NBER WORKING PAPER SERIES

THE IMMIGRANT AND NATIVE-BORN WAGE DISTRIBUTIONS: EVIDENCE FROM UNITED STATES CENSUSES

Kristin F. Butcher
John DiNardo

Working Paper 6630
http://www.nber.org/papers/w6630

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1998

We would like to thank David Card, Anne Morrison Piehl, Jean Wohlever, and seminar participants at Boston College, Princeton University, University of Illinois, USC, and the NBER summer institute, for helpful comments. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

The Immigrant and Native-born Wage Distributions:
Evidence from United States Censuses
Kristin F. Butcher and John DiNardo
NBER Working Paper No. 6630
July 1998
JEL No. J31, J61, J70

ABSTRACT

Over the past thirty years, immigration has increased, immigrant characteristics have changed, and the relative mean wages of immigrants *vis à vis* the native born have declined. Using data from four U.S. Censuses (1960 - 1990) we examine changes in the wage structure and their role in explaining comparisons between immigrants and the native-born in mean wages. *Inter alia*, we document that patterns of comparison between the immigrants and the native-born are not the same for men and for women, and that these differences in immigrant/native-born comparisons among men and women are a consequence of different evolutions in the wage structure. Although virtually ignored in the immigration literature, we return to a well-understood aspect of Blinder/Oaxaca differentials: the extent of measured discrimination depends on the "base" prices used for comparison. Contrary to previous work which finds little impact of the wage structure on immigrant/native-born wage differentials, we observe that if the "wage structure" had remained as it was in 1970, for example, the decline in immigrant wages relative to the native-born would generally be much smaller than has been observed.

Kristin F. Butcher
Department of Economics
Boston College
Chestnut Hill, MA 02167
butcherk@bc.edu

John DiNardo
Department of Economics
University of California -- Irvine
Irvine, CA 92717
and NBER
jdinardo@uci.edu

1 Introduction

During the 1980s approximately 8 million legal immigrants came to the United States. This is higher than during any decade since the early 1900s.[1] While the characteristics of immigrants have changed over the past 30 years, the nature of the United States economy has also changed. The stagnation of real wages for many workers and the increase in inequality since the late 1970s has been especially well–documented. In this paper we investigate the role that these changes in the wage distribution have played in comparisons of immigrants to the native-born.

Most previous work comparing wages of the native-born and immigrants has analyzed differences in mean wages of immigrants and native-born men. This body of work documents the relative decline in mean wages of immigrant men (c.f. Borjas (1990)). There are two problems with this narrow focus. First, when the distribution of wages is *not* changing over the time period being studied, mean wages are sufficient to characterize the immigrant/native differential. When the distribution of wages has changed dramatically overtime, such comparison may not be "sufficient" to completely characterize the changing nature of the differences in economic outcomes between immigrants and the native-born. In this paper, we adapt techniques from DiNardo, Fortin, and Lemieux (1996) that allow us to examine changes in the entire distribution of wages for immigrants and the native-born from 1960 to 1990.

A second problem with the previous literature is that it generally ignores the outcomes of immigrant and native-born women. The notable exceptions, Duleep and Sanders (1993), Baker and Benjamin (1997),Blau (1992) tend to focus more on labor force participation and fertility than on relative wages. Our analysis includes women, and we find substantial differences between the relative changes in the wage distributions for the two sexes – for example, we find that immigrant and native-born women's wage distributions are more similar than those for men. Such a finding casts doubt on the descriptive adequacy of a "single index model" of immigrant characteristics by country of origin.

The techniques employed in this paper allow us to investigate the im-

[1] However the fraction foreign-born was more than twice as high at the turn of the century than it is today.

portance of changes in immigrant and native-born characteristics versus the changes in the wage structure in determining changes in the relative wages of immigrants and the native-born. We show how the wage distribution for recent immigrants would change if we held characteristics constant at their 1960s or 1970s levels, but allowed the prices for those characteristics to change to their 1990s levels. We find that the earlier immigrants would have had wages much more similar to today's new arrivals, if they had faced the present day prices for their skills.

This "counterfactual" exercise further allows us to pinpoint which of the recent immigrants' characteristics has the greatest impact on changes in the wage distribution between 1970 (or 1960) and 1990. Both the scientific and popular literatures commonly cite the relative decline in immigrant/native education levels as the most important attribute explaining the increasing wage gap. Our analysis, however, suggests that it is race and ethnicity which most distinguishes recent immigrants, and that race and ethnicity (or unobservables correlated with race and ethnicity) "explain" much of the change in the comparative economic fortunes of recent immigrants once wage structure changes have been held constant.

Finally, we decompose the changes in the native-immigrant wage gap into a portion attributable to changes in characteristics, and a portion attributable to changes in skill prices. In doing so we return to a well-known aspect of Blinder/Oaxaca decompositions which seems to have been ignored in the current debate: the group's prices that are used to evaluate the differences matter. For example, when decomposing black-white wage differentials, it matters whether one uses the black or white coefficients to evaluate the differences in characteristics. Similarly, here it is crucial whether one uses the 1990 skill prices to evaluate the differences in the immigrant/native characteristics, or the skill prices from a time when the wage distribution was less dispersed.

The paper proceeds as follows: Section 2 provides a brief summary of the immigration literature and presents comparisons of average hourly wages of immigrants and the native-born. Section 3 switches the focus from summary statistics to the *entire* wage distribution. Section 4 investigates "counterfactual" wage distributions for recent immigrants; we show what the wage distribution for 1960 and 1970 recent immigrants would have looked like if they had faced the 1990 wage structure. Section 5 decomposes changes in the immigrant/native wage gap into a portion attributable to changes in

the wage structure and a portion attributable to changes in characteristics. Section 6 concludes.

2 Background and Summary Statistics

The focus of the economics literature has been on the rate of "assimilation" – the degree to which a foreign born individual will be paid comparably to an otherwise comparable worker. Using cross-sectional data, Chiswick (1978) observed that the degree of assimilation varied with the length of time an immigrant had lived in the U.S. Borjas (1987) argued that the same evidence is consistent with "declining cohort quality", noting that identification of an assimilation profile from cross–sectional Census data required (untestable) assumptions about cohort effects. Using an adaptation of the Roy Model, he argues that the correlation of country–of–birth effects with various characteristics of the source countries is consistent with the view that the "quality of immigrants" has declined and that immigrants do not achieve parity with the native–born over time. He traces the declining cohort quality to the shift in immigrant source countries that occurred after entrance criteria were revised in the 1965 Immigration Act.

In this study, we re-evaluate the evidence on changes in immigrant earnings and immigrant characteristics over time. Our analysis is similar in spirit to LaLonde and Topel (1992) who note that most immigrant/native-born comparisons assume "constant period effects"; to wit, if, for example, immigrants have always been concentrated in the bottom tail of the wage distribution, and the wage distribution widens, then the immigrant native-born comparisons will show that immigrant's earnings have declined. This will be the case even with no underlying change in the characteristics of immigrants *vis a vis* the native-born. LaLonde and Topel recommend comparing immigrants to native-born workers who would have experienced similar wage changes due to the overall change in the distribution of earnings, for example, native-born Hispanics. Using this comparison group, they show substantial assimilation of immigrants. Yuengert (1994) makes a related point demonstrating that the point in the distribution where earning comparisons are made has implications for immigrant/native-born wage comparisons that are different than those that arise from computing simple means.

Our paper is not specifically about assimilation of immigrants – the

growth of wages with time in the United States. However, the literature on assimilation and the changes in immigrant skills are inextricably linked, since in cross-sectional data, immigrants from different cohorts will have been in the United States for different lengths of time. In order to avoid making "longitudinal" inferences from cross–sectional information, we focus on the wages of recent immigrants – those who have been in the U.S. for 5 or fewer years. We adapt techniques appropriate for analysis of the entire distribution of wages to disentangle two influences on comparisons of native-born and immigrant wages between 1960 and 1990: 1) the enormous changes in the structure of wages 2) the striking changes in immigrant characteristics.

After describing the data we use, we summarize the average characteristics and average wages for native- born, immigrants, and recent immigrants for the period 1960 to 1990.

2.1 Data

The data used in this paper are from the 1960, 1970, 1980 and 1990 Public Use Microdata Samples of the United States censuses.[2] The 1960 data are a 1 percent sample of the population. The 1970 data are also a 1 percent sample (from the 5% State sample). The 1980 and 1990 data are from the 5% samples, however, a only a random subset of each of these is used in the analyses presented here.

In all cases, the data refer to men and women between the ages of 16 and 65, inclusive. Individuals born abroad to American parents and those born in United States outlying areas (e.g. Puerto Rico) have been dropped from the analysis. Immigrants are defined as those who were born outside the United States. We also drop the self-employed.

There have been several changes in the way certain information is coded across the four censuses used here. In all cases, we attempted to define our variables in a consistent way. These changes affect some of the analyses we are able to perform. For example, although the 1970-1990 Censuses give fairly detailed information on the time period during which an immigrant entered the United States, the 1960 Census does not. The 1960 Census does, however, allow us to identify those who entered within the last five years. For

[2]The data refer to the year prior to the census date. E.g., the wage data are for 1959,1969, 1979, and 1989. However, we will refer to everything by the census date.

consistency, we define "recent" immigrants as those who immigrated within the last five years in all our data.

Much of our analysis focuses on log hourly wages. We calculate these from information reported on total annual earnings from wages and salary, number of weeks worked, and usual hours worked per week.[3] Many analyses using census data use weekly wages rather than hourly wages, since these calculated hourly wages may be prone to measurement error. However, weekly wages confound wage rates and labor supply. Furthermore, our comparisons of mean hourly wages of immigrants and the native-born look very similar to results from studies that use mean weekly wages, instead.

2.2 Descriptive Statistics

Tables 1a and 1b show summary statistics for the native-born, all immigrants, and recent immigrants for men and women in all four decennial censuses. There are a number of striking differences between immigrants and natives across the years. The first row of each panel shows mean age. In 1960, immigrants were on average ten years older than the native-born. With the increase in immigration during the 1980s and 1990s, the average age among immigrants falls to parity with the native-born. As one would expect, recent immigrants are younger than the native-born and other immigrants. Their average age has hovered around 30 years old across the four time periods.

Changes in immigrants' labor market skills have received a great deal of attention(Borjas 1990). The direct measure of skills that is most readily available is educational attainment. Tables 1a and 1b give the fraction high school drop out, fraction with exactly a high school degree, fraction with some college education, and the fraction with a college degree or above.[4] Although the fraction of immigrants with less than a high school degree fell between 1960 and 1990 (0.69 to 0.41 for men, 0.68 to 0.39 for women), this educational group has declined faster among the native-born. Fifty-six

[3]In 1960 and 1970 the data are only available in bracketed form. We assign the midpoint of each bracket and calculate the hourly wage rates.

[4]These variables are defined off of the census information on highest grade attended and highest grade completed in the 1960- 1980 Censuses. There was a change in the education codes in the 1990 census. In all cases, highest grade completed < 12 = h.s. dropout; highest grade completed $= 12$ = h.s. degree; $13 \leq$ highest grade completed < 16 = some college; highest grade completed ≥ 16 = college graduate.

percent of native- born men had less than a high school degree in 1960. by 1990 this number was down to 22 percent. Thus, by 1990 immigrants are substantially more likely to be high school drop outs than are the native-born. This comparison looks approximately the same whether we look at recent immigrants or immigrants overall, and at men or at women.

Interestingly, immigrants have always had approximately the same rates of higher education as the native-born. The fraction with a college degree or above increased for all groups from 1960 to 1990, but they are approximately the same for the native-born, immigrants, and recent immigrants within each year.

Perhaps the way in which immigrants have changed the most over the past four decades is in their racial and ethnic composition. In 1960, only 5% of the immigrant and 7% of the recent immigrant men were Asian or other race.[5]By 1990, 23% of immigrant and 27% of recent immigrant men were Asian. The changes in fraction Hispanic have also been dramatic. In 1970, 22% of immigrant men were Hispanic.[6] By 1990, this number had approximately doubled to 45%. The shifts are similar for women. These changes reflect the changes in the national origin mix of immigrants after the 1965 change in immigration laws (which was extended to include the Western Hemisphere in 1968 and thus took full effect sometime in the late 1960s)(Borjas 1991).

There are several other comparisons between immigrants and natives that are worth noting. Immigrants are, and always have been, more likely to live in a metropolitan area than are their native-born counterparts. Since urbanization has been increasing in the United States, this difference is smaller in the later periods. However, by 1990, over 90% of the immigrants and recent immigrants reside in metropolitan areas, as opposed to just over 80% for the native-born. Marriage rates have dropped from 1960 to 1990, however, both immigrant men and women are more likely to be married than their native-born counterparts in all years.

We present a concise summary of the changes in the observable skills of

[5]In 1960 and 1970 we combined Asian and "other race" due to the small samples of both.

