Latinos in the Major Leagues from 1920 to 1939: A Human Capital Approach

Tom Cullen
University of Northern Iowa

Follow this and additional works at: https://scholarworks.uni.edu/mtie

Part of the Economics Commons

Let us know how access to this document benefits you

Copyright ©2015 by Major Themes in Economics

Recommended Citation
Available at: https://scholarworks.uni.edu/mtie/vol17/iss1/3

This Article is brought to you for free and open access by the CBA Journals at UNI ScholarWorks. It has been accepted for inclusion in Major Themes in Economics by an authorized editor of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.
Latinos in the Major Leagues from 1920 to 1939: 
A Human Capital Approach

Tom Cullen*

ABSTRACT. Economists frequently study the determinants of baseball players' wages. Researchers usually control for measures of productivity such as the player's batting average, home runs, and runs batted in, education, age, and experience. This study differs from standard studies in that it considers players' country of origin in the 1920s and 1930s, years before Jackie Robinson's debut. It uses information on player productivity from Baseball-Reference.com, player salary information from the Baseball Hall of Fame and Museum (graciously shared by its compiler), and information on college attendance provided by Who's Who in Baseball (held at the Joyce Sports Collection, Notre Dame University). It concludes that although there is documented evidence of racial discrimination against Latinos, the discrimination probably did not take the form of lower salaries.

I. Introduction

“Who is the greatest manager, really, Luque or Mike Gonzalez?-I think they are equal” (Hemingway 1952, 5). Hemingway, a resident of Cuba at the time he wrote this passage in The Old Man and the Sea, had first-hand experience with baseball in a Latin country. He even built a sandlot for aspiring local players and his children (Robinson 2008). Baseball in Latin America and the Caribbean markedly increased in popularity over thirty years before Hemingway's residence. Havana had a professional circuit before the U.S. National League was formed (Bjarkman 2015). The two managers Hemingway referenced, Dolf Luque and Mike Gonzalez, entered Major League Baseball in 1914 and flourished. Luque led the Major Leagues in earned run average and wins in 1923; Gonzalez was the first Latin player to record more than one home run in a season (Baseball-Reference 2014). Luque and Gonzalez were not alone. Players from Cuba, Mexico, and Venezuela were the largest foreign cohort in the Major Leagues from 1920 to 1939 (Baseball-Reference 2014).

Luque and Gonzalez's entrance to the majors marks a departure from Major League policy. Major League Baseball was segregated by

*The author would like to thank Dr. Michael Haupert for salary data. Studies of 20th Century American baseball would not be possible without his database.
gentlemen's agreement until Jackie Robinson and Larry Doby stepped on Major League diamonds in 1947 (Baseball-Reference 2014). Major League Baseball's commissioner, Kenesaw Landis, quashed attempts to integrate baseball during his tenure, which ended in 1944 (Pietrusza 2001, XV). Players from Latin America, Hawaii, and Europe, however, were allowed to play if they were light-skinned (Baseball-Reference 2014). Dark-skinned Cuban players like Pepe Abreu were relegated to the Negro Leagues, but Gonzalez and Luque were eligible for Major League rosters.

It is debatable, however, whether light-skinned players were better off. Several Cuban players had to be reassigned to the minor leagues because pitchers would throw at them rather than over home plate. Latin players faced death threats, as did the scouts that would bring them (Considine 1940, 19-24). The reason for the Latin influx into Major League Baseball is arbitrage. Joe Cambria, head of scouting for Washington, thought Cuba was the next source for a “$25 a week” star. Bob Considine, a journalist for Collier's Weekly, concluded that Latin players in Major League Baseball received “rock bottom pay” despite being “extraordinarily expert at the game” (Considine 1940, 19).

This study examines whether Latin players faced salary discrimination as Considine claimed. Brown, Gabriel, and Surdam provide a model of what determines player salaries. They contend that performance, age, position, and experience determine a player's salary, and cite Becker and Mincer's seminal works (Brown et. al. 2012, 449). This study attempts to link salaries of non-pitchers to individual statistics, college attendance, position, experience, age, and team as Brown et. al. did. In addition, it will control for country of origin.

