Employment Discrimination According to Sex," Southern Economic Journal, 39 (1973), 575-587.

[18]Jacob Mincer and Solomon Polachek, "Family Investments in Human Capital of Women," Journal of Political Economy, 82 (Part 2, 1974), S76-S108; Mary Corcoran and Greg J. Duncan, "Work History, Labor Force Attachment, and Earnings Differences Between the Races and Sexes," Journal of Human Resources, 14 (1979), 497-520.

[19]Stephan Thernstrom, The Other Bostonians (Cambridge: Harvard University Press, 1973). See especially pp. 135, 143, and 186-194.

[20]Greeley, The Irish Americans, p. 111.

[21]Hutchinson, p. 86.

[22]Industry wage levels for 1900 are reported in U.S. Bureau of the Census, Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition, Part 1 (Washington, D.C.: U.S. GPO, 1975), p. 172.

[23]U.S. Bureau of the Census, 16th Census of the United States, 1940, Population, "Characteristics of the Nonwhite Population by Race" (Washington, D.C.: U.S. GPO, 1943).

[24]Thomas, pp. 40-42.

[25]The 1950 census reported incomes in addition to wages, and in the census monograph Income of the American People, by Herman P. Miller (New York: John Wiley, 1956, p. 51), the incomes of the "proprietor, manager" occupation are shown to be nearly the same as those of the "professional" occupation: $4100 and $4250, respectively.

[26]Thomas, p. 600.

[27]Thomas, p. 611.

[28]Thomas, p. 42.

[29]Gene N. Levine and Darrel M. Montero, "Socioeconomic Mobility among Three Generations of Japanese Americans," Journal of Social Issues, 29 (no. 2, 1973), 33-48. This study was based on a sample drawn from lists of Japanese-American organizations in 1964-1966 to obtain about 900 surviving first-generation Japanese-Americans who lived in the continental United States. I do not know how well this sample represents the entire second-generation Japanese-American population, so I have not used it for deriving general statistics.