[6]In 1960 an individual is defined as "Hispanic" if he or she had a Spanish surname. This is not the same definition in the other years. The other years all use the self-reported ethnicity information.

immigrants in tables 2a and 2b. Here we grouped immigrants into quintiles based on predicted log hourly wage distribution for the native-born.[7] If immigrants were a random subset of the population, we would expect to find 20% of immigrants in each quintile (see Borjas (1995) for a similar analysis). The top panel of each table is for all immigrants; the bottom panel is for recent immigrants. Among all immigrants, the 1960 data show that only about 10% of men and 8.5% of women were in the bottom quintile. The highest fraction of immigrant men was in the fourth quintile (28%), while women were most highly concentrated in the third quintile (33%). There is a clear shift from 1960 to 1990. There was a 93% increase in the concentration of immigrant men in the bottom quintile between 1960 and 1990. A large increase occurred for women as well (62%). It is worth noting, however, that based on their characteristics, by 1990 only 20.3% of immigrant men are in the bottom quintile of the distribution. Similarly, 20.7% are in the top quintile. The immigrant men are somewhat over represented in the second lowest quintile, and underrepresented in the second highest quintile. Although the changes were in the same direction for women, they look "better" compared to their native-born counter-parts. In 1990, immigrant women are *under-represented* in the bottom quintile (13.8%) and over-represented in the top quintile (24.8%).

The story for recent immigrants is different. Recent immigrant men have always been over-represented in the bottom quintile, however, this has increased. By 1990, 35.6% of recent immigrant men are in the lowest quintile. They increased their representation in the second quintile, and reduced it in the top three. Recent immigrant women increased their proportion in the bottom quintile, and reduced it in the middle three. However, in 1990, recent immigrant women are more heavily over-represented in the top quintile than in previous decades (24%).

The numbers in tables 2a and 2b simply summarize how the *observable* characteristics of immigrants changed.

Tables 1a and 1b also show the unadjusted log hourly wages for all groups in 1990 dollars. In 1960 and 1970, immigrant men and women both earned slightly more than their native-born counterparts. Immigrant men's wages

[7] We ran a log hourly wage regression for the native-born on age, age squared, education dummies, race and ethnicity dummies, marital status, and metropolitan residence. We then predicted a log hourly wage for everyone. We then divided that into quintiles.

fell steadily relative to the native-born between 1960 and 1990. On average, immigrant men in 1960 earned 5.3% more per hour than the native born. By 1990, they earned 9.5% less on average. For immigrant women, the story is a bit different. Immigrant women earn about 2-3% more than the natives until 1990, when their wages are approximately the same.

Tables 1a and 1b also show the average hourly wages for recent immigrants in each year. Recent immigrants are an interesting group to study, since they are likely to be less heterogeneous than immigrants overall. They have had exactly the same amount of time to "assimilate" in each of the four periods. Similarly, return migration is unlikely to have changed the composition of this immigrant cohort. Average hourly wages for recent immigrants have always been substantially less than for the native-born. Recent immigrant men in 1960 and 1970 earned from 20-15% less per hour than native-born men. By 1990, this gap had climbed to 34% less. The pattern for recent immigrant women is similar: in 1960, recent immigrant women earned about 14% less than native-born women. In 1990, they earned 19% less. (See Borjas (1990) for similar figures).

Note that for both men and women, the recent immigrant- native-born wage differential is somewhat smaller in 1970 than in 1960. Recall that the changes in the immigration law that were passed in 1965 (and extended to the Western Hemisphere in 1968) almost certainly did not become fully effective until late in the 1960s. Moreover, to the extent that can be ascertained, income equality was closer to an all time "high" in 1969 than at any other time in the U.S. postwar history. [8] For these reasons, we will conduct much of our analysis later in the paper using recent immigrants in 1970.

There are two points to take away from this section. First, although the trend is the same, immigrant women and men differ in their relative wages in the United States labor market. The gaps between the wages of immigrant and native-born women are substantially smaller than those between immigrant and native-born men. Little attention has been focused on the performance of immigrant women in the United States labor market. In what follows, we will continue to show results for both women and men.

Secondly, our results corroborate previous research on the decline in wage earning attributes ("skills") of immigrants. In this section we focused on

[8]See for example, the series on gini indices in the P-60 Current Population Reports of the U.S. Bureau of the Census.

differences in mean log hourly wage and we find that immigrants, especially recent immigrants, are earning substantially less than their native-born counterparts and this gap has increased over the past 30 years. In what follows we analyze the role that changes in the distribution of earnings between 1960 and 1990 play in this decline in relative mean wages.

3 Immigrant and Native Wages Distributions

3.1 General Considerations

Until this point, our focus has been on simple summaries of the data. Such approaches have many advantages, not least of which is a considerable amount of "data reduction": summarizing vast quantities of information compactly. On the other hand, some aspects of wage structure changes can be obscured by a focus on simple summary statistics.

In this section, we again compare immigrant wages to native wages from our four Census samples except that we now focus on the *entire* distribution of wages. It might be interesting to consider the effect of supply and demand, minimum wages and their possible employment effects, and de-unionization, on the wage outcomes of immigrants and natives as in DiNardo, Fortin, and Lemieux (1996). Similarly, it would be interesting to incorporate the possible interaction between changes in the level of immigration and the wage structure of natives as in Card (1990), Butcher and Card (1991), Borjas, Freeman, and Katz (1992) and Borjas, Freeman, and Katz (1996). *Inter alia*, limitations of the necessary time-consistent information from the four samples lead us instead to limit our focus to the distribution of observed wages.

3.2 Methodological Concerns

The non-parametric density estimates we consider in this paper use the kernel density estimator introduced by Rosenblatt (1956) and Parzen (1962). The kernel density estimate $\widehat{f}_h$ of a univariate density f based on a random sample

$W_1, \ldots, W_n$ of size n, with weights $\theta_1, \ldots, \theta_n$ ($\sum_i \theta_i = 1$), is

$$\hat{f}_h(w) = \sum_{i=1}^{n} \frac{\theta_i}{h} K\left(\frac{w - W_i}{h}\right), \tag{1}$$

where h is the bandwidth and $K(\cdot)$ is the kernel function.[9]

A potentially important issue in kernel density estimation is choice of bandwidth. Put simply, larger bandwidths result in more bias and less variance (over-smoothing), while smaller bandwidths result in less bias and more variance (under-smoothing). Although there are a number of different methods for *automatically* choosing the bandwidth ranging from cross-validation[10] to "plug-in" methods[11], there is no consensus on what is "optimal." Instead we apply the simple dictum: since it is generally easier to smooth with the eye than "unsmoothe" with the eye, we choose bandwidths that err on the side of being "too small." Furthermore, when we consider more than one density estimate at a time, we apply the same bandwidth to each. The estimates in this section use bandwidths from 0.0477 to 0.0988, using smaller bandwidths for larger samples. As the general shape of the densities remain the same for a for a fairly large range of bandwidths, the issue of bandwidth seems to have little practical importance in our exercise.

Less important is the issue of kernel choice, and for all our estimates we use a Gaussian kernel.

3.3 Simple Density Estimates: All Immigrants and Natives

Figure 1 presents density estimates for the wage distributions of native-born men and women from 1960 through 1990. All wage observations have been converted to 1990 dollars. [12] These estimates highlight the dramatic changes in wage structure over these decades. The change is greatest for men, where

[9] In this section, the issue of weights is of little practical consequence since only the 1990 Census includes population weights. It will be become a concern later when we consider "counterfactual" density estimates.

[10] See Silverman (1986) for a discussion

[11] See Sheather and Jones (1991) for one example.

[12] We used the implicit price deflator for GNP total personal consumption on expenditures.

the increasing density in the lower tail is striking. The distributional changes for women are subtly different : for men, the "peak" of the distribution is lower in 1990 than in any other year, but there is an increase in density in the upper tail. Also, where the distribution for men shows a steady increase in the variance, the distributions for women from 1960 through 1980 are more similar. This is in part due to the effect of the minimum wage on the distribution of wages for women in 1980 (DiNardo, Fortin, and Lemieux 1996).

In figures 2a and 2b we compare the density of all immigrants to natives for each year of our data, for men and women, respectively. As in the previous figure, the log hourly wages have been converted to 1990 dollars.

Considering native men first, it is clear that most measures of central tendency show that wages rise between 1960 and 1970 and fall thereafter, as is consistent with a large body of research. The distribution of immigrant wages shows some interesting patterns over time. In both 1960 and 1970, the immigrant and native born wage distributions are very similar. After 1970 however, the two distributions begin to diverge. Much of the difference between the two distributions is slightly below and slightly above the median. The difference between the wage distribution of the two groups, in terms of measures such as mean wages, is less driven by differences at the extremes of the distribution, but rather by differences between "middle-class" immigrants and natives (using the term middle-class in the sense of "middle portion of the distribution.")

Again for women, in 1960 and 1970 the densities for immigrants are very similar to their native counterparts. However, the relative position of the two densities appears to *converge in 1980.* Despite the measurement error in our constructed average hourly earnings measures, the large impact of the minimum wage (which was at a peak in real terms in 1979) is evident in the spike for both immigrants and natives in 1980. This is consistent with the evidence presented in DiNardo, Fortin, and Lemieux (1996). In 1990, the distributions appear to begin to pull apart slightly, with immigrants relatively over–represented slightly to the left of mode and natives relatively over–represented slightly to the right of mode.

3.4 Simple Density Estimates: Recent Immigrants and Natives

We have thus far restricted our attention to a comparison between all immigrants and natives. In this subsection we consider the difference between recent immigrants and natives. In figures 3a and 3b we therefore repeat the exercise we performed in the previous two figures, except this time we overlay the distribution of wages of recent immigrants and natives. As was mentioned earlier, a comparison of immigrants who have recently arrived (within the past five years in this analysis) is a direct way to investigate changes in immigrant skills across cohorts. The wages of immigrants overall are a influenced by a myriad of factors, including the length of time immigrants have been in the country, and the size of the arriving cohort in each previous time period, attrition of a cohort due to death and decisions to emigrate.

Men

The top two panels of figure 3a display our estimates for our male 1960 and 1970 census samples. In contrast to figure 2a where we considered all immigrant and native-born men, the leftward "shift" of the distribution of recent immigrant wages is quite pronounced. *Inter alia*, this is attributable to the fact that recent immigrants are much younger than natives taken as a whole, and are the youngest among all immigrants (see Table 1a.) Likewise, the bottom two panels of figure 3a display our estimates for our male 1980 and 1990 census samples.

One interesting way to describe the difference between the two male wage densities is to compare the difference in the modal recent immigrant and native wages. In 1960 the modal recent immigrant earns about 29 percent less than the modal native. By 1970, this difference falls to to 10 percent although, this computation slightly overstates the similarity as the leftmost mode is 30 percent lower than the modal native. In terms of magnitude, however, the big changes occur after 1970. Between 1970 and 1980 the difference between the modal immigrant and his native counterpart *rises* 50 percentage points to 62 percent and then rises another 19 points. By 1990, the modal recent immigrant is now making *81 percent less than his native counterpart.*

Women

In several respects the changes in the relative distributions for women's wages are similar, but there is one noteworthy difference. The difference between the modal recent immigrant and native for the years 1960, 1970, 1980, and 1990 are 0.20, 0.07, 0.23, and 0.49 respectively. The difference in the male and female immigrant/recent immigrant wage gap in 1980 is particularly noteworthy. Consistent with the evidence for women discussed in section 3.3, the minimum wage seems to play the largest role. This corroborates evidence from the CPS, where wages at or very near the minimum wage represent the *modal* wage for all women. Therefore, it is not surprising that the mode of wages for recent immigrants and native women show little difference in our 1980 Census data.

Summary

A powerful visual impression one receives from the densities in Figures 2a-3b is that over time, men's and women's distributions changed in very different ways, in part because of the minimum wage. As for men, the gap between recent immigrants and natives is spreading at the same time that the distribution is is widening.

Although the big changes in immigration law were in effect for recent immigrants in 1980 and 1990, the summary statistics in Tables 1 and 2 do not tell a simple story about a uniform decline among immigrant women and men, in the characteristics that are rewarded in the United States labor market. Rather, the visual evidence from the simple density functions is consistent with the important impact of the minimum wage, which is more binding for women than men. This suggests that trends in the immigrant/native wage gap differ in ways directly attributable to changes in wage structure instead of changes in "skills." Any simple story about the changes in immigrant/native mean wages that relies solely on changes in admission criteria has difficulty encompassing the differences in results for men and women presented here.

