Wage Gaps And Demographic Behavior Essay, Research Paper

WHY ARE RACIAL AND ETHNIC WAGE GAPS LARGER FOR MEN THAN FOR WOMEN?

EXPLORING THE ROLE OF SEGREGATION USING THE

NEW WORKER-ESTABLISHMENT CHARACTERISTICS DATABASE

Kimberly Bayard, Judith Hellerstein, David Neumark, and Kenneth Troske\*

September 1998

\*Bayard is a Ph.D. student in economics at the University of Maryland. Hellerstein is Assistant Professor of Economics at the University of Maryland, and a Faculty Research Fellow of the NBER. Neumark is Professor of Economics at Michigan State University, and a Research Associate of the NBER. Troske is Assistant Professor of Economics at the University of Missouri. This research was supported by NSF grant SBR95-10876 through the NBER. The research in this paper was conducted while the authors were research associates with the Center for Economic Studies, U.S. Bureau of the Census. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Bureau of the Census or the Center for Economic Studies.

Abstract

We examine the possible sources of the larger racial and ethnic wage gaps for men than for women in the U.S. Specifically, using a newly created employer-employee matched data set containing workers in essentially all occupations, industries, and regions, we examine whether these wage differences can be accounted for by differences between men and women in the patterns of racial and ethnic segregation within occupation, industry, establishments and occupation-establishment cells. To the best of our knowledge, this is the first paper to examine segregation by race and ethnicity at the level of establishment and job cell. Our results indicate that greater segregation between Hispanic men and white men than between Hispanic women and white women accounts for essentially all of the higher Hispanic-white wage gap for men. In addition, our estimates indicate that greater segregation between black and white men than between black and white women accounts for a sizable share (one-third to one-half) of the higher black-white wage gap for men. Our results imply that segregation is an important contributor to the lower wages paid to black and Hispanic men than to white men with similar individual characteristics. Our results also suggest that equal pay types of laws may offer some scope for reducing the black-white wage differential for men, but little scope for reducing the Hispanic-white wage differential for men.

I. Introduction

Labor economists have long been occupied with explorations of the sources of wage differences by sex, race, and ethnicity. It is well known that wages earned by minorities and by females fall short of wages earned by white males, after accounting for differences in standard human capital proxies and other variables for which measures are readily available in many micro-level data sets (schooling, age or experience, marital status, urban residence, region, etc.).

Aside from this general fact, an additional fact about racial and ethnic wage gaps is that they are considerably larger for men than for women. This is true in the raw data, as well as once we account for numerous determinants of wages or earnings. For example, based on 1981 CPS data, Cain (1986, Table 13.4) reports that for all workers, black-white earnings ratios are 0.67 for men vs. 0.97 for women, while Hispanic-white earnings ratios are 0.72 for men and 0.90 for women. For full-time, year-round workers, black-white earnings ratios are 0.69 for men vs. 0.90 for women, while Hispanic-white earnings ratios are 0.72 for men and 0.87 for women.(1) As a second example, as we report later in this paper, in log wage regressions including controls for schooling, age, etc., based on the 1990 Census of Population, the estimated black-white (actually, black vs. non-black, non-Hispanic) earnings differential is -0.121 for men vs. -0.022 for women, while the Hispanic-white differential is -0.115 for men vs. -0.045 for women. Finally, in a cross-section of 1990 and 1991 observations from the NLSY, in log wage regressions with no controls Neal and Johnson (1996) report that black men earn 24.4 percent less than white men vs. an 18.5 percent shortfall for black women, while Hispanic men earn 11.3 percent less than white men vs. a 2.8 percent (and insignificant) shortfall for Hispanic women.

When Neal and Johnson control for AFQT (interpreted as a catch-all for pre-market factors affecting wages), the black-white difference for men falls to -7.2 percent, while black women are estimated to earn 3.5 percent more than white women (an insignificant difference).(2) Thus, even if one believes the Neal and Johnson claim that pre-market factors account for a sizable fraction of racial and ethnic wage differences, the fact that the difference in the black-white wage gap between men and women persists suggests that this difference is a “labor market” rather than a “pre-market” phenomenon.

In our view the larger racial and ethnic wage gaps for men than for women are a rather striking set of stylized facts that have largely been ignored by researchers attempting to understand the sources of racial and ethnic wage differences. In this paper we examine more closely the possible sources of the differences in the wage gap, paying particular attention to whether these differences can be accounted for by differences between men and women in the patterns of racial and ethnic segregation.(3) More generally, we believe that research on why racial and ethnic wage gaps differ by sex may ultimately prove useful in helping to understand the sources of these gaps. For example, if one believes that the observed wage differentials are the result of employer or customer discrimination (e.g., Darity and Mason, 1998) then one needs to try to explain why this discrimination is apparently more severe with respect to male employees. In general, if one believes that some other unmeasured characteristic is responsible for these wage differences, then evidence that this characteristic is more important for men than for women would bolster one’s case.