II. Literature Review

Kevin Christiano's (1988) study warrants review on multiple levels. First, he established a method to test race discrimination in Major League Baseball. If salary discrimination were present, a comparable minority player would be paid less than a white counterpart. He found no evidence of discrimination for players negotiating contracts in 1987. He concluded that free agency, the ability of a player to switch teams after contract expiration, mitigated salary discrimination. Baseball before 1976, the Reserve Clause era, allowed teams to discriminate (Christiano 1988, 136-138). The Reserve Clause was a clause in player contracts that allowed Major League franchises to unilaterally retain a player's services
Cullen: Latinos in the Major Leagues

(Rottenberg 1956, 246). If Major League Baseball used a salary competition regime, discrimination probably would not exist.

[T]o discriminate by race is to risk the loss of an entire class of workers. It is to pass up all the benefits that workers of the undesired racial category might bring. Even worse, discrimination leaves these benefits to one's competitors. Thus, according to theory, with competition come powerful disincentives to the practice of racism (Christiano 1988, 138).

During the Reserve Clause era, Major League franchises did not use salary competition. The Reserve Clause allowed franchises to exercise monopsony power over their respective players. Moreover, it allowed franchises to select players of desired races (Christiano 1988, 138). Christiano acknowledged the possibility of discrimination the Reserve Clause era.

Considine's (1940) article in Collier's Weekly provides first-hand evidence of discrimination. Considine claimed that Latin foreign players were “so naïve and loved the game so much that they would sign up for absolutely nothing” (i.e. they would accept very low salaries). Considine claimed the migration of Latin players from Cuba, Venezuela, and Mexico was the experiment of an ambitious scout, Joe Cambria, for the Washington Senators. Cambria viewed Latin players as a cheap alternative to native talent, deeming them “future Joe DiMaggio(s) for $25 a week” (Considine 1940, 19). Considine concluded that Latin players faced lower pay and poor treatment:

They get beanballs thrown at their heads by closed-shop (and closedbrained) rivals. They face pitchers who willingly will throw away their arms bearing down on them in an effort to escape the "ignominy" of yielding a hit to them. They get a measure of grasssingingeing abuse from the "jockeys" on the enemy bench. From their own team they get rock-bottom pay, and from many of their own teammates they get a wintry ostracism. Those who want to befriend them are halted by the differences in languages (Considine 1940, 19).

The model presented below will attempt to test Considine's claims of discrimination in the form of lower salaries.
Determining salaries requires estimating human capital. Human capital estimation requires a review of Jacob Mincer's seminal work. Mincer laid out an empirical specification for wage determination: 

$$\ln E_s = \ln E_0 + rS + \beta_1t - \beta_2t^2$$

where \( \ln E_s \) is the full earnings capacity for a worker in logarithmic form, \( \ln E_0 \) is the earnings capacity for a worker with no experience and no schooling, \( S \) is the number of years of schooling, and \( t \) is the number of years of experience (Mincer 1974, 84). Mincer used a log-linear functional form because income naturally has a skewed distribution (Mincer 1974, 24-25).

Although Becker's approach was similar to Mincer's, Becker's use of age is integral in this study. Becker's age-earnings profile links age, schooling, and earnings over the typical life cycle. He found that earnings rise with age at a decreasing rate (Becker 1964, 228-233).

Lee Douthit's interview is a pillar supporting the theoretical framework behind this study. 1920 to 1939 had a marked increase in the number of players who attended college (Drowne and Huber 2004, 32). Douthit, a University of California graduate, told Baseball Magazine that college attendance had a positive effect on prospective players entering the major leagues.

[A college player] … enjoys other advantages. His strict training at the hands of a college coach has taught him the necessity of keeping in physical condition. The glamour of big cities means less to him than to the country boy who has never been away from home. He is more likely to view baseball as a profession. In short, he is further advanced than the average sandlot ball player (Douthit 1929, 448).