4 Counterfactual Density Estimates

LaLonde and Topel (1992) discuss the role of the changing wage structure in comparisons between immigrants and the native-born. Although their

analysis is substantially different in form, they make a similar point. If immigrants have always been concentrated in the lower part of the earnings distribution, and the distribution of earnings becomes more dispersed, then immigrants will look worse relative to natives, even without a change in immigrants' characteristics. Although immigrants overall are not concentrated in the very bottom of the distribution, their emphasis on the importance of wage structure changes is consistent with our findings in the previous section. Additionally, since recent immigrants are, and have always been (during the time periods covered in this paper), much younger and in ethnic/racial groups that are disadvantaged relative to natives or immigrants overall, they are more likely to be concentrated at the lower end of the earnings distribution in each year of our data (as seen in figures 3a and 3b). For this group, the changes in the wage structure are likely to be particularly important.

In this section we take the analysis one step further: we can investigate how the distribution of earnings for recent immigrants in the past would look if they faced the wage structure of the 1990s. Recall that the United States changed its immigration law in 1965, and the changes in these laws took effect in the late 1960s. This change in criteria is widely cited as a primary cause of the "deterioration" in the characteristics of immigrants. While recent immigrants in 1960 were clearly admitted under the old criteria, since the law was not effective immediately, most of the recent immigrants in 1970 were admitted under the old criteria as well. As Tables 1a and 1b show, their mean wages are higher than in the other years. In addition, the data in the 1970 Census, particularly on the operational definition of Hispanic, is more comparable with the 1980 and 1990 Census than is the 1960 Census. For these reasons, we will perform our comparisons in this section between recent immigrants in 1970 and recent immigrants in 1990. However, we have completed all comparisons using the 1960 Census data as well, and these results are very similar. (See Appendix figure 1 for an example).

4.1 Methods

In this section we consider the following thought experiment: What would the distribution of 1970 recent immigrant wages look like if they were treated (on the basis of observable X's only) as 1990 recent immigrants. To do so we adapt the method discussed in DiNardo, Fortin, and Lemieux (1996). In what follows we briefly review the approach in this setting, and describe the

necessary modifications. Additional detail can be found in DiNardo, Fortin, and Lemieux (1996).

One way to begin would be to postulate two different wage equations, one for 1970 recent immigrants and another for 1990 recent immigrants:

$$y^{70} = X^{70}\beta^{70} + \epsilon^{70} \tag{2}$$

$$y^{90} = X^{90}\beta^{90} + \epsilon^{90} \tag{3}$$

Given this formulation, the simplest way to proceed is to estimate separate regressions for the 1970 sample and the 1990 sample to get estimates of β^{70} and β^{90} respectively. The well-known Oaxaca (1973)/ Blinder (1973) decomposition involves computation of the following:

$$\widehat{\overline{Y}_{70}^{90}} = \overline{X}_{70}\widehat{\beta^{90}} \tag{4}$$

$$\widehat{\overline{Y}_{90}^{70}} = \overline{X}_{90}\widehat{\beta^{70}} \tag{5}$$

where $\widehat{\beta^{90}}$ and $\widehat{\beta^{70}}$ are the OLS estimates from equation (3) and (2) respectively, and $\overline{X}_{90}$ and $\overline{X}_{70}$ are the means of X variables in 1990 and 1970 for recent immigrants. $\widehat{\overline{Y}_{70}^{90}}$ is the mean salary of 1970 recent immigrant workers had they been paid with the wage function of recent immigrants in 1990. $\widehat{\overline{Y}_{90}^{70}}$ is the mean salary of 1990 recent immigrants if they had been paid according to the wage function of 1970 recent immigrants.

In practical applications, the Oaxaca/Blinder approach is generally restricted to a comparison of means. In fact, when the distributions one is comparing are unimodal, symmetric and have similar variances, the Oaxaca approach comes quite close to being a "sufficient statistic" for the effect of changes in the structure of wages.

As has been well documented, however, changes in the shape of the distribution of wages over the time period we consider have been quite dramatic. With a simple modification of non-parametric density estimation, however, it is easy to analyze such changes. One goal is to estimate the density that would have prevailed in 1990, had the distribution of wage determining characteristics been as it was among 1970 recent immigrants.

Note that the definition of conditional probability yields the following

representation of the overall distribution of wages.

$$g(w) = \int f(w|x)h(x)dx$$

When the conditional expectation is linear in the x's $f(w|x)$ is closely related to the regression function.

Since the counterfactual distribution we wish to generate involves combining distributions with different "dates" it will be helpful to establish notation for the observed distributions that incorporate this dating.

Define the observed density of wages in 1990 given by:

$$g(w|t = 90) = \int f^{90}(w|x)h(x|t = 90)dx \tag{6}$$

where $f^{90}(w|x) \equiv f(w|x, t = 90)$. As before, in the special case where the conditional expectation and the linear projection are the same, $f^{90}(w|x)$ is closely related to the wage equation for recent immigrants in 1990.

Likewise, the observed density of wages for recent immigrants in 1970, can be written as:

$$g(w|t = 70) = \int f^{70}(w|x)h(x|t = 70)dx \tag{7}$$

where $f^{70}(w|x) \equiv f(w|x, t = 70)$

We are interested in the distribution of wages if 1970 recent immigrants were paid under the wage structure prevailing for recent immigrants in 1990, or more formally:

$$g_{70}^{90}(w) = \int f^{90}(w|x)h(x|t = 70)dx \tag{8}$$

As is turns out, estimation of the above density can be made simple by noting that Bayes Law implies:

$$\begin{aligned} h(x) &= \frac{h(x|t = 70)Pr(t = 70)}{Pr(t = 70|x)} \\ h(x) &= \frac{h(x|t = 90)Pr(t = 90)}{Pr(t = 90|x)} \end{aligned} \tag{9}$$

Since we are in effect considering a sample of wage realizations from a

"population" of 1970 and 1990 recent immigrants, we can without loss of generality set $Pr(t = 70) = Pr(t = 90)$. That is, the cumulative probability of being a 1970 immigrant from this "population" is the same as the cumulative probability of being a 1990 immigrant, and these two terms can be ignored in the analysis that follows.

Writing the above equations in terms of $h(x|t = 70)$ and $h(x|t = 90)$ reveals that equation (6) (i.e. the actual density of 1990 recent immigrant wages) is exactly like the desired counterfactual distribution described in equation (8) except for the term: $h(x|t = 70)$ which is instead replaced by $h(x|t = 90)$ in the actual 1990 distribution of wages.

This "problem" can be fixed by merely multiplying the "weight" in equation (6) by:

$$\frac{h(x|t = 70)}{h(x|t = 90)}$$

substituting equation (9) into equation (8) yields the following:

$$g_{70}^{90}(w) = \int \theta f^{90}(w|x)h(x|t = 90)dx \tag{10}$$

where

$$\theta = \frac{Pr(t = 70|x)}{Pr(t = 90|x)}$$

But notice that equation (10) is identical to equation (6) except for the "weight" θ. In essence, we have reduced the problem of estimating the desired counterfactual density to calculating the appropriate weight and applying this weight to the observed distribution of recent immigrant wages in 1990. The term $Pr(t = 70|x)$ can be estimated non–parametrically, by dividing up the sample by the characteristics x and calculating the proportion of individuals in each cell, or by a discrete choice model like the logit, where the x's are entered in a reasonably flexible way. We estimate θ by a simple logit of the form:

$$\Pr(t = 70) = \Lambda(f(\text{age, age squared, schooling, etc.,})) \tag{11}$$

with our combined sample of 1970 and 1990 immigrants. The choice of

discrete model is of little consequence, but the logit has a practical advantage relative to the probit that the sum of predicted values equals the sum of empirically observed values. The resulting weights are then normalized to sum to one.

If the data is already weighted by some factor ψ as in 1990, this procedure is identical except that the weight is merely $\psi \cdot \theta$, where the weight ψ is already normalized to sum to one.

Note that considering the entire distribution of wages does not preclude analyzing more conventional summaries of distributions. For instance, the weights we calculate can be used to compute *any* statistic of interest. Even the simple Oaxaca/Blinder counterfactual can be computed by taking the sample of 1990 immigrants and applying our counterfactual weights.[13]

4.2 Results

Our initial estimation equation included 3 schooling categories, 3 race and ethnicity categories, marital status, metropolitan residence, residence in a high immigrant state, seven industry dummies, five place of birth categories, and interactions between high immigrant state and some of the other variables, and between metropolitan area and some of the other variables. By including these variables, we are re-weighting the 1990 recent immigrants to have the same characteristics, *as measured by these included variables*, as the 1970 recent immigrants. We can also investigate the importance of these characteristics to our counterfactual density estimates by including or excluding them. The most discussed changes in immigrants' characteristics are the changes in education, race and ethnicity, and region of birth. Our first set of counterfactual density estimates includes the full set of variables listed above. We then show the importance of race and ethnicity and region of origin to the changes in recent immigrant wages between 1970 and 1990.

Figure 4 presents our counterfactual density estimates. Each picture overlays two densities: one, the actual estimated density for recent immigrants in 1990 and two, our estimate of the distribution of wages that would have obtained for 1970 immigrants if they had faced 1990 skill prices constructed as

[13]The Oaxaca/Blinder counterfactual computed the usual way takes a value almost identical to the Oaxaca/Blinder counterfactual using our weights.

described in the previous subsection. The left panel is for recent immigrant women, and the right panel for men.

The first set of pictures shows the results when we include all of the characteristics described above. For both women and men the "actual" density is somewhat left-shifted compared to the "counterfactual" density, however there is substantial overlap. The overlap is greater for women than for men. [14] This overlap suggests that even if 1990 recent immigrants had characteristics identical to those who arrived in 1970, we would still see a decline in the earnings of recent immigrants. The most notable difference between these densities, however, is in the dispersion. If recent immigrant men from the 1970s had faced the skill prices prevailing in 1990, the distribution of their wages would have been much more dispersed than it actually is among the 1990 recent immigrant men.

The results for women in the top panel show substantial overlap between the actual and counterfactual densities. The counterfactual density shows somewhat higher concentrations to right of the mode than the actual density. Again, this suggests that if recent immigrant women in 1990 had instead had the characteristics of recent immigrant women in 1970, we would nonetheless observe virtually the identical wage distribution.

Our next idea explores the question: which characteristics most distinguish (in terms of their "effect" on wages) the recent immigrants of the 1970s and the 1990s? The immigrant characteristics that receive the most attention in popular discussion are education, race and ethnicity, and region of origin. We will use our techniques to investigate the sensitivity of our counterfactual estimates to the exclusion of race/ethnicity and region of origin variables. We approach this question agnostic on the precise mechanism by which demographic variables have measurable impacts on wages. Current practice is to assume that wages largely reflect marginal productivity and that variables correlated with wages do so because they are correlated with productivity(DiNardo and Pischke 1997). Absent direct evidence on marginal productivity, we think a satisfactory resolution of this debate is unlikely to occur soon.

We first investigate the sensitivity of these estimates to race and ethnicity

[14]The minor differences between the "actual" densities here and in Figures 3a and 3b is due to differences in smoothing: the bandwidth for the actual densities in Figure 4 was chosen to be the same as those for the counterfactuals in Figure 4.

changes. The second panel in Figure 4 shows the result of excluding race and ethnicity from the logit used to predict whether one is a 1970 or 1990 immigrant. Thus, education, state of residence, industry, and region of birth are the important predictors of whether one is a 1970 or 1990 recent immigrant. Note that the more variables we include in the logit step, the more precisely we are re-weighting the 1990 recent immigrants to "look like" the 1970 recent immigrants. By excluding race and ethnicity, we are asking how the wages of individuals who look like the 1970 recent immigrants *except for their racial and ethnic composition* compare to 1990 recent immigrants, when faced with the 1990 wage structure.

The results in the second panel show the tremendous importance of race and ethnicity in determining the changes in wages between 1970 and 1990. Once these characteristics are no longer included, the differences in the actual and counterfactual densities diminish. This is especially the case for women.

In the above exercise region of origin is included in the prediction equation. Since region of birth and race and ethnicity are closely related, these might be quite similar exercises. The last set of results further excludes these five region of birth categories from the prediction equation. We see that the differences between the actual and counterfactual distributions shrink even more. The importance of these covariates is highlighted in Figure 5 which displays the difference-in-density estimates which correspond to Figure 4.

This exercise demonstrates that there is tremendous overlap between the wages of recent immigrants in 1990 and the wages recent immigrants in 1970 would have received had they faced the 1990 prices for their skills or characteristics. To the extent that there are differences between the actual and counterfactual densities, they appear to be related to the race and ethnicity changes in recent immigrants between 1970 and 1990. When we discriminate between recent immigrants based neither on race/ethnicity, nor region of origin, all differences disappear. The comparison between the three sets of results in Figure 4 demonstrate that changes in race and ethnicity are strongly correlated with changes in wages.

5 Changes in the Wage Structure or Changes in Immigrant Skills?

In this section we decompose the change in the gap in wages between recent immigrants and the native born into a portion due to changes in the wage structure and a portion due to changes in relative skills. Adapting the weights constructed in the previous section allows us to perform this decomposition at different points in the wage distribution. In this section, we use the most complete set of observable "skills" - including race/ethnicity and region of birth for the recent immigrants - to characterize the changes in the attributes that affect wages.