This inquiry fits into an extensive literature on the role of segregation in generating racial, ethnic, and sex differences in labor markets, but takes this literature in a new direction. In the literature on sex differences in wages, considerable attention has focused on the role of occupational segregation, in particular the concentration of women in low-wage occupations (e.g., Johnson and Solon, 1986; Sorensen, 1989; Macpherson and Hirsch, 1995). However, relatively little attention has been paid to the role of occupational segregation in generating racial and ethnic differences in wages (for an exception, see Sorensen, 1989), in part because occupational segregation between races and ethnic groups is much less pronounced than occupational segregation between the sexes (King, 1992; Watts, 1995).

Furthermore, even less attention has been paid to the role of segregation along other dimensions such as industry, employer, and job cell (occupation within employer). The main reason for the lack of such work is that the data sets labor economists typically use to study wage differences are household data sets, which allow one to measure the percent female or black in an occupation or industry, but not the sex, race, or ethnic composition of firms, establishments or jobs. Economists interested in studying these other dimensions of segregation have had to turn to other special data sources in which information on the workforce is available or can be constructed. For example, Groshen (1991) uses data from the Bureau of Labor Statistics Industry Wage Surveys, with which one can measure the percent female by establishment as well as job cell. Blau (1977) studies BLS Area Wage Surveys, which cover clerical, professional, and technical occupations, and which allow the estimation of percent female along the same dimensions. Bayard, et al. (1998) construct a data set (called the New Worker-Establishment Characteristics Database, or NWECD) based on a match of employees to their establishments, and carry out an analysis of the roles of sex segregation by occupation, industry, establishment, and job cell, similar to Groshen’s. While there are differences in the findings reported in these studies, all find that in addition to being concentrated in low-wage occupations, women are also concentrated in low-wage establishments and low-wage job cells.(4)

In this paper, we use the NWECD to study the role of racial and ethnic segregation in generating wage differences between whites, blacks, and Hispanics. The NWECD is uniquely suited to this analysis, as the Industry and Area Wage Surveys contain no information on race and ethnicity. Thus, to the best of our knowledge, this is the first paper that looks at segregation by race and ethnicity at the level of the establishment and job cell.(5) We consider evidence on the effects of racial and ethnic segregation on wages, and the extent to which racial and ethnic wage differences remain after controlling for segregation. Such evidence helps to assess whether equal pay policies are likely to reduce these wage differences (assuming that these remaining differences reflect discrimination).(6) We are particularly interested in the question posed in the title of this paper, namely whether more severe racial and ethnic segregation among men can explain why racial and ethnic wage gaps are bigger among men than among women.

II. The Data

The NWECD is created from two data sources, the Sample Detail file (SDF), which contains all individual responses to the 1990 Decennial Census one-in-six Long Form, and the 1990 Standard Statistical Establishment List (SSEL), which is an administrative database containing information for all business establishments operating in the United States in 1990. We construct the NWECD by using detailed location and industry information available in both data sets to match worker records in the SDF to employer records in the SSEL. In this section we discuss the details of the matching process, assess the accuracy of the match, and discuss the representativeness of these matched data.

The Matching Process

Households receiving the 1990 Decennial Census Long Form were asked to report the name and address of the employer in the previous week for each employed member of the household. In addition, respondents were asked for the name and a brief (one or two word) description of the type of business or industry of the most recent employer for all members of the household. Based on the responses to these questions the Census Bureau assigned geographic and industry codes to each record in the data and it is these codes that are available in the SDF. In addition to this information, the SDF contains the standard set of demographic characteristics collected on the long-form of the Decennial Census. To construct the NWECD we first selected records for the slightly more than 17 million respondents who indicated they were employed in the previous week.

The SSEL is an annually updated list of all business establishments with one or more employees operating in the United States that the Census Bureau uses as a sampling frame for its various Economic Censuses and Surveys. As such, the SSEL contains the name and address of each establishment, geographic codes based on its location, and a four-digit SIC code. In addition, the SSEL contains data on the number of employees and total annual payroll for the establishment, a unique establishment identifier, as well as an identifier that allows the establishment to be linked to other establishments that are part of the same enterprise. To construct the NWECD, we selected the 5.6 million records from the 1990 SSEL. We focus on the private sector, excluding establishments in Public Administration.

Matching workers to employers proceeded in four steps. First, we standardized the geographic and industry codes in the two data sets. Next, we selected all establishments that were unique in an industry-location cell. Third, all workers who indicated they worked in the same industry-location cell as a unique establishment were matched to the establishment. Finally, we eliminated all matches based on imputed data. The resulting data set is what we call the NWECD.

There are a number of issues involved in the matching process that merit further discussion. The first set of issues concerns standardizing the geographic and industry codes. The Census Bureau divides the country into a hierarchy of geographic areas. For our purposes the relevant areas are state, county, place, tract, and block. The Census Bureau assigns a unique code to every state in the country. Within each state the Census Bureau assigns a unique code to every county. The Census Bureau also assigns a unique place code to population centers with 2,500 or more people. Because these population centers are unique within a state, but can cross county boundaries, we can distinguish between areas in the same place located in different counties. Finally, the Census Bureau divides up populated counties into unique tracts and divides tracts up into unique blocks.(7) Thus, for an establishment located in a metropolitan area, the Census Bureau assigns a unique geographic code which identifies the state, county, place, tract, and block of the establishment.