Douthit claimed college players earn higher salaries because of superior development; he warned, however, about skepticism from fellow players and owners (Douthit 1929, 448).

Clifford Harvey, of Baseball Magazine, affirmed Douthit's claims. Harvey noted that 100 former college graduates were in the Major League Baseball in 1929, 18 of which made the “tremendous jump” to the major leagues without minor league training (Harvey 1930, 297). He concluded that college would increase the chances of a player 'making good' since college made “natural ball player[s]” (Harvey 1930, 297).

Alexandre Olbrecht provided a link between education and player salaries using human capital estimation. He found that college players...
who negotiated contracts in 2005 were paid an 8.3% premium; however, his result was not statistically significant (Olbrecht 2007, 981-985). He provided a theory as to why college players receive a premium on contracts:

[H]uman capital theory suggests that educational investments will make a worker more productive, raise his marginal product and eventually his lifetime earnings. By attending college, a player might develop into a more rounded individual who is more cognizant of team goals and committed to sacrificing personal statistics for wins. Teams cognizant of this theory may be willing to pay wage premiums for the attributes described. Ultimately individuals deciding between college and the minor leagues are more interested in which track will yield the largest lifetime income stream. A documented and significantly large earnings premium could serve to offset the opportunity costs of schooling and be relevant information to players deciding between the minors and college (Olbrecht 2007, 982).

Olbrecht's reasoning restates Douthit and Harvey's claims about human capital. While he applied it to modern players, Olbrecht's use of college attendance as a proximate variable for educational attainment is the avenue to verify Douthit and Harvey's claims.

Scully attempted several times (in 1974, 1989, and 1995) to determine whether Reserve Clause players were paid their marginal contributions. He linked franchise revenues to win percentage, which is determined by player productivity (Scully 1989, 154). He applied Mincer's model to salaries during the Reserve Clause era and concluded the Reserve Clause depressed Major League salaries.

Brown, Gabriel, and Surdam applied Mincer's study among others to determine player salaries for the New York Yankees from 1919-1941 (Brown et al. 2012, 449). Although they used panel estimation, more sophisticated statistics, and examined only Yankees players, they offered a framework to employ offensive statistics in human capital estimation. The following model uses three offensive metrics: home runs, batting average, and runs batted in.

Simon Rottenberg, in his evaluation of the Reserve Clause, pointed to franchises inflating salaries to dominate the talent pool. The absence of an amateur draft left rookies free to negotiate with franchises
(Baseball-Reference 2014). Rottenberg contended that the largest population centers with large revenue bases would attract the most capable players and would erode competitive balance. Rottenberg inferred that the Reserve Clause was intended to mitigate this effect, but concluded that the talent distribution was inequitable (Rottenberg 1956, 247). Moreover, the New York market (the Yankees and the Giants) dominated the talent pool during the era using their large revenue bases to offer higher salaries (Rottenberg 1956, 247). Players who played for either the Giants or the Yankees received a salary premium, according to Rottenberg.

Glenn, McGarrity and Weller also examined player position's effect on tenure and salary. They found catchers and players who played shortstop or second base (middle infielders) were traded less frequently because of increased firm-specific human capital. In other words, because of catchers' and middle infielders' defensive prowess, they stay with franchises longer. A possible result could be higher pay (Glenn et. al. 2001, 92). Holding offensive productivity fixed, catchers and middle infielders could be paid a premium.

III. Model

This human capital model is based on the premise that a player's pay is determined by offensive statistics and other observable characteristics such as age, position, and experience. A traditional approach of measuring human capital would be to measure a player's marginal revenue product, or the additional revenue generated from an additional unit of an input (Rottenberg 1956, 916). The Reserve Clause allowed Major League franchises to fix salaries of players at no less than 75% of the current year's salary (Rottenberg 1956, 246). Players' only recourse was not to play. There is a strong presumption that the Reserve Clause allowed franchises to exercise monopsony power and depress salaries. If teams had complete freedom to fix salaries, the question is whether teams set salaries with a standard method. Surdam, Brown, and Gabriel found that Reserve Clause players for the New York Yankees were paid on a systematic, productivity-related basis (Brown et. al. 2012, 449). Players were paid according to productivity, although they did not receive their marginal revenue product (Scully 1995, 43).

