I Z A

The Measurement of Racial Discrimination in Pay between Job Categories: Theory and Test

Örn B. Bodvarsson
John G. Sessions

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Measurement of Racial Discrimination in Pay between Job Categories: Theory and Test

Őrn B. Bodvarsson
St. Cloud State University
and IZA

John G. Sessions
University of Bath
and IZA

ABSTRACT

The Measurement of Racial Discrimination in Pay between Job Categories: Theory and Test[*]

The traditional model of taste discrimination in labor markets presumes perfect substitution, making it unsuitable for the measurement of discrimination across job assignments. We extend the model to explain cross-assignment discrimination and test it on data from Major League Baseball. A competitive firm with a Generalized Leontief production function fills each job assignment with whites and nonwhites in an environment of customer prejudice. According to the model, cross-assignment discrimination depends upon racial productivity differences, the productivity x prejudice interaction, technology, relative labor supply and racial integration. We find strong evidence of *ceteris paribus* racial salary differences between hitters and pitchers.

JEL Classification: J7

Keywords: wages, discrimination, imperfect substitutability, integration

Corresponding author:

Őrn B. Bodvarsson
Department of Economics
St. Cloud State University
720 Fourth Avenue South
St. Cloud, MN 56301-4498
USA
E-mail: obbodvarsson@stcloudstate.edu

I.   Introduction

Discrimination is defined as the *unequal treatment of equals*. In the labor market "equals" implies that minority and majority workers are perfect substitutes and "unequal treatment" implies that, despite the absence of productivity differences, members of the minority group are valued differently from members of the majority group. In the traditional theory of racial discrimination in the labor market, due originally to Becker (1971) and Arrow (1973), whites and nonwhites are perfect substitutes and it is shown that prejudice can result in unequal labor market outcomes. Despite this model's presumption of perfect substitution and the many tests of it[1], white and nonwhite labor groups are actually not perfectly substitutable. Grant and Hamermesh (1981) found that black adults are imperfect substitutes for white men and complements to white women and youths. Borjas (1983) provided evidence showing that black male workers are imperfect substitutes for white male workers, but Hispanic and white male workers are complementary. Borjas (1987) showed that black natives are imperfect substitutes for white natives and Kahanec (2006) used U.S. census data to confirm that nonwhite labor is complementary to white labor.

If whites and nonwhites are imperfectly substitutable, their human capital endowments must differ. But why? Earlier literature only provides some hints. Welch (1967) argued that blacks and whites working in the same firm are unlikely to be perfect substitutes because, due to long term discrimination, blacks may acquire less schooling and attend lower quality schools. Bodvarsson and Partridge (2001) argued that white and nonwhite professional basketball players could be imperfect substitutes because of differences in pre-NBA training and experience. Borjas (2008) suggested that black and white workers may not be perfect substitutes in a firm when there are anti-discrimination policies such as Affirmative Action in place.

---

[1] See Altonji and Blank (1999) for a comprehensive survey of the empirical literature.

One crucial condition for perfect substitution is that blacks and whites be employed in the same job assignment in a firm: Since different assignments require different skills, workers in assignment A must be imperfect substitutes for workers in assignment B, despite both contributing to production of the same good. The traditional model is equipped to answer the question: Are whites and nonwhites *within* the same job assignment paid differently? For example, is there a *ceteris paribus* pay difference between white pilots (flight attendants) and non-white pilots (flight attendants)? But, the model cannot address the question of "cross-job" discrimination, e.g. are white flight attendants (pilots) *ceteris paribus* paid differently from nonwhite pilots (flight attendants), or are non-white doctors *ceteris paribus* paid differently from white nurses?

The measurement of cross-job discrimination is an unexplored area. The literature offers very little insight into how the structure of the production function influences the structure of discrimination. Furthermore, very little is known about what sorts of empirical specifications must be used when whites and nonwhites are imperfect substitutes. Failure to properly account for the structure of the production function will, as Hashimoto and Kochin (1980) would argue, ultimately lead to biased estimates of discrimination.

In this paper, we examine cross-job discrimination. We articulate a production function where jobs are distinct inputs and extend Becker's (1971) *Market Discrimination Coefficient* (MDC) to measure the *ceteris paribus* racial pay gap across jobs. We find that discrimination varies in counterintuitive ways depending upon the production function and group differences in productivity and labor supply. We test the model with data from Major League Baseball (MLB) for the 1990s and we find strong evidence of *ceteris paribus* racial wage differences between hitters and pitchers.

II.   A theory of cross-job discrimination

II.1 *The problem setting*

Suppose production requires the completion of Jobs 1 and 2. The successful completion of Job 1 requires a different set of skills than what is required for the successful completion of Job 2. There are four inputs– white workers doing job 1 (2) and nonwhite workers doing job 1 (2). The firm assigns each worker to a particular job depending upon his/her observed skills and credentials. We allow for racial differences in productivity both within and across jobs. We assume that white customers are prejudiced against nonwhite workers (nonwhite customers are color-blind), and the firm must assign nonwhite workers to jobs for which prejudiced white customers must see them produce. The labor market is perfectly competitive and product price is normalized at unity.

Technology is characterized by a Generalized Leontief Production function[2]:

$$(1) \quad Q = \sum_{j=1}^{2} \sum_{i=1}^{2} \gamma_{ij} [X_i^W (DX_j^{NW})]^{\frac{1}{2}}, \quad D < 1,$$

where Q is output, X_i^W is the quantity of white labor input i, X_j^{NW} is the quantity of nonwhite labor input j, and γ_{ij} is the technology coefficient. The parameter D measures the strength of customer prejudice.[3] Customer prejudice may be viewed as a situation where customers discount the marginal revenue product (MRP) of nonwhite workers. The lower is D, the more intense the prejudice and the lower is nonwhite MRP. If D equals 1, there is no prejudice. While it is traditional to think of customer discrimination as implying a price discount on the output of nonwhite workers, the approach above is equivalent; D reflects the idea that nonwhite input is valued less when customers are prejudiced compared to the case where customers are not prejudiced. Note also that production is constant returns

---

[2] See Diewert (1971)

to scale and we restrict the values of the technology coefficients so that $\gamma_{ij} = \gamma_{ji}$. The sign of a coefficient indicates whether inputs i and j are substitutes ($\gamma_{ij} < 0$) or complements ($\gamma_{ij} > 0$).