One important point that we make in this section is that any Blinder/Oaxaca decomposition of differentials into "wage structure" versus "skills" is not unique, for reasons stressed in the original literature on the subject.

Simply put, when the wage structure has been changing a great deal it *must* play an important role in any "gap." This is well recognized in the Blinder/Oaxaca discrimination literature. The extent of discrimination against women, for example, will differ depending on whether the women's coefficients are used as in:

$$(\overline{X_m} - \overline{X_w})\beta_w$$

or men's coefficients as in:

$$(\overline{X_m} - \overline{X_w})\beta_m$$

where in the usual notation $\overline{X_w}$ represents the average value of the wage determining characteristics for women, β_w represents the coefficients from a wage regression using only women, and where the remaining terms are defined analogously for men.

A similar point applies with equal to force to immigrant/native wage differential decompositions as has been stressed by LaLonde and Topel (1992). It will be useful to establish some notation. The change in the wage gap (ΔG) can be written as follows:

$$(N_{90}^{90} - I_{90}^{90}) - (N_{70}^{70} - I_{70}^{70}) = \Delta G \tag{12}$$

where N_x^y refers to the mean (or other moment) of native-born wages with skills as in year x and wage structure as in year y. Using the weights from

a logit as before, the appropriate moment is straightforward to compute.[15] The top panels of tables 3a and 3b report all the possible permutations of N_x^y and I_x^y for different moments of the distribution. In what follows, we will use the elements of the top panel to demonstrate how evaluations of the importance of wage structure changes are sensitive to essentially arbitrary assumption about the price of skill.

Note that Equation 12 can be rewritten as follows:

$$\left[(N_{90}^{90} - N_{70}^{90}) - (I_{90}^{90} - I_{70}^{90})\right] - \left[(N_{70}^{70} - N_{70}^{90}) - (I_{70}^{70} - I_{70}^{90})\right] = \Delta G \qquad (13)$$

where we have merely added and subtracted the term $N_{70}^{90} - I_{70}^{90}$. The first term in brackets *may* be described as the change in the wage gap attributable to the change in the native-born/immigrant gap in "skills." More precisely, it is the change in the gap that would have occurred had *1990* skill prices prevailed over the entire period.

The second term in brackets *may* be described as the change in the gap attributable to changes in the wage structure. This part of the gap is merely the change in native-born/immigrant wage gap that would have occurred if neither group had experienced a change in their X variables. Put differently, it is the change in the value of skill gap between immigrants and natives that existed in *1970.*

Consider the implications of this decomposition for men between 1970 and 1990. In 1990, the gap between recent immigrants and the native born (at the mean) was 34 percent. In 1970, this gap was only 15 percent which implies that the gap increased 19 percentage points over the period 1970 to 1990.

In the first column of the lower panel of Table 3a, we present the amount of the gap that is attributable to "skills" as in the decomposition in Equation 12. At the mean (the last row of the table) the change in the wage gap over the period 1970-1990 that is attributable to the relatively lower "skills" of recent immigrants is 21 percent, which explains more than 100 percent of the difference. Like most other analysts this evidence might lead us to conclude

[15]This requires running a logit to predict who is a 1970 native, in order to construct weights analogous to those described above for the native born. We use the same explanatory variables, except for region of birth.

that "changes in the U.S. wage structure were not sufficiently large to account for a sizable part of the declining relative wage of immigrants" (Borjas 1994).

It is evident, however, that this is not the only way to do such a decomposition. Suppose instead that we add and subtract the term $N_{90}^{70} - I_{90}^{70}$. Doing so yields:

$$\left[(N_{90}^{90} - N_{90}^{70}) - (I_{90}^{90} - I_{90}^{70})\right] - \left[(N_{70}^{70} - N_{90}^{70}) - (I_{70}^{70} - I_{90}^{70})\right] = \Delta G \qquad (14)$$

In this case, the first term in brackets could be labeled the part of ΔG attributable to changes in the wage structure: that is, it measures how the gap would have changed if the skills of immigrants and natives had always been at their 1990 level. The second term could be labeled the part of the gap attributable to the change in skills, although in contrast to the previous decomposition we evaluate the change in skills in 1970s prices instead of 1990s prices. It might be helpful to view this decomposition as identical to the previous case, except that we have reversed the roles of 1990 and 1970.

This "role reversal" has quite important implications for evaluating the importance of the wage structure in the increasing wage gap between recent immigrants and the native born. Consider the bottom panel of Table 3a. The second column shows that for men, at the mean, had the wage structure remained as it was in 1970 the increase in the wage gap would have been only 9.2 percent. As the actual wage gap increased 18.8 percent, more than 50 percent of the increase is now attributable to the wage structure.

Depending on where in the wage distribution one looks, the effects can be even more dramatic. Consider the 75 percentile of the male wage distribution. Over the 1970 to 1990 period, the wage gap increased from 10.6 to 43 percent. Evaluating the increase in the skill gap at 1970 skill prices yields only 10.4 percent. Using this particular decomposition, therefore, wage structure changes explain 68 percent of the increase in the wage gap.

We should stress, however, that there is nothing unique about our sample that leads to the conclusion that wage structure changes are important. Using the decomposition given by Equation 12, virtually all of the increase in the recent immigrant native born wage differential can be attributed to "skill." If we evaluate the change in skill in 1990 prices, the gap at the respective 75th percentiles would have been 31.8 percent. As the actual gap was 32.4 percent, this decomposition suggests the changing skill gap explains almost 98 percent

of the increase in the native-born/recent immigrant wage differential!

The sensitivity of the decomposition to the choice of base period is also evident in the results for women, displayed in table 3b. At the 50th and 75th percentile, the change in the gap due to skills is twice as large when evaluated at 1990 prices (Equation 12) than when evaluated at 1970 prices (Equation 13). Interestingly, at the 25th percentile the proportion attributable to skill is larger when evaluated at 1970 prices (13.8) than when evaluated at 1990 prices (11.3).

What this exercise has demonstrated is that the choice of "prices" used to evaluate the changes in skill is tantamount to deciding which is more important – changes in skills or changes in wage structure. Most previous work that has looked at this question has implicitly evaluated the skills at 1990 prices. As the gender–specific distributions are at their widest in 1990, these previous evaluations have necessarily attributed little to the tremendous changes in the wage structure. While no choice of price is ontologically prior, one intuitive way to appreciate the effect of the wage structure is to ask the question – would the gap have increased as much if the wage structure had remained unchanged from 1970. By this metric, it is clear that much of the increase in the native–born/recent immigrant wage gap would not have occurred.

6 Summary and Conclusion

We take away 5 conclusions from our analysis:

1. Patterns of comparison between the wages of immigrants and the native–born are not the same for men and for women.

2. Moreover, these differences in immigrant/native–born comparisons among men and women are clearly consequences of different evolutions in the wage structure. For example, the minimum wage exerts a greater effect on the wage distribution for women than for men.

3. If recent immigrants in 1970 had faced the 1990 wage structure, the distribution of their wages would have been markedly similar to the wage distribution that actually existed for recent immigrants in 1990.

4. If the wage structure or "skill" prices had remained the same between 1970 and 1990, the recent immigrant/native-born wage differentials (at any point in the wage distribution) would have been much lower than has been observed.

5. Absent an ontologically defensible time-invariant measure of the "value of skill", dramatic changes in the wage structure have important consequences for group comparisons of wages.

This last observation is well-known in the discrimination literature: Oaxaca/ Blinder decompositions of the white-black wage gap into a part attributable to differences in the X's (regressors) and a part attributable to changes in the coefficients depend on whose coefficients are used to evaluate the difference in the regressors. If for example, the coefficients for whites are used, the role of the regressors ("skill") is much larger than if the black coefficients are used. A meaningful discussion of the breakdown of the change in the immigrant/native wage gap, or any other wage gap, must include an assertion of which prices are "correct."

In sum, our analysis and results suggest that the dramatic changes in the distribution of wages between 1960 and 1990 have indeed had an important effect on wage comparisons between immigrants and the native–born. This suggests, among other things, that the emphasis on post-1965 changes in United States immigration policy in explaining the relative position of immigrants and natives may be misplaced.

References

BAKER, M., AND D. BENJAMIN (1997): "The Role of the Family in Immigrants' Labor Market Activity: An Evaluation of Alternative Explanations," *American Economic Review*.

BLAU, F. (1992): "The Fertility of Immigrant Women: Evidence from High-Fertility Source Countries," in *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, ed. by G. J. Borjas. and R. B. Freeman, A National Bureau of Economic Research project report, chap. 3, pp. 93–133. University of Chicago Press, Chicago.

BLINDER, A. S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 8, 436–455.

BORJAS, G. (1987): "Self-Selection and The Earnings of Immigrants," *American Economic Review*, 77(4), 531–553.

——— (1990): *Friends or Strangers: The Impact of Immigrants on the U.S. Economy*. BasicBooks.

——— (1991): "National Origin and the Skills of Immigrants in the Postwar Period," NBER Working Paper 3675, National Bureau of Economic Research, Cambridge, MA.

——— (1994): "The Economics of Immigration," *Journal of Economic Literature*, 32(4), 1667–1717.

——— (1995): "The Internationalization of the U.S. Labor Market and the Wage Structure," *Economic Policy Review*, 1, 3–8.

BORJAS, G., R. FREEMAN, AND L. KATZ (1992): "On the Labor Market Effects of Immigration and Trade," in *Immigration and the Workforce*, ed. by G. Borjas, and R. Freeman. University of Chicago Press, Chicago.

——— (1996): "Searching for the Effect of Immigration on the Labor Market," NBER Working Paper 5454, National Bureau of Economic Research.

BUTCHER, K. F., AND D. CARD (1991): "Immigration and Wages: Evidence from the 1980's," *American Economic Review*, 81(2), 292–296.

CARD, D. (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43(2), 245–257.

CHISWICK, B. R. (1978): "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, 86(5), 897–921.

DINARDO, J., N. M. FORTIN, AND T. LEMIEUX (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, forthcoming.

DINARDO, J., AND J. PISCHKE (1997): "The Returns to Computer Use Revisited: Have Pencils Cha nged the Wage Structure Too?," *Quarterly Journal of Economics*.

DULEEP, H. O., AND S. SANDERS (1993): "The Decision to Work by Married Immigrant Women," *Industrial and Labor Relations Review*, 46, 677–690.

LALONDE, R. J., AND R. H. TOPEL (1992): "The Assimilation of Immigrants in the U.S. Labor Market," in *Immigration and the Work Force: economic consequences for the United States and source areas*, ed. by G. J. Borjas, and R. B. Freeman, A National Bureau of Economic Research project report, chap. 3, pp. 67–92. University of Chicago Press, Chicago.

OAXACA, R. (1973): "Male–Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693–709.

PARZEN, E. (1962): "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, 33, 1065–1076.

ROSENBLATT, M. (1956): "Remarks on Some Non–parametric Estimates of a Density Function," *The Annals of Mathematical Statistics*, 27, 832–837.

SHEATHER, S., AND M. JONES (1991): "A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation," *Journal of the Royal Statistical Society, B*, 53, 683–690.

SILVERMAN, B. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.

YUENGERT, A. M. (1994): "Immigrant Earnings, Relative to What? The Importance of Earnings Function Specification and Comparison Points." *Journal of Applied Econometrics*, 9, 71–90.