Scully, Olbrecht, and Brown et. al. all used batting average and home
runs to study players of the period. Runs batted in were not included in other studies, but has explanatory power. One concern about adding runs batted in is multicollinearity. But while each home run constitutes a run batted in, runs batted in are not perfectly correlated with home runs. Players like Paul Waner ranked in the bottom quartile of home runs, but ranked in the top 50 of runs batted in. Also, runs scored are directly related to wins. Wins increase attendance, which boosts revenue. Higher revenue, as was the case with the New York Yankees and Giants, translates to higher salaries (Blass 1992, 261).

Applying measures of productivity is the first step. The next step is education, as is common in human capital estimations. The effect of college attendance on salaries has been answered to a degree. Olbrecht found that college players from 2005 received an 8.3% premium on contracts. This finding is promising, but does not preclude further investigation. As Douthit pointed out, owners and teammates were leery of integrating college players (Douthit, 449). College players may have possessed stronger bargaining leverage and could find lucrative employment elsewhere. The Reserve Clause could have encouraged college players to use this option.

The next variable is player position. The importance of player position, or the defensive assignment of a player, has a mixed history in the literature. Christiano, Glenn et. al., Brown et. al, and Blass used outfield as a variable in salary determination (Brown 454); (Blass 2001, 265); (Christiano, 139-140). Olbrecht and Scully did not include the variable. This study includes middle infielders and catchers as Glenn et. al. did. As stated above, catchers and middle infielders are traded less often and have higher firm-specific human capital (Glenn et. al. 2001, 92).

Experience, age, and age-squared appear in virtually every human capital salary model. Mincer pointed out that age increases the explanatory value of earnings analyses, but concluded experience was a better fit (Mincer 1974, 83-84). Becker included age and experience to determine earnings in his seminal work (Becker 1964, 228-233). Scully pointed out that experience is one of the four factors that are crucial to salary determination (Scully 1989, 156). Every study cited in this study uses some combination of age and experience to estimate salaries.

This study applies the above measures to the human capital framework developed by Becker and Mincer. According to the framework, a worker is paid according to his productivity, which is linked
to the amount of a worker's human capital. With Haupert's database of
salaries for the period, a link can be estimated between a player's human
capital and salary. The empirical specification of salaries is as follows:

\[ \ln(S) = \beta_0 + \beta_1 \text{COLLEGE} + \beta_2 \text{CATCHER} + \beta_3 \text{MIDDLE INFIELD} \\
+ \beta_4 \text{EXPERIENCE} + \beta_5 \text{NEW YORK} + \beta_6 \text{BATTING AVERAGE} \\
+ \beta_7 \text{RBI} + \beta_8 \text{HOME RUN} + \beta_9 \text{AGE} + \beta_{10} \text{AGE}^2 + \beta_{11} \text{LATIN} + \epsilon \quad (1) \]

The dependent variable is the natural log of a player's salary for a
corresponding season. This is standard in the literature for salary
determination in the human capital framework (Mincer 1974, 84). The
key independent variable, LATIN, denotes a player's country of origin in
binary form; 1 for those born in Cuba, Mexico, or Venezuela and 0 for
those born elsewhere. COLLEGE, is a binary variable that is a 1 for
players who attended college and 0 for those who did not. CATCHER and
MIDDLE INFIELD are also binary variables that assign a 1 value for
those who played those positions and 0 for those who did not.