Define r_i^W and r_j^{NW} as the market prices of white input i and nonwhite input j, respectively. The firm's profit function is

$$(2) \quad \pi = \sum_{j=1}^{2}\sum_{i=1}^{2}\gamma_{ij}[X_i^W(DX_j^{NW})]^{\frac{1}{2}} - \sum_{i=1}^{2}r_i^W X_i^W - \sum_{j=1}^{2}r_j^{NW} X_j^{NW}.$$

When firms maximize profits, the labor market establishes the following system of labor demand functions:

$$(3) \quad r_1^W = \gamma_{11} + \gamma_{12}D^{\frac{1}{2}}(\frac{X_1^{NW}}{X_1^W})^{\frac{1}{2}} + \gamma_{13}(\frac{X_2^W}{X_1^W})^{\frac{1}{2}} + \gamma_{14}D^{\frac{1}{2}}(\frac{X_2^{NW}}{X_1^W})^{\frac{1}{2}},$$

$$(4) \quad r_1^{NW} = \gamma_{22}D + \gamma_{12}D^{\frac{1}{2}}(\frac{X_1^W}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{23}D^{\frac{1}{2}}(\frac{X_2^W}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{24}D(\frac{X_2^{NW}}{X_1^{NW}})^{\frac{1}{2}},$$

$$(5) \quad r_2^W = \gamma_{33} + \gamma_{13}(\frac{X_1^W}{X_2^W})^{\frac{1}{2}} + \gamma_{23}D^{\frac{1}{2}}(\frac{X_1^{NW}}{X_2^W})^{\frac{1}{2}} + \gamma_{34}D^{\frac{1}{2}}(\frac{X_2^{NW}}{X_2^W})^{\frac{1}{2}},$$

$$(6) \quad r_2^{NW} = \gamma_{44}D + \gamma_{14}D^{\frac{1}{2}}(\frac{X_1^W}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{24}D(\frac{X_1^{NW}}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{34}D^{\frac{1}{2}}(\frac{X_2^W}{X_2^{NW}})^{\frac{1}{2}}.$$

Compensation to workers in each race/job group depends upon four factors – the group's productivity, customer prejudice, technology, and relative labor supplies. According to equation (4), for example, the wage paid to nonwhite workers doing job 1 depends upon: (i) the group's productivity (reflected by

---

[3] Kahn (1991) used a similar approach to incorporate discrimination into a production function.

$\gamma_{22})^4$; (ii) prejudice (D); (iii) the degrees of substitutability or complementarity between whites and nonwhites doing job 1 (γ_{12}), nonwhites doing job 1 and whites doing job 2 (γ_{23}), and nonwhites doing job 1 and nonwhites doing job 2 (γ_{24}); and (iv) the numbers of nonwhite workers doing job 1 per white worker doing job 1 ($\frac{X_1^{NW}}{X_1^W}$), white workers doing job 2 per nonwhite worker doing job 1 ($\frac{X_2^W}{X_1^{NW}}$), and nonwhite workers doing job 2 per nonwhite worker doing job 1 ($\frac{X_2^{NW}}{X_1^{NW}}$).

Becker's (1971) *Market Discrimination Coefficient* (MDC) measures the *ceteris paribus* racial earnings gap. For cross-assignment discrimination, the MDC for whites doing job 1 relative to nonwhites doing job 2 is

$$(7) \quad MDC_{NW_2}^{W_1} = \frac{r_1^W(D<1)}{r_2^{NW}(D<1)} - \frac{r_1^W(D=1)}{r_2^{NW}(D=1)} =$$

$$\frac{\gamma_{11} + D^{\frac{1}{2}}[\gamma_{12}(\frac{X_1^{NW}}{X_1^W})^{\frac{1}{2}} + \gamma_{14}(\frac{X_2^{NW}}{X_1^W})^{\frac{1}{2}}] + \gamma_{13}(\frac{X_2^W}{X_1^W})^{\frac{1}{2}}}{D[\gamma_{44} + \gamma_{24}(\frac{X_1^{NW}}{X_2^{NW}})^{\frac{1}{2}}] + D^{\frac{1}{2}}[\gamma_{14}(\frac{X_1^W}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{34}(\frac{X_2^W}{X_2^{NW}})^{\frac{1}{2}}]} - \frac{[\gamma_{11} + \gamma_{12}(\frac{X_1^{NW}}{X_1^W})^{\frac{1}{2}} + \gamma_{14}(\frac{X_2^{NW}}{X_1^W})^{\frac{1}{2}} + \gamma_{13}(\frac{X_2^W}{X_1^W})^{\frac{1}{2}}]}{[\gamma_{44} + \gamma_{24}(\frac{X_1^{NW}}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{14}(\frac{X_1^W}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{34}(\frac{X_2^W}{X_2^{NW}})^{\frac{1}{2}}]},$$

and for whites doing job 2 relative to nonwhites doing job 1 is

$$(8) \quad MDC_{NW_1}^{W_2} = \frac{r_2^W(D<1)}{r_1^{NW}(D<1)} - \frac{r_2^W(D=1)}{r_1^{NW}(D=1)} =$$