Table 1a: Descriptive Statistics for Men : Native-born, Immigrants, and Recent Immigrants, by Year
(Standard Errors)

	1960			1970		
	Native-born	Immigrants	Recent Immigrants	Native-born	Immigrants	Recent Immigrants
Age	37.913	47.752	33.059	37.189	41.746	32.654
	(0.0196)	(0.0765)	(0.1797)	(0.0190)	(0.0885)	(0.1494)
H.S. Drop out	0.561	0.678	0.556	0.436	0.502	0.474
	(0.0006)	(0.0027)	(0.0080)	(0.0007)	(0.0029)	(0.0066)
H.S. Degree	0.248	0.148	0.157	0.307	0.212	0.178
	(0.0006)	(0.0021)	(0.0059)	(0.0006)	(0.0024)	(0.0050)
Some College	0.098	0.081	0.127	0.135	0.126	0.124
	(0.0004)	(0.0016)	(0.0054)	(0.0005)	(0.0019)	(0.0043)
College Graduate	0.091	0.093	0.160	0.122	0.160	0.224
	(0.0004)	(0.0017)	(0.0059)	(0.0004)	(0.0021)	(0.0055)
Black	0.099	0.017	0.034	0.100	0.036	0.067
	(0.0004)	(0.0008)	(0.0029)	(0.0004)	(0.0011)	(0.0033)
Asian[1]	--	--	--	--	--	--
Other Race[1]	0.007	0.050	0.074	0.009	0.081	0.158
	(0.0001)	(0.0013)	(0.0029)	(0.0001)	(0.0016)	(0.0048)
Hispanic[2]	0.015	0.078	0.163	0.032	0.218	0.342
	(0.0002)	(0.0015)	(0.0060)	(0.0002)	(0.0024)	(0.0062)
Married	0.715	0.780	0.596	0.678	0.744	0.640
	(0.0006)	(0.0024)	(0.0079)	(0.0006)	(0.0026)	(0.0063)
Metropolitan Area	0.569	0.808	0.787	0.622	0.824	0.853
	(0.0007)	(0.0023)	(0.0066)	(0.0006)	(0.0022)	(0.0046)
Log Hrly Wage[3] 1990$'s	2.218	2.271	2.017	2.447	2.483	2.293
	(0.0011)	(0.0042)	(0.0136)	(0.0011)	(0.0046)	(0.0111)
Fraction of the Population	--	0.058	0.007	--	0.050	0.010
		(0.0003)	(0.0001)		(0.0003)	(0.0001)
Sample Size	484963	29982	3816	560088	29209	5806

	1980			1990[5]		
	Native-born	Immigrants	Recent Immigrants	Native-born	Immigrants	Recent Immigrants
Age	36.268	37.030	29.977	37.100	36.405	30.196
	(0.0151)	(0.0536)	(0.0836)	(0.0143)	(0.0420)	(0.0728)
H.S. Drop out	0.304	0.400	0.419	0.224	0.407	0.444
	(0.0005)	(0.0019)	(0.0038)	(0.0004)	(0.0016)	(0.0034)
H.S. Degree	0.345	0.229	0.199	0.290	0.174	0.170
	(0.0005)	(0.0017)	(0.0031)	(0.0005)	(0.0013)	(0.0025)
Some College	0.179	0.167	0.178	0.276	0.200	0.172
	(0.0004)	(0.0015)	(0.0030)	(0.0005)	(0.0013)	(0.0025)
College Graduate	0.172	0.204	0.205	0.210	0.219	0.214
	(0.0004)	(0.0016)	(0.0031)	(0.0004)	(0.0014)	(0.0028)
Black	0.108	0.065	0.071	0.115	0.077	0.075
	(0.0003)	(0.0010)	(0.0020)	(0.0003)	(0.0009)	(0.0018)
Asian	0.006	0.155	0.260	0.006	0.231	0.272
	(0.00001)	(0.0014)	(0.0034)	(0.0001)	(0.0014)	(0.0030)
Other Race	0.020	0.169	0.228	0.028	0.226	0.257
	(0.0005)	(0.0015)	(0.0032)	(0.0002)	(0.0014)	(0.0030)
Hispanic	0.036	0.351	0.368	0.045	0.445	0.478
	(0.0002)	(0.0019)	(0.0037)	(0.0002)	(0.0017)	(0.0034)
Married	0.606	0.658	0.517	0.569	0.613	0.474
	(0.0005)	(0.0019)	(0.0039)	(0.0005)	(0.0016)	(0.0034)
Metropolitan Area	0.800	0.945	0.945	0.822	0.957	0.956
	(0.0004)	(0.0009)	(0.0018)	(0.0004)	(0.0007)	(0.0014)
Log Hrly Wage[3] 1990 $'s	2.386	2.364	2.127	2.337	2.242	1.995
	(0.0008)	(0.0032)	(0.0068)	(0.0009)	(0.0028)	(0.0059)
Fraction of the Population	--	0.066	0.017	--	0.091	0.022
		(0.0003)	(0.0001)		(0.0003)	(0.0001)
Sample Size[4]	900991	63221	16760	885629	89174	21850

Notes: Data are from the PUMS of the 1960-1990 U.S. Censuses. Ages 16-65 are included. [1]Asian and Other Race are combined in 1960 and 1970. [2]In 1960 Hispanic are those with "spanish" surnames; in the other years it refers to those who self-identify as hispanic. [3] Only those with valid log hourly wages are included in this calculation. [4]The sample size is for the entire sample, not just those with valid wages.

Table 1b: Descriptive Statistics for Women : Native-born, Immigrants, and Recent Immigrants, by Year
(Standard Errors)

	1960			1970		
	Native-born	Immigrants	Recent Immigrants	Native-born	Immigrants	Recent Immigrants
Age	38.140	46.535	31.972	37.629	41.411	32.166
	(0.0193)	(0.0755)	(0.1666)	(0.0189)	(0.0764)	(0.1390)
H.S. Drop out	0.523	0.680	0.575	0.414	0.516	0.524
	(0.0007)	(0.0026)	(0.0075)	(0.0006)	(0.0027)	(0.0060)
H.S. Degree	0.321	0.207	0.253	0.383	0.296	0.252
	(0.0007)	(0.0023)	(0.0066)	(0.0006)	(0.0024)	(0.0052)
Some College	0.099	0.074	0.115	0.125	0.112	0.111
	(0.0004)	(0.0015)	(0.0048)	(0.0004)	(0.0017)	(0.0038)
College Graduate	0.057	0.039	0.057	0.078	0.077	0.112
	(0.0003)	(0.0011)	(0.0035)	(0.0003)	(0.0014)	(0.0038)
Black	0.107	0.015	0.023	0.111	0.036	0.076
	(0.0004)	(0.0007)	(0.0023)	(0.0004)	(0.0010)	(0.0032)
Asian	--	--	--	--	--	--
Other Race	0.007	0.033	0.091	0.009	0.075	0.156
	(0.0001)	(0.0010)	(0.0043)	(0.0001)	(0.0014)	(0.0044)
Hispanic	0.014	0.060	0.073	0.032	0.205	0.347
	(0.0002)	(0.0013)	(0.0039)	(0.0002)	(0.0021)	(0.0057)
Married	0.714	0.735	0.713	0.666	0.718	0.687
	(0.0006)	(0.0025)	(0.0068)	(0.0006)	(0.0024)	(0.0056)
Metropolitan	0.580	0.812	0.802	0.629	0.817	0.845
Area	(0.0007)	(0.0022)	(0.0060)	(0.0006)	(0.0021)	(0.0044)
Log Hrly Wage	1.830	1.853	1.690	2.048	2.081	1.974
1990 $'s	(0.0017)	(0.0063)	(0.0164)	(0.0013)	(0.0057)	(0.0131)
Fraction of the	--	0.060	0.008	--	0.055	0.011
Population		(0.0003)	(0.0001)		(0.0003)	(0.0001)
Sample Size	507058	32294	4399	593616	34858	6926

	1980			1990		
	Native-born	Immigrants	Recent Immigrants	Native-born	Immigrants	Recent Immigrants
Age	36.903	38.600	31.328	37.625	38.341	31.416
	(0.0150)	(0.0512)	(0.0939)	(0.0142)	(0.0431)	(0.0789)
H.S. Drop out	0.292	0.411	0.459	0.206	0.392	0.421
	(0.0005)	(0.0019)	(0.0040)	(0.0004)	(0.0016)	(0.0035)
H.S. Degree	0.410	0.301	0.239	0.322	0.223	0.210
	(0.0005)	(0.0017)	(0.0034)	(0.0005)	(0.0014)	(0.0029)
Some College	0.178	0.162	0.154	0.296	0.214	0.183
	(0.0004)	(0.0014)	(0.0029)	(0.0005)	(0.0014)	(0.0027)
College Graduate	0.119	0.127	0.147	0.175	0.172	0.186
	(0.0003)	(0.0013)	(0.0028)	(0.0004)	(0.0013)	(0.0027)
Black	0.122	0.065	0.073	0.128	0.080	0.080
	(0.0003)	(0.0009)	(0.0021)	(0.0003)	(0.0009)	(0.0019)
Asian	0.006	0.166	0.311	0.006	0.253	0.320
	(0.0001)	(0.0014	(0.0037)	(0.0001)	(0.0014)	(0.0033)
Other Race	0.020	0.138	0.191	0.027	0.183	0.205
	(0.0001)	(0.0013)	(0.0031)	(0.0002)	(0.0013)	(0.0028)
Hispanic	0.035	0.316	0.336	0.044	0.395	0.423
	(0.0002)	(0.0018)	(0.0038)	(0.0002)	(0.0016)	(0.0035)
Married	0.601	0.681	0.630	0.569	0.642	0.580
	(0.0005)	(0.0018)	(0.0039)	(0.0005)	(0.0016)	(0.0035)
Metropolitan	0.803	0.940	0.952	0.825	0.958	0.962
Area	(0.0004)	(0.0009)	(0.0017)	(0.0004)	(0.0007)	(0.0013)
Log Hrly Wage	1.991	2.026	1.871	2.042	2.047	1.849
1990 $'s	(0.0008)	(0.0033)	(0.0076)	(0.0008)	(0.0029)	(0.0070)
Fraction of the	--	0.069	0.016	--	0.089	0.020
Population		(0.0003)	(0.0001)		(0.0003)	(0.0001)
Sample Size	939397	69125	15643	921739	90347	20225

Notes: see previous table.

Table 2a: Fraction of Immigrants in Each Quintile of the Native-born Log Hourly Wage Distribution
Men
(Standard Errors)

	1960	1970	Absolute Diff. 1960 to 1970	Pct. Change 1960 to 1970[1]	1980	1990	Absolute Diff. 1960 to 1990	Pct. Change 1960 to 1990
Quintile				All Immigrants				
1	0.105	0.153	0.048	45.714	0.158	0.203	0.098	93.333
	(0.0022)	(0.0026)	(0.0034)	(0.0395)	(0.0016)	(0.0016)	(0.0027)	(0.0432)
2	0.150	0.178	0.028	18.667	0.215	0.231	0.081	54.000
	(0.0026)	(0.0027)	(0.0037)	(0.0275)	(0.0019)	(0.0016)	(0.0031)	(0.0284)
3	0.245	0.235	-0.010	-4.082	0.221	0.207	-0.038	-15.510
	(0.0031)	(0.0030)	(0.0043)	(0.0172)	(0.0019)	(0.0016)	(0.0035)	(0.0126)
4	0.283	0.208	-0.075	-26.502	0.185	0.152	-0.131	-46.290
	(0.0032)	(0.0029)	(0.0043)	(0.0131)	(0.0018)	(0.0014)	(0.0035)	(0.0079)
5	0.217	0.227	0.010	4.608	0.221	0.207	-0.010	-4.608
	(0.0029)	(0.0030)	(0.0042)	(0.0195)	(0.0019)	(0.0016)	(0.0033)	(0.0147)
Quintile				Recent Immigrants				
1	0.254	0.246	-0.008	3.150	0.281	0.356	0.102	40.157
	(0.0084)	(0.0068)	(0.0108)	(0.0422)	(0.0042)	(0.0039)	(0.0093)	(0.0488)
2	0.152	0.190	0.038	25.00	0.265	0.229	0.077	50.658
	(0.0069)	(0.0061)	(0.0092)	(0.0699)	(0.0041)	(0.0034)	(0.0077)	(0.0715)
3	0.198	0.186	-0.012	-6.061	0.170	0.147	-0.051	-25.758
	(0.0077)	(0.0061)	(0.0077)	(0.0477)	(0.0035)	(0.0029)	(0.0082)	(0.0325)
4	0.187	0.169	-0.018	-9.626	0.126	0.108	-0.079	-42.246
	(0.0075)	(0.0059)	(0.0095)	(0.0480)	(0.0031)	(0.0025)	(0.0079)	(0.0268)
5	0.208	0.208	0.000	0.000	0.158	0.160	-0.048	-23.077
	(0.0078)	(0.0064)	(0.0101)	(0.0482)	(0.0034)	(0.0030)	(0.0084)	(0.0324)

Notes: The quintiles are defined as follows: First we ran a log hourly wage regression for the native-born, holding constant age, age squared, 3 education categories (college graduate omitted), race and ethnicity controls (white non-Hispanic omitted), marital status, and metropolitan residence. We then predict wages for everyone, and divide the predicted wages into quintiles.

[1]The standard errors for the percentage changes are calculated as follows: We first generate 10,000 bootstrapped estimates of the percent change assuming that the components are independent normal variables with means and standard deviations given in columns (1) and (2) of the table. Let q_i represent the i th quintile from the distribution of bootstrapped estimates of the percent change. Let $L = \varphi^{-1}(-2)$ and $U = \varphi^{-1}(2)$, where $\varphi^{-1}(.)$ is the inverse cumulative standard normal distribution function. The "standard errors" in columns 4 and 8 are then computed as $\frac{|q_L - q_U|}{4}$.