**Table 1—Expected Variable Coefficients, and Variable Definitions**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Description</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOG REAL SALARY</td>
<td>Salary Adjusted for Inflation in Logarithmic Form</td>
<td>Dependent</td>
</tr>
<tr>
<td>LATIN</td>
<td>Players of Latin Origin</td>
<td>-</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>Players Who Attended College</td>
<td>+</td>
</tr>
<tr>
<td>CATCHER</td>
<td>Players Who Were Primarily Catchers</td>
<td>+</td>
</tr>
<tr>
<td>MIDDLE INFIELD</td>
<td>Players Who Played Primarily Infield</td>
<td>+</td>
</tr>
<tr>
<td>NEW YORK</td>
<td>Played in New York Market</td>
<td>+</td>
</tr>
<tr>
<td>EXPERIENCE</td>
<td>Years of Experience in Major League Baseball</td>
<td>+</td>
</tr>
<tr>
<td>BATTING AVERAGE</td>
<td>Hits/At Bats; Minimum 100 at Bats</td>
<td>+</td>
</tr>
<tr>
<td>RBI</td>
<td>Number of RBI’s</td>
<td>+</td>
</tr>
<tr>
<td>HOME RUN</td>
<td>Number of Home Runs</td>
<td>+</td>
</tr>
<tr>
<td>AGE</td>
<td>Player’s Age at Start of Season</td>
<td>+</td>
</tr>
<tr>
<td>AGE-SQUARED</td>
<td>Player’s Age Squared</td>
<td>-</td>
</tr>
</tbody>
</table>

1. Position dummies were determined whether the player played the majority of games
   at a position.
2. Whether the player played for either the New York Yankees or the New York Giants
   for any part of a given season
Holding offensive performance fixed, these players were more important defensively. NEW YORK is another binary variable that denotes a player's presence in the New York market. BATTING AVERAGE, RBI, and HOME RUN are statistics that document a player's offensive productivity. BATTING AVERAGE is the number of hits divided by the number of at-bats. RBI and HOME RUN are statistics that record the number of runs batted in and home runs for a given season. Age and age-squared are discrete variables that document a player's age at the beginning of a given season.

The model above was estimated with panel regression with robust standard errors to correct for clustering. Adjusting for clustering is intended to correct for multiple observations on one player.

IV. Data

Michael Haupert, of the University of Wisconsin LaCrosse, generously shared his database of nominal salaries from 1920 to 1939. Haupert assembled his database from contract information held at the National Baseball Hall of Fame and Museum. Several adjustments had to be made to conform to standard human capital estimation. Observations needed to be adjusted for inflation or deleted if offensive production statistics were not available.

Mincer's second variable, education, had to be transformed into college attendance. Mincer's model used years of schooling rather than college attendance. College attendance is a proxy since it is the closest, verifiable representation of schooling. College records were assembled from the Who's Who in Major League Baseball from 1920-1939 at University of Notre Dame's Rare and Special Collections.

As stated above, offensive productivity is preferable in human capital estimation for baseball. An all-encompassing statistic such as WARP (Wins Above Replacement Player) would be ideal since it captures a player’s relative value (Baseball-Reference 2014). But sophisticated measures like WARP were not available to management from 1920 to 1939. This study uses traditional measures of productivity that were available to management. These data were available through Baseball-Reference.com and were merged with Haupert's salary database. Players with less than 100 at-bats were excluded, as were all pitchers. Players with few at-bats skew regression coefficients, particularly, offensive statistics (Brown et. al., 456), and pitchers are judged on a different set of statistics.