$$\frac{\gamma_{33} + \gamma_{13}(\frac{X_1^W}{X_2^W})^{\frac{1}{2}} + D^{\frac{1}{2}}[\gamma_{23}(\frac{X_1^{NW}}{X_2^W})^{\frac{1}{2}} + \gamma_{34}(\frac{X_2^{NW}}{X_2^W})^{\frac{1}{2}}]}{D[\gamma_{22} + \gamma_{24}(\frac{X_2^{NW}}{X_1^{NW}})^{\frac{1}{2}}] + D^{\frac{1}{2}}[\gamma_{12}(\frac{X_1^W}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{23}(\frac{X_2^W}{X_1^{NW}})^{\frac{1}{2}}]} - \frac{[\gamma_{33} + \gamma_{13}(\frac{X_1^W}{X_2^W})^{\frac{1}{2}} + \gamma_{23}(\frac{X_1^{NW}}{X_2^W})^{\frac{1}{2}} + \gamma_{34}(\frac{X_2^{NW}}{X_2^W})^{\frac{1}{2}}]}{[\gamma_{22} + \gamma_{24}(\frac{X_2^{NW}}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{12}(\frac{X_1^W}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{23}(\frac{X_2^W}{X_1^{NW}})^{\frac{1}{2}}]}.$$

---

[4] Note that γ_{22} is not the marginal productivity of this job/race group, but is correlated with it. If γ_{22} rises (falls), the marginal productivity curve will shift up (down). For example, an increase in γ_{22} could result from a technological advance, an increase in the average human capital endowment of each worker, or some other exogenous change.

According to equations (7) and (8), discrimination depends upon: (i) prejudice; and (ii) racial productivity differences within and across jobs; (iii) the degrees of substitutability or complementarity between whites and nonwhites within and across jobs; and (iv) the relative supplies of white and nonwhite labor within and across jobs.

Equations (7) and (8) yield important predictions:

(i) *The ceteris paribus racial pay gap across jobs is larger the greater is customer prejudice* $(\dfrac{\partial MDC_{NW_2}^{W_1}}{\partial D}$ *and* $\dfrac{\partial MDC_{NW_1}^{W_2}}{\partial D} < 0)$;

(ii) *If white workers in one job become more productive, discrimination against nonwhite workers in the other job increases.*

This prediction is confirmed by:

$$(9)\ \frac{\partial MDC_{NW_2}^{W_1}}{\partial \gamma_{11}} = \frac{1}{D[\gamma_{44} + \gamma_{24}(\frac{X_1^{NW}}{X_2^{NW}})^{\frac{1}{2}}] + D^{\frac{1}{2}}[\gamma_{14}(\frac{X_1^{W}}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{34}(\frac{X_2^{W}}{X_2^{NW}})^{\frac{1}{2}}]} -$$

$$\frac{1}{[\gamma_{44} + \gamma_{24}(\frac{X_1^{NW}}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{14}(\frac{X_1^{W}}{X_2^{NW}})^{\frac{1}{2}} + \gamma_{34}(\frac{X_2^{W}}{X_2^{NW}})^{\frac{1}{2}}]} > 0$$

and

$$(10)\ \frac{\partial MDC_{NW_1}^{W_2}}{\partial \gamma_{33}} = \frac{1}{D[\gamma_{22} + \gamma_{24}(\frac{X_2^{NW}}{X_1^{NW}})^{\frac{1}{2}}] + D^{\frac{1}{2}}[\gamma_{12}(\frac{X_1^{W}}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{23}(\frac{X_2^{W}}{X_1^{NW}})^{\frac{1}{2}}]} -$$

$$\frac{1}{[\gamma_{22} + \gamma_{24}(\frac{X_2^{NW}}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{12}(\frac{X_1^{W}}{X_1^{NW}})^{\frac{1}{2}} + \gamma_{23}(\frac{X_2^{W}}{X_1^{NW}})^{\frac{1}{2}}]} > 0.$$

If there is an exogenous increase in the productivity of white workers doing job 1 (γ_{11} rises) or white workers doing job 2 (γ_{33} rises), white wages rise. According to equations (9) and (10), the white wage with prejudice (measured by the numerator in the left-hand ratio in equations (7) or (8)) will rise proportionately more than will the white wage in the absence of prejudice (measured by the numerator in the right-hand ratio in equations (7) and (8)). Regardless of the signs and magnitudes of the technology coefficients and the relative supplies of labor, discrimination rises. For example, a technological advance that makes white workers doing one job more efficient results in greater discrimination against nonwhite workers performing the other job;

(iii) *If nonwhite workers in a job become more productive, then they experience less discrimination* ($\dfrac{\partial MDC_{NW_2}^{W_1}}{\partial \gamma_{44}} < 0$ and $\dfrac{\partial MDC_{NW_1}^{W_2}}{\partial \gamma_{22}} < 0$).[5]

When nonwhite workers experience an increase in productivity, they benefit in two ways. First, their wage rises. Second, from equations (7) and (8), the productivity increase reduces the wage ratio with discrimination and without, but the left-hand wage ratio falls more than the right-hand ratio in either equation;

(iv) *Discrimination experienced by nonwhites in a job depends upon the racial compositions within and across both jobs.*

As an example, consider equation (7). The *ceteris paribus* racial pay gap between whites in job group 1 and nonwhites in job group 2 depends upon the racial composition of group 1 ($\dfrac{X_1^{NW}}{X_1^{W}}$), the racial composition of group 2 ($\dfrac{X_2^{W}}{X_2^{NW}}$), the supply of nonwhites in group 2 relative to the supply of whites in group 1 ($\dfrac{X_2^{NW}}{X_1^{W}}$), the supply of whites in group 2 relative to the supply of whites in group 1 ($\dfrac{X_2^{W}}{X_1^{W}}$), the