Table 2b: Fraction of Immigrants in Each Quintile
of the Native-born Log Hourly Wage Distribution
Women
(Standard Errors)

	1960	1970	Absolute Diff. 1960 to 1970	Pct. Change 1960 to 1970	1980	1990	Absolute Diff. 1960 to 1990	Pct. Change 1960 to 1990
Quintile				All Immigrants				
1	0.085	0.131	0.046	54.118	0.139	0.138	0.053	62.235
	(0.0029)	(0.0029)	(0.0041)	(0.0631)	(0.0018)	(0.0015)	(0.0033)	(0.0584)
2	0.186	0.284	0.098	52.688	0.188	0.218	0.032	17.204
	(0.0040)	(0.0039)	(0.0056)	(0.0389)	(0.0020)	(0.0018)	(0.0044)	(0.0272)
3	0.326	0.179	-0.147	-45.509	0.232	0.218	-0.108	-33.129
	(0.0048)	(0.0033)	(0.0058)	(0.0127)	(0.0022)	(0.0018)	(0.0051)	(0.0112)
4	0.207	0.193	-0.014	-6.763	0.195	0.178	-0.029	-14.010
	(0.0042)	(0.0034)	(0.0054)	(0.0254)	(0.0020)	(0.0017)	(0.0045)	(0.0194)
5	0.197	0.214	0.017	8.629	0.246	0.248	0.051	25.888
	(0.0041)	(0.0035)	(0.0054)	(0.0288)	(0.0022)	(0.0019)	(0.0045)	(0.0276)
Quintile				Recent Immigrants				
1	0.098	0.215	0.117	119.388	0.249	0.258	0.160	163.265
	(0.0076)	(0.0079)	(0.0110)	(0.1936)	(0.0050)	(0.0045)	(0.0088)	(0.2156)
2	0.288	0.313	0.025	8.681	0.199	0.221	-0.067	-23.264
	(0.0120)	(0.0090)	(0.0150)	(0.0551)	(0.0046)	(0.0043)	(0.0127)	(0.0358)
3	0.212	0.130	-0.082	-38.679	0.207	0.179	-0.033	-15.566
	(0.0104)	(0.0065)	(0.0123)	(0.0421)	(0.0047)	(0.0039)	(0.0111)	(0.0450)
4	0.222	0.115	-0.107	-48.198	0.113	0.104	-0.118	-53.163
	(0.0106)	(0.0062)	(0.0123)	(0.0381)	(0.0036)	(0.0031)	(0.0110)	(0.0266)
5	0.179	0.227	0.048	26.682	0.232	0.238	0.059	32.961
	(0.0099)	(0.0081)	(0.0128)	(0.0839)	(0.0049)	(0.0044)	(0.0108)	(0.0776)

Notes: See previous table.

Table 3a: Native-Born and Recent Immigrant Log Wages
Evaluated at Skill Prices and Characteristics from 1970 to 1990
Men

	Native Born				Recent Immigrants			
	'90 Prices '90 X's	'70 Prices '70 X's	'90 Prices '70 X's	'70 Prices '90 X's	'90 Prices '90 X's	'70 Prices '70 X's	'90 Prices '70 X's	'70 Prices '90 X's
	N_{90}^{90}	N_{70}^{70}	N_{70}^{90}	N_{90}^{70}	I_{90}^{90}	I_{70}^{70}	I_{70}^{90}	I_{90}^{70}
25th percentile	1.879	2.120	1.878	2.166	1.534	1.904	1.668	1.866
50th percentile	2.377	2.488	2.337	2.538	1.906	2.287	2.101	2.270
75th percentile	2.807	2.796	2.733	2.902	2.377	2.690	2.621	2.692
Mean	2.337	2.447	2.300	2.517	1.995	2.293	2.170	2.271
(std. dev.)	(0.719)	(0.646)	(0.685)	(0.688)	(0.725)	(0.701)	(0.737)	(0.745)
Obs.	675249	362864	675249	362864	15154	4014	15154	4014

	Changes in the Native - Recent Immigrant Wage Gap Evaluated at 1990 and 1970 Skill Prices	
	1990 Skill Prices	1970 Skill Prices
	$[N_{90}^{90} - I_{90}^{90}] - [N_{70}^{90} - I_{70}^{90}]$	$[N_{90}^{70} - I_{90}^{70}] - [N_{70}^{70} - I_{70}^{70}]$
25th percentile	0.135	0.084
50th percentile	0.235	0.067
75th percentile	0.318	0.104
Mean	0.212	0.092

Notes: These numbers were calculated using the 1970 and 1990 PUMS of the U.S. Census. In the top panel, the columns with 1990 prices and 1990 X's are simply the log wages for 1990. Columns with 1990 Prices and 1970 X's are calculated by re-weighting the 1990 individuals to have the same characteristics as the 1970 individuals. Other columns are defined analogously. The bottom panel calculates changes in the native-born recent immigrant wage gap. In both cases, the characteristics changes are the actual changes between 1970 and 1990. These changes in characteristics are evaluated at the skill price distribution in 1990 in the first column, and the skill price distribution in 1970 in the second column. In the equations above, the superscript denotes the skill prices and the subscript denotes the characteristics. See the text for a more detailed explanation.

Table 3b: Native-Born and Recent Immigrant Log Wages
Evaluated at Skill Prices and Characteristics from 1970 to 1990
Women

	Native Born				Recent Immigrants			
	'90 Prices '90 X's	'70 Prices '70 X's	'90 Prices '70 X's	'70 Prices '90 X's	'90 Prices '90 X's	'70 Prices '70 X's	'90 Prices '70 X's	'70 Prices '90 X's
	N_{90}^{90}	N_{70}^{70}	N_{70}^{90}	N_{90}^{70}	I_{90}^{90}	I_{70}^{70}	I_{70}^{90}	I_{90}^{70}
25th percentile	1.609	1.677	1.504	1.788	1.419	1.599	1.427	1.572
50th percentile	2.040	2.049	1.907	2.166	1.781	1.933	1.783	1.991
75th percentile	2.474	2.407	2.303	2.543	2.238	2.289	2.238	2.337
Mean	2.042	2.048	1.916	2.179	1.849	1.974	1.849	1.985
(std. dev.)	(0.662)	(0.686)	(0.648)	(0.700)	(0.679)	(0.670)	(0.639)	(0.716)
Obs.	624373	226994	624373	226994	9410	2625	9410	2625
	Changes in the Native - Recent Immigrant Wage Gap Evaluated at 1990 and 1970 Skill Prices							
	1990 Skill Prices $[N_{90}^{90} - I_{90}^{90}] - [N_{70}^{90} - I_{70}^{90}]$				1970 Skill Prices $[N_{90}^{70} - I_{90}^{70}] - [N_{70}^{70} - I_{70}^{70}]$			
25th percentile	0.113				0.138			
50th percentile	0.135				0.059			
75th percentile	0.171				0.088			
Mean	0.126				0.120			

Notes: See previous table.

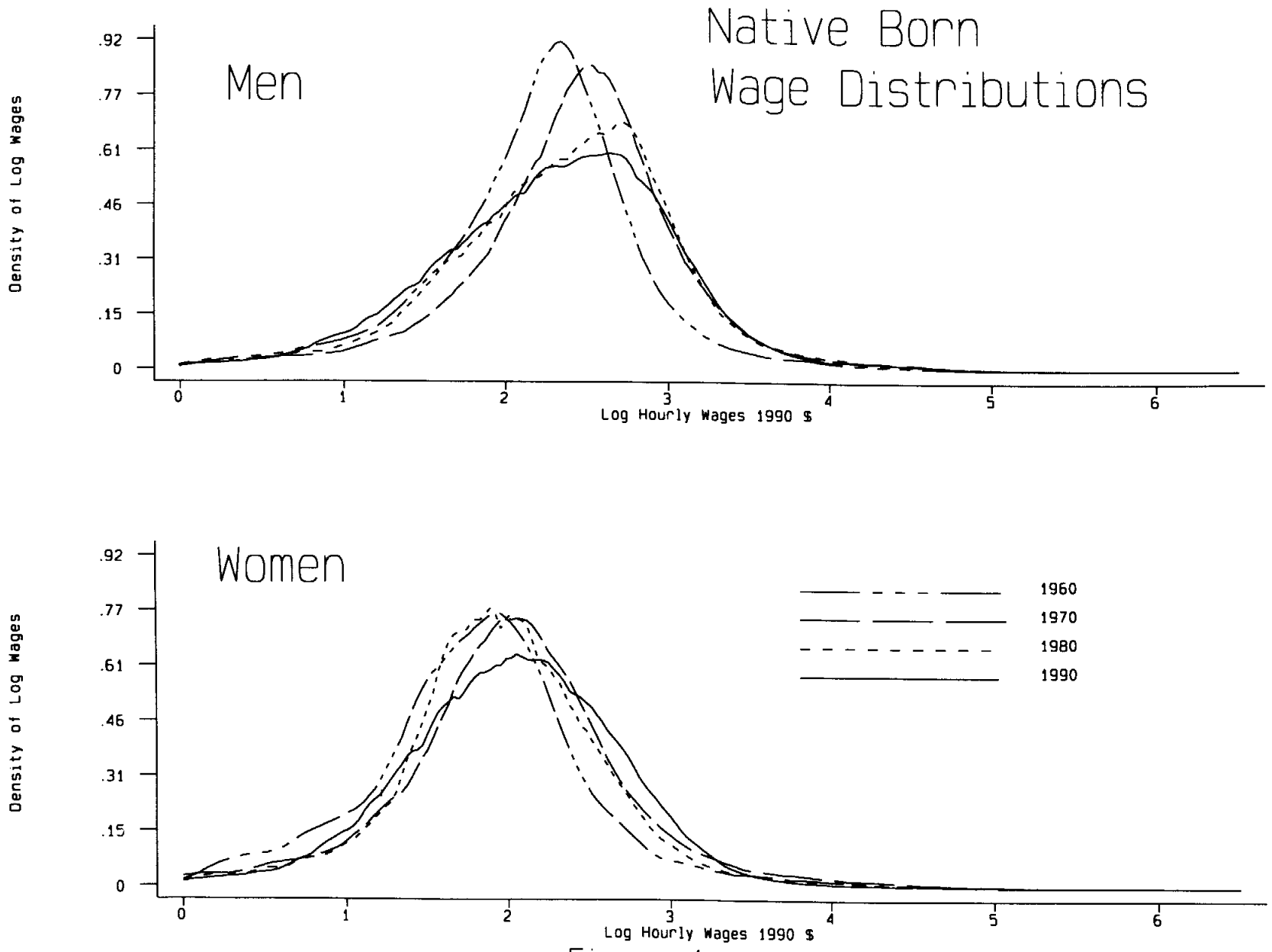

Native Born
Wage Distributions
Men
Women
Density of Log Wages
.92
.77
.61
.46
.31
.15
0
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
1960
1970
1980
1990

Figure 1

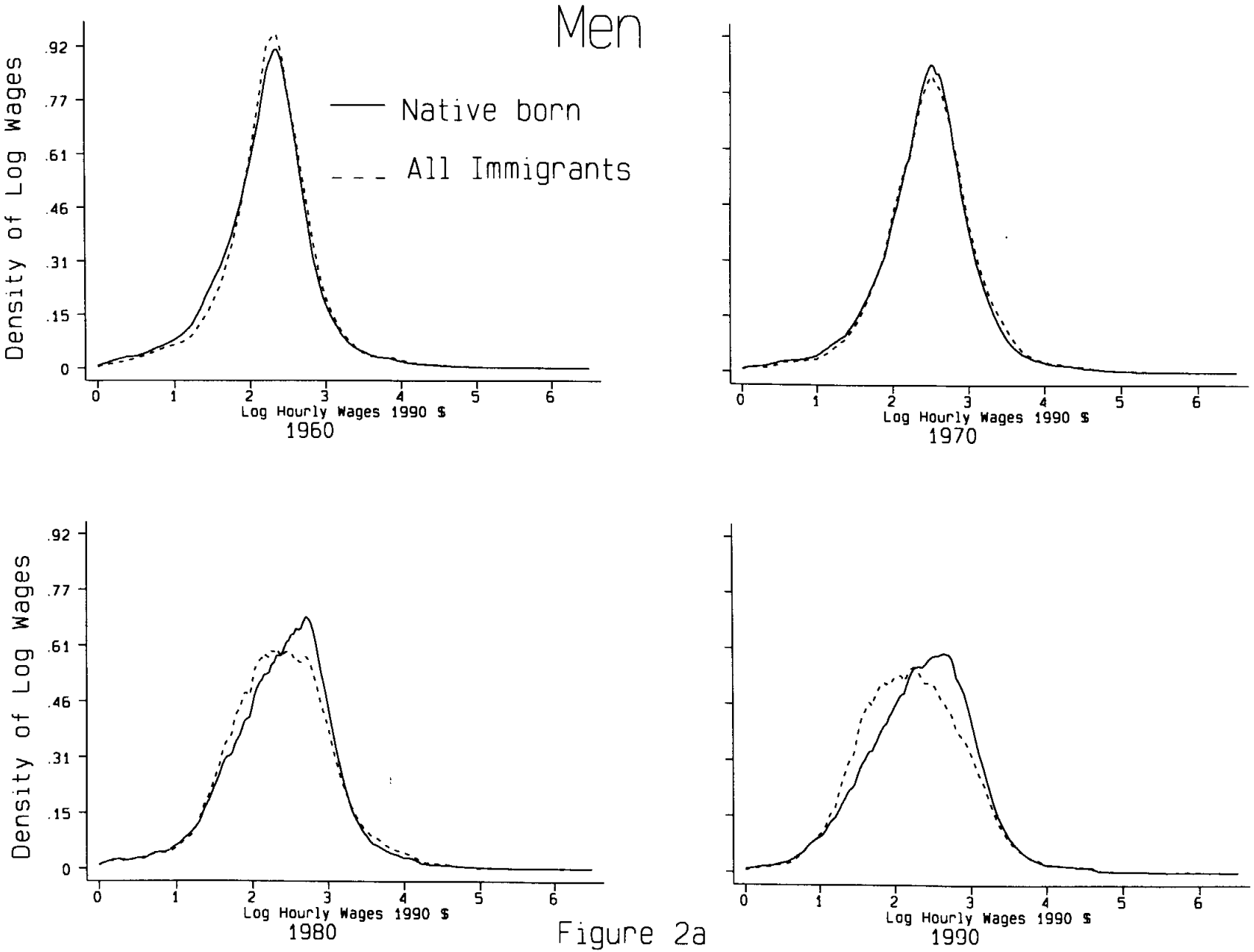

Men
Native born
All Immigrants
Density of Log Wages
.92
.77
.61
.46
.31
.15
0
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
1960
1970
1980
1990