---

supply of nonwhites in group 1 relative to the supply of nonwhites in group 2 ($\frac{X_1^{NW}}{X_2^{NW}}$) and the supply of

whites in group 1 relative to the supply of nonwhites in group 2 ($\frac{X_1^{W}}{X_2^{NW}}$). Compare these results with

what would be predicted by the traditional model. In Becker (1971), an increase in the relative supply of

nonwhites results in a greater *ceteris paribus* pay differential between whites and nonwhites. This is the

simple result for an economy where whites and nonwhites are identical and there is effectively just one

job. In our model, the relationship between discrimination and the relative supply of nonwhite labor is

much more complicated. Not only does the amount of discrimination experienced by nonwhites in one

job depend upon how many nonwhites there are in that job, but also on how many nonwhites there are in

the other job. Furthermore, how dominant whites and nonwhites are in each job and how racially

integrated one job is relative to another will influence the level of discrimination experienced by

nonwhites in a job. This is all due to the production function, which dictates the interrelationships in

production between the different labor groups. Note that we cannot sign the relationship between the

MDC and any labor supply ratio without knowing the signs of the technology coefficients (γ_{ij});

(v) *Prejudice and productivity interact in the determination of racial pay differences across job groups; the marginal effect of prejudice on pay depends upon whether whites and nonwhites are substitutes or complements and on the magnitudes of the elasticities of substitution* ($\frac{\partial^2 MDC_{NW_2}^{W_1}}{\partial \gamma_{11} \partial D}, \frac{\partial^2 MDC_{NW_1}^{W_2}}{\partial \gamma_{22} \partial D} \neq 0$).[6]

This implication is important because it suggests that in an empirical specification of cross-job

discrimination, interaction terms between race and productivity must be included in order to avoid

estimation bias. As an example of how prejudice and productivity interact in our model, note from

equation (9) that the reduction in discrimination experienced by nonwhite workers in group 1 (relative to

white workers in group 2) as a result of a productivity increase will be lower the greater is the degree of

---

[6] Proofs available upon request from the authors.

customer prejudice ($\frac{\partial^2 MDC_{NW_2}^{W_1}}{\partial \gamma_{11} \partial D} < 0$). Prejudice taxes the benefit nonwhites enjoy from being more

productive and the tax is greater the greater is customer distaste for output made by nonwhite workers.

III.   A Test Case: Major League Baseball During the 1990s

III.1 *Description of the test case*

In this section, we test a number of the implications of the model presented above. We focus on an

industry where: (a) there are accurate data on salaries and productivity for individual workers across

distinct job assignments and these data are available for different firms; (b) the productivities of job

assignment groups within the firm are interrelated; (c) there is racial integration; and (d) there is

potential for customer discrimination. One industry conveniently satisfying these criteria is Major

League Baseball (MLB) in the USA.[7] In MLB, each team (firm) requires two distinctly complementary

types of player skill – hitting (an offensive skill) and pitching (a defensive skill) - in the production of

baseball entertainment. Pitchers have historically been disproportionately white, whereas the pool of

hitters has tended to be more racially balanced.

III. 2 *Empirical Analysis*

To ascertain the level of discrimination across positions we need to control for a player's MRP, that is,

for his contribution to the team's ticket, broadcasting and merchandise revenues. This is problematic on

two accounts: First, the team production nature of baseball makes it is impossible to directly measure

individual revenue contributions. We thus proxy MRP by various position-specific career statistics

(computed on a game-by-game basis since the beginning of the player's MLB career) that proxy ability

---

[7] Wage discrimination in professional sports has received considerable attention among labor economists because of the abundant statistical evidence on a player's personal attributes, compensation and productivity. Most of the studies have focused on racial discrimination with respect to pay, hiring, retention and positional segregation. For a relatively recent examination of the research in this area, see Kahn's [2000] expository survey.

and skills. This leads us to our second problem, since different career statistics are commonly recorded for hitters and pitchers. To be sure, the career statistics commonly recorded for hitters are: *At Bats, Stolen Bases, Bases on Balls, Total Bases, Slugging Average* and *Batting Average*. The statistics commonly recorded for pitchers are: *Wins, Losses, Games Started, Complete Games, Saves, Homeruns, Walks, Strikeouts, Innings Pitched, Earned Run Average (ERA)* and *Strikeout Rate.*[8]

To surmount this second hurdle, we adopt the following two-stage approach. We first assume that player productivity contains both an 'on-field' component and an 'off-field' component, and that these impact upon earnings as follows:

(11)
$$w_h = \mathbf{B}_0^h \mathbf{X}^h + \mathbf{B}_1^h \mathbf{Z}$$
$$w_p = \mathbf{B}_0^p \mathbf{X}^p + \mathbf{B}_1^p \mathbf{Z}$$

where w_h and w_p denote the salaries of hitters and pitchers respectively, $\mathbf{X}^i$, $i = h, p$, is a vector of (largely) position-specific 'on-field' productivity measures, $\mathbf{Z}$ is a vector of common 'off-field' productivity measures, and the $\mathbf{B}$'s denote parameter vectors. Our aim is to derive an estimating equation of the form:

(12) $$w_i = \mathbf{B}_0^i \mathbf{X} + \mathbf{B}_1^i \mathbf{Z}$$

where $i = h, p$ denotes hitters and pitchers and $\mathbf{X}$ denotes some common, 'imputed' measure of 'on-field' productivity. We therefore estimate the following 'first-stage' regressions:

(13)
$$w_h = \mathbf{A}_0^h \mathbf{X}^h$$
$$w_p = \mathbf{A}_0^p \mathbf{X}^p$$

---

[8] A discussion of these and other baseball terms referred to in the paper may be found at any one of a number of websites, including www.baseball1.com and www.mlb.com.

That is, we regress hitter and pitcher wages on only those variables that we presume affect a player's on-field productivity. We assume that $\mathbf{X}^h$ and $\mathbf{X}^p$ include the relevant position-specific statistics described above as well as a player's age, years of MLB experience and years of MLB-squared. We also include in the vector $\mathbf{X}^h$ dummy variables to identify those hitters who are 'designated hitters' as well as those who are 'infielders', 'outfielder' or 'catchers' when their team is fielding.

We interpret predicted earnings, $\hat{w} = \left(\hat{w}_h, \hat{w}_p \right)$, from regressions (13) as a position-independent, imputed measure of on-field productivity for both hitters and pitchers. We then estimate the following second-stage regression:

(14) $w_i = \beta_0^i \hat{w} + \mathbf{B}_1^i Z$

The measures of on-field productivity we incorporate into $\mathbf{Z}$ include age, MLB experience and experience-squared, tenure with current club, whether the player plays in the American or National League, whether he is a member of a Canadian team, whether he is a free agent and whether he is subject to final offer arbitration.[9] We also include characteristics of the Greater Metropolitan area in which the player's team is located (i.e. percentage of population that is white, black and Hispanic, average annual income and population).

Our empirical analysis is set out in Tables 1-4 (Appendix). Tables 1 and 2 present descriptive statistics for hitters and pitchers, respectively. Our full sample includes 1093 hitters (549 White, 367 Black and 177 Hispanic) and 1204 pitchers (942 White, 127 Black and 135 Hispanic). Salary, experience, performance and position data were drawn from the *Lahman Baseball Database* (www.baseball1.com) over four seasons - 1992, 1993, 1997 and 1998. The Major League expanded by

---

[9] In MLB, player salaries are set under two different regimes, one competitive, the other monopsonistic. The monopsonistic regime applies to players with fewer than six years of MLB experience. These players are subject to the *reserve clause* and are constrained to negotiate their pay with only one team. The competitive regime applies to players with at least 6 years of MLB experience. They are eligible to file for *free agency* and may negotiate with any team in the league. Monopsony power effectively begins to erode, however, as early as the fourth year because then a player is eligible for *final offer arbitration*.

two teams between 1992 and 1993 and again by two teams between 1997 and 1998. The salary data do not include information about contract length, bonus clauses or endorsements. Salaries for players on the Canadian teams were converted to U.S. dollars. The experience data were used to determine the player's eligibility for free agency and final offer arbitration and the player's race was inferred from inspection of *Topps* baseball cards for all four seasons. For the U.S. teams, metropolitan area population and per-capita income were obtained from the website of the Bureau of Economic Analysis (www.bea.gov). For the Canadian teams, similar data were obtained from the Statistics Canada website (www.statcan.ca). Per-capita income data for the Canadian cities were converted to U.S. dollars.

It is apparent from Table 1 that there are no major differences between the personal and professional characteristics of white, black and Hispanic hitters, nor in the characteristics of the greater metropolitan area in which they play. In terms of career characteristics, however, it is apparent that black hitters record significantly more *At Bats*, *Stolen Bases, Bases on Balls* and *Total Basses* than either white or Hispanic hitters. Whites record more *At Bats* and *Bases on Balls* but fewer *Stolen Bases* and *Total Bases* than Hispanics. Blacks are also significantly less likely than whites or Hispanics to play as an infielder, catcher or designated hitter. In Table 2, the domination of white pitchers is immediately apparent. It is evident that white pitchers are on average older than black and (especially) Hispanic pitchers, and that they also enjoy higher average earnings. In terms of career characteristics, white pitchers record significantly higher *Wins*, *Losses*, *Games Started*, *Complete Games*, *Shutouts*, *Saves*, *Homeruns*, *Walks*, *Strikeouts* and *Innings Pitched* than either blacks or Hispanics, with Hispanics recording generally lower figures than blacks.

Table 3 reports second-stage regressions with white pitchers, black pitchers, Hispanic pitchers, white hitters, black hitters and Hispanic hitters defined as the default race-position category respectively. It is apparent from Table 3 that there are significant racial differences in earnings both within and across

---

Arbitration rights tend to relieve players of monopsonistic exploitation because arbitrators strive to award competitive salaries.

positions. Our estimated coefficients suggest that even after controlling for both on- and off-field productivity, the median annual earnings of white pitchers are 15.1 percent higher those of black pitchers, 11.1 percent higher than those of black hitters and 8.8 percent higher than those of Hispanic hitters. There is also some evidence of reverse discrimination with the median annual earnings of white hitters being 16.1 percent lower than those of Hispanic pitchers, Finally, in terms of the two minority groups, the median annual earnings of Hispanic pitchers are 15.6 per cent higher than those of Black pitchers and 11.8 percent (at the 90 percent level) higher than Black hitters.

In Table 4 we explore our theoretical prior that wage discrimination across player job assignments interacts with productivity differences between white, black and Hispanic hitters and pitchers. We test this prediction by creating a *Relative Productivity* variable that equals the difference between a player's individual productivity and the mean productivity of players in the other racial/position group multiplied by the player's individual productivity. Thus, in Column (1) of Table 4, where we focus on white pitchers relative to black hitters, our *Relative Productivity* (*White Pitcher:Black Hitter*) variable is defined as: *Individual White Pitcher Productivity* x (*Individual White Pitcher Productivity - Mean Black Hitter Productivity*), where productivity is estimated according to the two-stage process outlined in equations (11)-(14).

It would appear from Table 4 that wage discrimination is indeed affected by changes in relative productivities. Our theoretical prior is that discrimination against non-whites in a particular job assignment should increase as whites in the other job assignment become more productive. Our empirical results suggest that, as regards white pitchers and black hitters and white pitchers and Hispanic hitters, the opposite is the case, with the earnings of white pitchers relative to non-white hitters falling as the relative productivity of white pitchers increases.