Figure 2a

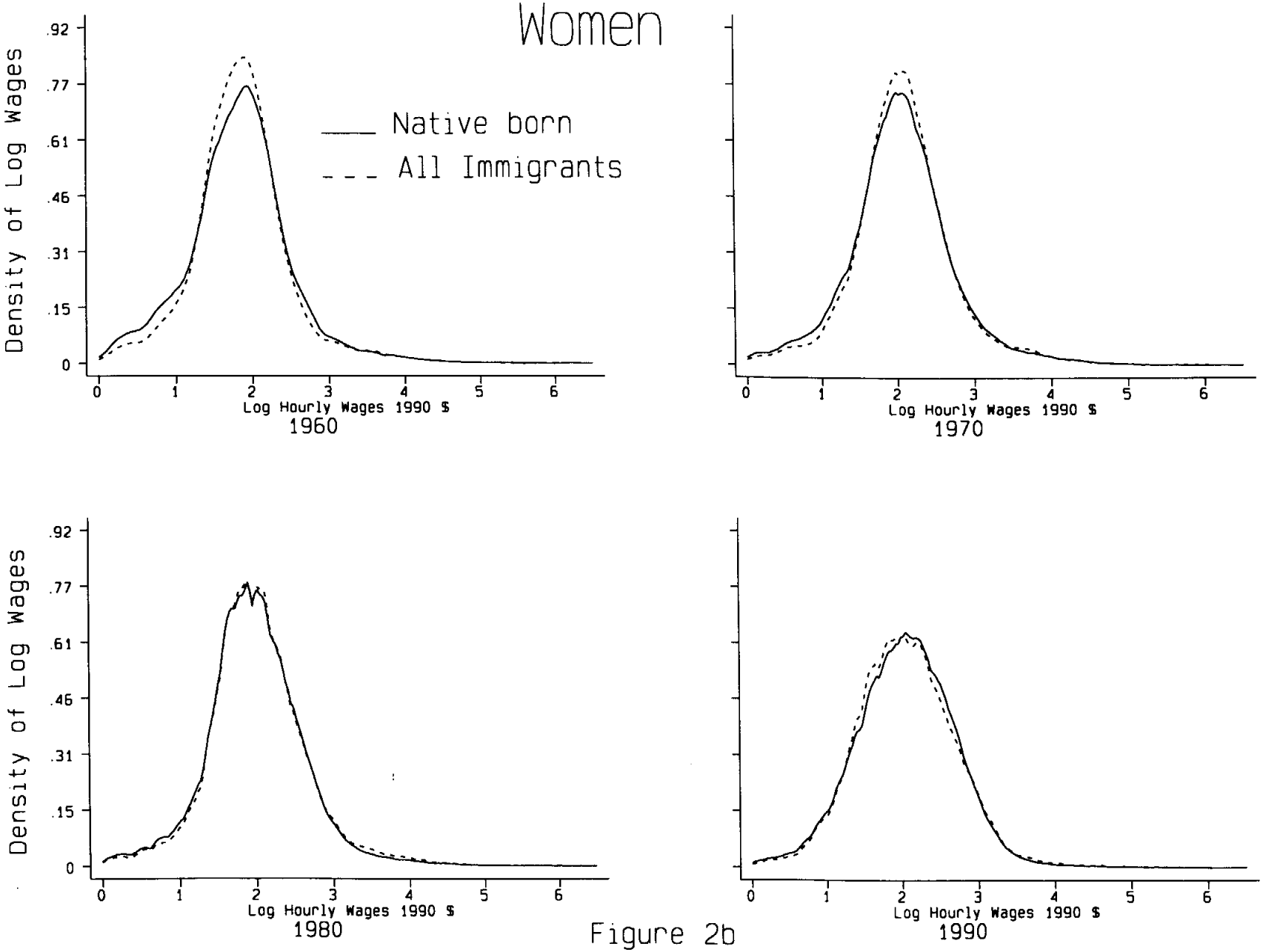

Women
Native born
All Immigrants
Density of Log Wages
.92
.77
.61
.46
.31
.15
0
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
1960
1970
1980
1990

Figure 2b

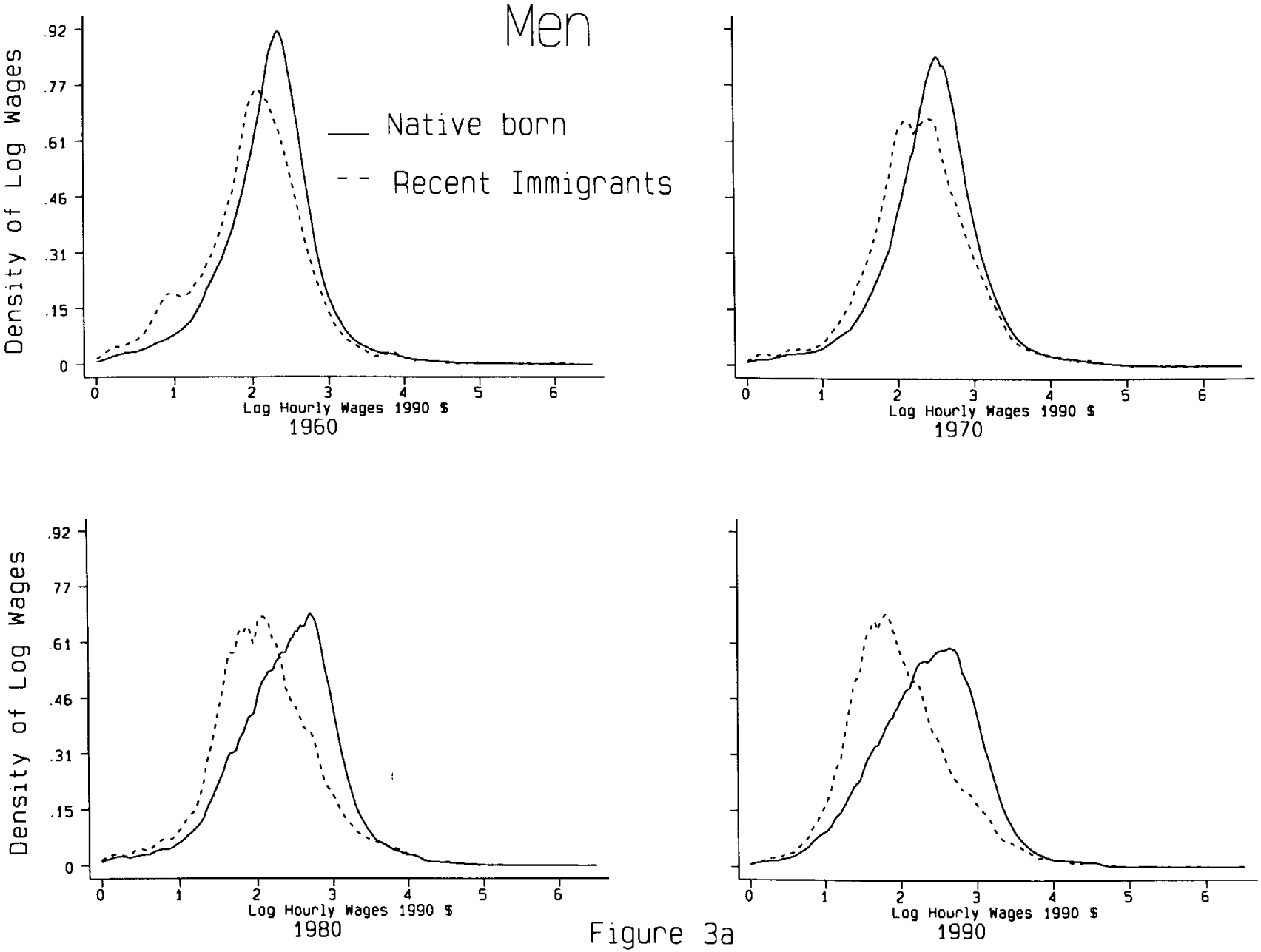

Men
Native born
Recent Immigrants
Density of Log Wages
.92
.77
.61
.46
.31
.15
0
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
1960
1970
1980
1990

Figure 3a

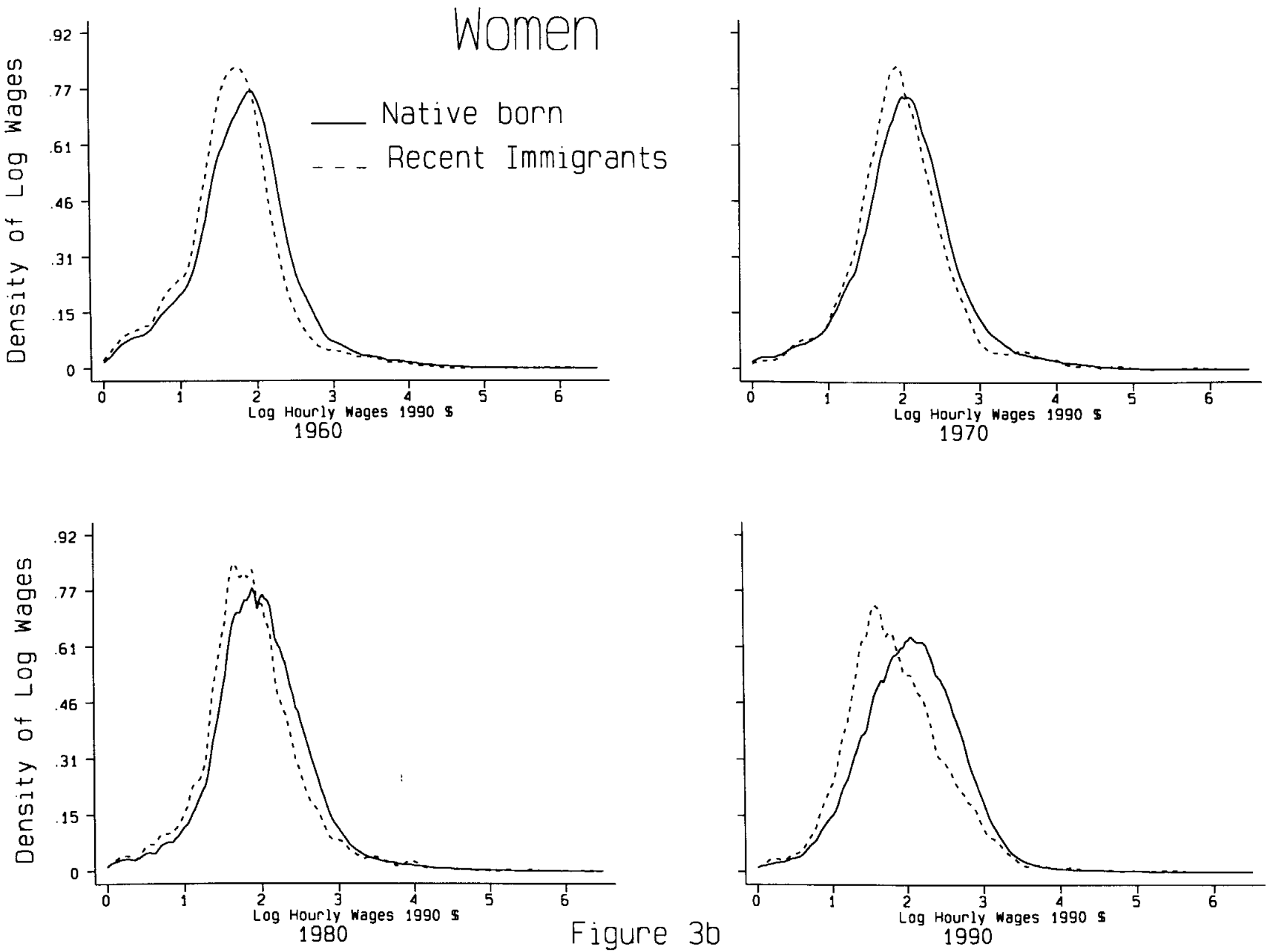

Women
Native born
Recent Immigrants
Density of Log Wages
.92
.77
.61
.46
.31
.15
0
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
1960
1970
1980
1990