14

IV.   Concluding Remarks

In this study, we address a previously un-researched problem in the literature on taste discrimination in pay: Ascertaining the extent to which racial or gender differences in pay across job assignments are attributable to prejudice. Nearly all wage discrimination studies have focused on discrimination within the same job assignment, thus treating whites and nonwhites (or males and females) as perfect substitutes. We extend the theory to the case of discrimination across job assignments where assignments are viewed as distinct inputs. Our theoretical findings underscore the importance of carefully considering the production function when there are productivity differences between majority and minority workers. An important finding from our theoretical analysis is that the magnitude of racial productivity differences influences the amount of discrimination. Furthermore, when whites and nonwhites are interrelated in production, either within or across job assignments, race and productivity will interact. This is an important implication, for it means that *whenever* white and nonwhite workers have productivity differences, the researcher should include productivity x race interactions in any empirical specification.

We tested our model using data from Major League Baseball, an industry characterized by complementary job assignments, a history of racial integration and discrimination. We found convincing evidence of ceteris paribus racial differences in pay across player positions, even after controlling for a wide array of demographic variables and position-specific productivity. Moreover, we find strong evidence of our theoretical prior that racial pay differentials across assignments are affected by changes in relative productivities. A suggested next step in this research would be to ascertain the generalizability of these findings for other industries and forms of labor market structure.

References

Altonji, J.G. and R.M. Blank (1999). 'Race and Gender in the Labor Market.' In O.A. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Vol. 3C, Amsterdam: Elsevier, pp. 3143-259.

Arrow, K. J. (1973). 'The Theory of Discrimination.' In O.A. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*. Princeton NJ: Princeton University Press, pp. 3-33.

Becker, G. S. (1971). *The Economics of Discrimination, Second Edition.* Chicago: University of Chicago Press.

Bodvarsson, Ö. B. and M. D. Partridge. (2001). 'A Supply and Demand Model of Employer, Customer and Coworker Discrimination.' *Labor Economics*, 8, June, pp. 389-416.

Borjas, G. J. (1983). "The Substitutability of Black, Hispanic, and White Labor." *Economic Inquiry*, 21, January, pp. 93-106.

_____ (1987). "Immigrants, Minorities, and Labor Market Competition." *Industrial and Labor Relations Review*, 40 (3), April, pp. 382-92.

_____ (2008). *Labor Economics*, 4th edition. Dubuque, IA: McGraw-Hill Irwin.

Diewert, W. E. (1971). "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function." *Journal of Political Economy*, 79, May/June, pp. 481-507.

Grant, J.H. and D.S. Hamermesh (1981). "Labor Market Competition Among Youths, White Women and Others." *The Review of Economics and Statistics*, 63(3), August, pp. 354-60.

Grossman, J.B. (1982). "The Substitutability of Natives and Immigrants in Production." *The Review of Economics and Statistics*, 64(4), November, pp. 596-603.

Hashimoto, M. and L. Kochin (1980). "A bias in the statistical effects of discrimination." *Economic Inquiry*, 8, pp. 478-86.

Kahanec, M. (2006). "The Substitutability of Labor of Selected Ethnic Groups in the U.S. Labor Market." *IZA Discussion Paper No. 1945*, January (www.iza.org).

Kahn, L. M. (1991). 'Customer Discrimination and Affirmative Action.' *Economic Inquiry*, 29, July, pp. 555-571.

_____(2000). 'A Level Playing Field? Sports and Discrimination.' In William S. Kern (ed.), *The Economics of Sports*, W.E. Upjohn Institute, Kalamazoo, MI.

Welch, F. (1967). "Labor-Market Discrimination: An Interpretation of Income Differences in the Rural South." *Journal of Political Economy*, 75(3), June, pp. 225-40.

Appendix

Table 1: Descriptive Statistics: Hitters

Variable	All Mean	All Std. Dev	White Mean	White Std. Dev	Black Mean	Black Std. Dev	Hispanic Mean	Hispanic Std. Dev
Personal Characteristics								
Log Annual Salary	13.890	1.13	13.865	1.10	13.938	1.13	13.866	1.22
Age	30.304	3.70	30.596	3.49	30.488	3.95	29.023	3.55
White	0.502	0.500	-	-	-	-	-	-
Black	0.336	0.472	-	-	-	-	-	-
Hispanic	0.162	0.369	-	-	-	-	-	-
Professional Characteristics								
MLB Experience	7.061	3.89	7.062	3.87	7.223	4.07	6.723	3.55
MLB Experience-Squared	64.957	69.31	64.785	70.06	68.684	74.23	57.763	54.59
Tenure with Current Club	2.673	3.00	3.062	3.38	2.305	2.62	2.226	2.24
Free Agent	0.600	0.49	0.598	0.49	0.605	0.49	0.599	0.49
Eligible for Final Offer Arbitration	0.296	0.46	0.304	0.46	0.294	0.46	0.271	0.45
American League	7.061	3.89	0.521	0.50	0.469	0.50	0.588	0.49
National League	0.486	0.50	0.479	0.50	0.057	0.23	0.124	0.33
Canadian Team	0.073	0.26	0.067	0.25	7.223	4.07	6.723	3.55
Performance								
At Bats	2506.414	2001.58	2419.738	1940.51	2699.202	2198.95	2375.525	1720.23
Stolen Bases	69.746	112.52	44.800	72.35	111.055	157.89	61.480	69.63
Bases on Balls	254.275	247.74	253.131	233.32	285.349	293.87	193.39	161.14
Total Bases	1060.200	913.52	1016.772	880.39	1162.845	1013.19	982.073	771.85
Slugging Average	0.407	0.06	0.404	0.06	0.416	0.06	0.397	0.07
Batting Average	0.267	0.03	0.264	0.02	0.271	0.02	0.266	0.02
Infielder	0.459	0.50	0.556	0.50	0.281	0.45	0.531	0.50
Outfielder	0.383	0.49	0.217	0.41	0.657	0.48	0.333	0.47
Catcher	0.116	0.32	0.189	0.39	0.016	0.13	0.096	0.30
Designated Hitter	0.059	0.24	0.046	0.21	0.079	0.27	0.056	0.23
Greater Metro Area Characteristics								
Percentage White	80.507	6.89	80.938	6.77	80.683	6.72	78.808	7.39
Percentage Black	13.273	6.58	12.959	6.60	13.676	6.62	13.409	6.44
Percentage Hispanic	10.621	10.65	10.719	10.80	10.331	10.58	10.918	10.36
Average Annual Income ($)	25562.990	3789.65	25508.570	3757.99	25551.300	3731.59	25756.00	4016.17
Population[1]	5514009	4657988	5313189	4509095	5513759	4729589	6137413	4927354
Year Dummies								
1992	0.250	0.43	0.255	0.44	0.243	0.43	0.249	0.43
1993	0.235	0.42	0.248	0.44	0.237	0.43	0.192	0.40
1997	0.260	0.44	0.248	0.43	0.270	0.44	0.277	0.45
1998	0.255	0.44	0.250	0.43	0.251	0.43	0.282	0.45
Sample Size	1093		549		367		177	

Note: 1. Population denotes the greater metro area population.
Source: All variables except Race and Greater Metro Area Characteristics (GMAC) extracted from the Lahman Baseball Database (Version 5.0, Release Date: Dec. 15, 2002). Race is derived form observed Topps Baseball Cards, years 92, 93, 94, 97, 99 (only years available). GMAC derived from the Statistical Abstract 1997-1999, the BEA, CA1-3, and from Statistical Canada..

Table 2: Descriptive Statistics: Pitchers

Variable	All Mean	Std. Dev	White Mean	Std. Dev	Black Mean	Std. Dev	Hispanic Mean	Std. Dev
Personal Characteristics								
Log Annual Salary	13.409	1.19	13.451	1.20	13.238	1.16	13.276	1.18
Age	29.815	4.09	30.190	4.02	29.016	4.00	27.948	4.03
White	0.782	0.41	-	-	-	-	-	-
Black	0.105	0.31	-	-	-	-	-	-
Hispanic	0.162	0.37	-	-	-	-	-	-
Professional Characteristics								
MLB Experience	5.988	4.20	6.158	4.20	5.772	4.49	5.000	3.75
MLB Experience-Squared	53.468	76.64	55.562	78.38	53.331	75.31	38.985	63.34
Tenure with Current Club	1.924	2.07	1.935	2.10	1.843	1.97	1.926	1.99
Free Agent	0.467	0.50	0.482	0.50	0.441	0.50	0.385	0.49
Eligible for Final Offer Arbitration	0.306	0.46	0.314	0.46	0.236	0.43	0.319	0.47
American League	0.513	0.50	0.518	0.50	0.543	0.50	0.452	0.50
National League	0.487	0.50	0.475	0.50	0.528	0.50	0.556	0.50
Canadian Team	0.069	0.25	0.063	0.24	0.055	0.23	0.126	0.33
Performance								
Starter	0.442	0.50	0.441	0.50	0.402	0.49	0.489	0.50
Wins	37.446	44.33	39.007	45.27	34.386	42.41	29.430	38.34
Losses	34.179	37.05	35.904	38.37	29.236	30.11	26.785	32.12
Games Started	74.12	105.53	77.769	108.53	59.646	92.16	62.274	93.98
Complete Games	10.15	22.24	10.981	23.33	6.433	14.87	7.844	19.65
Shutouts	2.875	6.08	3.065	6.32	1.984	4.74	2.385	5.35
Saves	19.488	51.87	20.941	52.93	19.362	62.60	9.474	26.16
Homeruns	56.517	62.57	58.842	64.46	50.409	52.94	46.044	56.11
Walks	225.779	249.73	231.782	257.66	224.095	217.58	185.474	217.41
Strikeouts	436.641	514.13	450.726	530.21	436.047	490.18	338.919	402.35
Innings Pitched	627.59	702.43	655.160	720.78	558.969	620.14	499.785	627.21
ERA	4.025	0.96	3.995	0.94	4.175	1.11	4.094	0.97
Strikeout Rate	0.078	0.02	0.078	0.02	0.083	0.02	0.079	0.02
Greater Metro Area Characteristics								
Percentage White	80.714	6.84	80.695	6.91	80.335	6.56	81.201	6.59
Percentage Black	13.038	6.46	12.946	6.49	14.026	6.46	12.750	6.19
Percentage Hispanic	10.975	10.77	10.899	10.61	10.909	10.40	11.573	12.20
Average Annual Income ($)	25488.2	3939.85	25491.51	3895.30	25852.23	3898.44	25122.19	4271.98
Population[1]	5551948	4683875	5481401	4631793	6035905	4915887	5588930	4829139
Year Dummies								
1992	0.221	0.42	0.236	0.42	0.189	0.39	0.148	0.36
1993	0.239	0.43	.248	0.43	0.244	0.43	0.170	0.38
1997	0.264	0.44	.256	0.44	0.276	0.45	0.311	0.46
1998	0.276	0.45	.260	0.44	0.291	0.46	0.370	0.48
Sample Size	1204		942		127		135	

Note: 1. Population denotes the greater metro area population.
Source: All variables except Race and Greater Metro Area Characteristics (GMAC) extracted from the Lahman Baseball Database (Version 5.0, Release Date: Dec. 15, 2002). Race is derived form observed Topps Baseball Cards, years 92, 93, 94, 97, 99 (only years available). GMAC derived from the Statistical Abstract 1997-1999, the BEA, CA1-3, and from Statistical Canada