Figure 3b

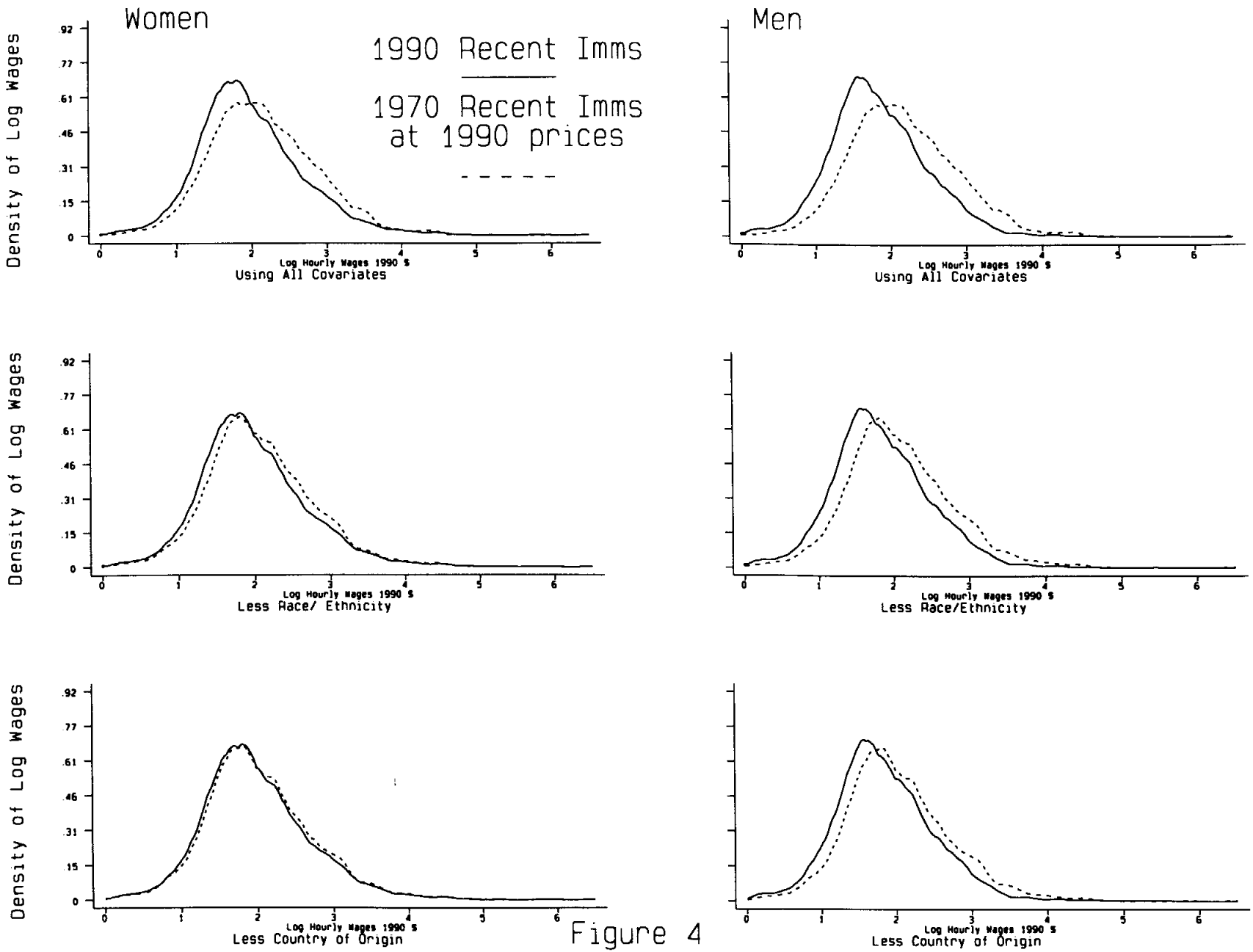

Women
Men
1990 Recent Imms
1970 Recent Imms
at 1990 prices
Density of Log Wages
.92
.77
.61
.46
.31
.15
0
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
Using All Covariates
Less Race/ Ethnicity
Less Race/Ethnicity
Less Country of Origin

Figure 4

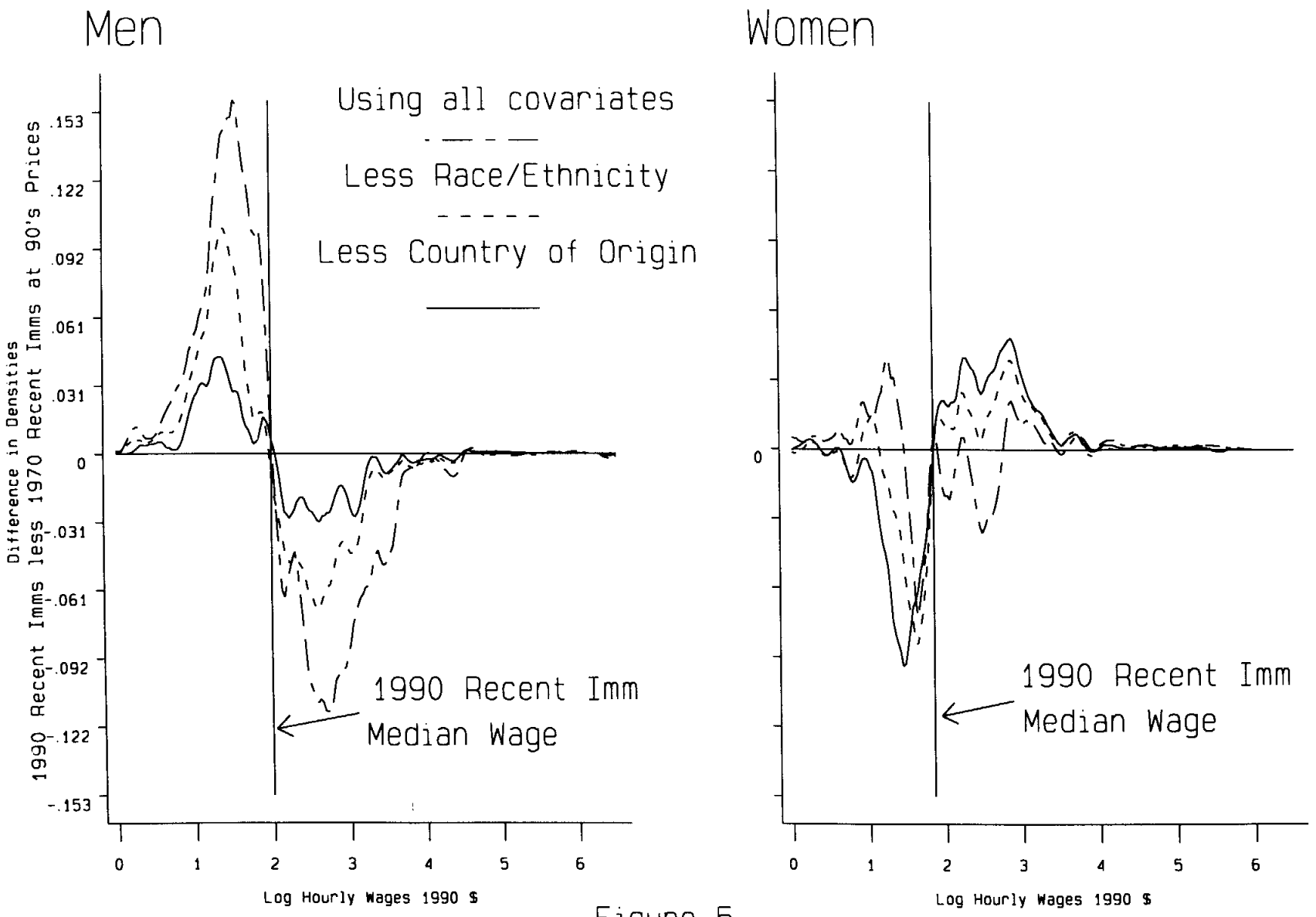
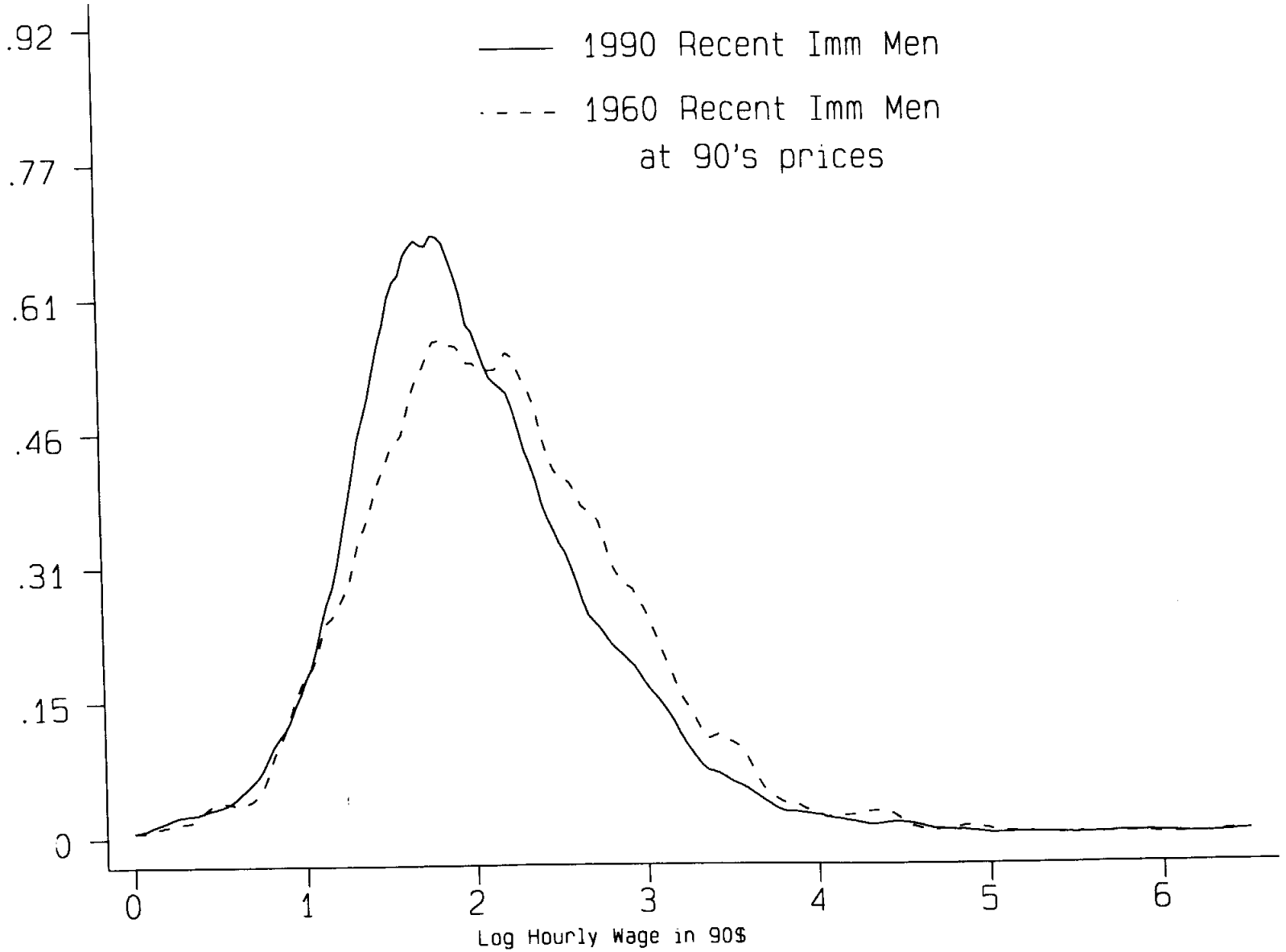

Men
Women
Using all covariates
Less Race/Ethnicity
Less Country of Origin
Difference in Densities
1990 Recent Imms less 1970 Recent Imms at 90's Prices
.153
.122
.092
.061
.031
0
-.031
-.061
-.092
-.122
-.153
1990 Recent Imm
Median Wage
1990 Recent Imm
Median Wage
0
1
2
3
4
5
6
Log Hourly Wages 1990 $
Log Hourly Wages 1990 $

Figure 5



Appendix Figure 1