Table 3: Discrimination Controlling for Position Specific Productivity

	(1) All Default – White Pitcher		(2) All Default - Black Pitcher		(3) All Default – Hispanic Pitcher		(4) All Default – White Hitter		(5) All Default - Black Hitter		(6) All Default - Hispanic Hitter	
	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat
Imputed Productivity	0.863	34.04	0.863	34.04	0.863	34.04	0.863	34.04	0.863	34.04	0.863	34.04
Race Dummies												
White Pitcher	-	-	0.164	2.69	-0.005	-0.09	0.170	4.97	0.106	2.69	0.092	1.79
Black Pitcher	-0.164	-2.69	-	-	-0.170	-2.15	0.006	0.09	-0.058	-0.87	-0.072	-0.98
Hispanic Pitcher	0.005	0.09	0.170	2.15	-	-	0.176	2.81	0.112	1.70	0.097	1.35
White Hitter	-0.170	-4.97	-0.006	-0.09	-0.176	-2.81	-	-	-0.064	-1.52	-0.078	-1.46
Black Hitter	-0.106	-2.69	0.058	0.87	-0.112	-1.70	0.064	1.52	-	-	-0.014	-0.25
Hispanic Hitter	-0.092	-1.79	0.072	0.98	-0.005	-0.09	0.078	1.46	0.014	0.25	-	-
Professional Characteristics												
Age	-0.024	-3.27	-0.024	-3.27	-0.024	-3.27	-0.024	-3.27	-0.024	-3.27	-0.024	-3.27
MLB Experience	0.152	3.45	0.152	3.45	0.152	3.45	0.152	3.45	0.152	3.45	0.152	3.45
MLB Experience-Squared	-0.010	-4.79	-0.010	-4.79	-0.010	-4.79	-0.010	-4.79	-0.010	-4.79	-0.010	-4.79
Tenure	0.056	10.00	0.056	10.00	0.056	10.00	0.056	10.00	0.056	10.00	0.056	10.00
Free Agent	0.879	6.14	0.879	6.14	0.879	6.14	0.879	6.14	0.879	6.14	0.879	6.14
Final Offer Arbitration	0.471	5.94	0.471	5.94	0.471	5.94	0.471	5.94	0.471	5.94	0.471	5.94
American League	-0.006	-0.23	-0.006	-0.23	-0.006	-0.23	-0.006	-0.23	-0.006	-0.23	-0.006	-0.23
Canadian	-0.022	-0.21	-0.022	-0.21	-0.022	-0.21	-0.022	-0.21	-0.022	-0.21	-0.022	-0.21
Greater Metro Area Characteristics												
Percent White	0.001	0.34	0.001	0.34	0.001	0.34	0.001	0.34	0.001	0.34	0.001	0.34
Percent Black	0.005	1.24	0.005	1.24	0.005	1.24	0.005	1.24	0.005	1.24	0.005	1.24
Percent Hispanic	0.005	3.39	0.005	3.39	0.005	3.39	0.005	3.39	0.005	3.39	0.005	3.39
Average Annual Income	0.000	1.45	0.000	1.45	0.000	1.45	0.000	1.45	0.000	1.45	0.000	1.45
Population	0.000	0.11	0.000	0.11	0.000	0.11	0.000	0.11	0.000	0.11	0.000	0.11
Year Dummies												
1993	0.051	1.31	0.051	1.31	0.051	1.31	0.051	1.31	0.051	1.31	0.051	1.31
1997	0.046	0.97	0.046	0.97	0.046	0.97	0.046	0.97	0.046	0.97	0.046	0.97
1998	0.130	2.44	0.130	2.44	0.130	2.44	0.130	2.44	0.130	2.44	0.130	2.44
Constant	0.994	1.63	0.830	1.37	0.100	1.65	0.824	1.34	0.888	1.44	0.902	1.48
R-Squared	0.7356		0.7356		0.7356		0.7356		0.7356		0.7356	
F-Statistic	420.83 $_{22, 2273}$		420.83 $_{22, 2273}$		420.83 $_{22, 2273}$		420.83 $_{22, 2273}$		420.83 $_{22, 2273}$		420.83 $_{22, 2273}$	
Root Mean Squared Error	0.61302		0.61302		0.61302		0.61302		0.61302		0.61302	
Observations	2296		2296		2296		2296		2296		2296	

Table 4: Discrimination Controlling for Position Specific Productivity and Relative Productivity

	(1) White Pitchers / Black Hitters		(2) White Pitchers / Hispanic Hitters		(3) Black Pitchers / White Hitters		(4) Black Pitchers / Hispanic Hitters		(1) Hispanic Pitchers / White Hitters		(2) Hispanic Pitchers / Black Hitters	
	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat	Coef	T-Stat
Imputed Productivity	0.894	18.94	0.998	18.13	0.882	17.03	0.974	13.87	0.874	17.32	0.960	16.49
Race Dummies												
White Pitcher	0.108	2.64	0.102	1.90								
Black Pitcher					-0.061	-0.92	-0.125	-1.68				
Hispanic Pitcher									0.240	2.60	0.201	1.98
Relative Productivity												
White Pitcher: Black Hitter	-0.003	-0.90										
White Pitcher: Hispanic Hitter			-0.013	-2.99								
Black Pitcher: White Hitter					-0.018	-3.85						
Black Pitcher: Hispanic Hitter							-0.018	-3.49				
Hispanic Pitcher: White Hitter									0.007	0.95		
Hispanic Pitcher: Black Hitter											0.003	0.50
Constant	0.804	-0.84	-0.786	-0.79	0.047	0.04	-3.335	2.03	0.596	0.53	-1.285	-0.96
R-Squared	0.7478		0.7727		0.7194		0.7837		0.7203		0.7346	
F-Statistic	322.06 $_{19, 1288}$		314.29 $_{19, 1098}$		127.73 $_{19, 656}$		99.30 $_{19, 284}$		132.23 $_{19, 664}$		105.17 $_{19, 482}$	
Root Mean Squared Error	0. .60507		0.59381		0.61113		0.59296		0.61166		0.61851	
Observations	1308		1118		676		304		684		502	

Notes: 1. Other explanatory regressors were those set out in Table 3; 2. 'Relative Productivity' is defined as, e.g., 'White Pitcher: Black Hitter' = Individual White Pitcher Productivity x (Individual White Pitcher Productivity - Mean Black Hitter Productivity).