**Table 2–Descriptive Statistics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Description</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>REAL SALARY</td>
<td>Salary Adjusted for Inflation</td>
<td>$70,093.46</td>
<td>$895.52</td>
<td>$6,620.86</td>
<td>$5,369.63</td>
</tr>
<tr>
<td>LATIN*</td>
<td>Players of Latin Origin</td>
<td>1</td>
<td>0</td>
<td>0.013</td>
<td>0.114</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>Players Who Attended College</td>
<td>1</td>
<td>0</td>
<td>0.332</td>
<td>0.471</td>
</tr>
<tr>
<td>CATCHER**</td>
<td>Players Who Were Primarily** Catchers</td>
<td>1</td>
<td>0</td>
<td>0.161</td>
<td>0.367</td>
</tr>
<tr>
<td>MIDDLE INFIELD</td>
<td>Players Who Played Primarily** 2nd Base or Shortstop</td>
<td>1</td>
<td>0</td>
<td>0.268</td>
<td>0.443</td>
</tr>
<tr>
<td>NEW YORK</td>
<td>Played in New York Market***</td>
<td>1</td>
<td>0</td>
<td>0.193</td>
<td>0.395</td>
</tr>
<tr>
<td>EXPERIENCE</td>
<td>Years of Experience in the Majors</td>
<td>23</td>
<td>0</td>
<td>5.974</td>
<td>4.191</td>
</tr>
<tr>
<td>BATTING AVERAGE</td>
<td>Hits/At Bats; Minimum 100 At Bats</td>
<td>0.424</td>
<td>0.067</td>
<td>0.290</td>
<td>0.039</td>
</tr>
<tr>
<td>RBI</td>
<td>Number of RBI’s</td>
<td>185</td>
<td>0</td>
<td>55.607</td>
<td>32.713</td>
</tr>
<tr>
<td>HOME RUN</td>
<td>Number of Home Runs</td>
<td>60</td>
<td>0</td>
<td>6.177</td>
<td>8.148</td>
</tr>
<tr>
<td>AGE</td>
<td>Player’s Age at Start of Season</td>
<td>41</td>
<td>18</td>
<td>28.547</td>
<td>4.072</td>
</tr>
<tr>
<td>AGE-SQUARED</td>
<td>Player’s Age Squared</td>
<td>1681</td>
<td>324</td>
<td>831.530</td>
<td>237.400</td>
</tr>
</tbody>
</table>

* Player was born in Cuba, Mexico, or Venezuela
** Position dummies were determined whether the player played the majority of games at a position.
*** Whether the player played the year for either the New York Yankees or the New York Giants
V. Results/Discussion

Table 3–OLS Regression Results: Hitters From 1920-1939:
Dependent Variable: ln(Salary)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>4.333***</td>
<td>0.598</td>
<td>0.000</td>
</tr>
<tr>
<td>LATIN</td>
<td>0.162</td>
<td>0.155</td>
<td>0.295</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>0.080***</td>
<td>0.036</td>
<td>0.026</td>
</tr>
<tr>
<td>CATCHER</td>
<td>0.075</td>
<td>0.050</td>
<td>0.139</td>
</tr>
<tr>
<td>MIDDLE INFIELD</td>
<td>0.115***</td>
<td>0.034</td>
<td>0.001</td>
</tr>
<tr>
<td>EXPERIENCE</td>
<td>0.070***</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>NEW YORK</td>
<td>0.134***</td>
<td>0.040</td>
<td>0.001</td>
</tr>
<tr>
<td>BATTING AVERAGE</td>
<td>0.462</td>
<td>0.366</td>
<td>0.207</td>
</tr>
<tr>
<td>RBI</td>
<td>0.004***</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>HOME RUN</td>
<td>0.015***</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>AGE</td>
<td>0.227***</td>
<td>0.044</td>
<td>0.000</td>
</tr>
<tr>
<td>AGE-SQUARED</td>
<td>-0.004***</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td>2358</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R-Squared</td>
<td>0.513</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Statistic (11, 2346)</td>
<td>200.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Denotes significance at p<.10
** Denotes significance at p<.05
*** Denotes significance at p<.01

All variables except for LATIN, CATCHER, and BATTING AVERAGE were statistically significant at the p<.05 level. R-Squared
and Adjusted R-Squared were .515 and .513, respectively. Approximately 51% of variation of salaries from 1920 to 1939 has been explained by this model. The regression could be improved if we had information on other variables such as eyesight, work ethic, and speed. Unfortunately, data on these things are either unavailable or would raise a problem of multicollinearity (Scully 1989, 156).

The main variable of interest, LATIN, was not statistically significant. A possible reason could be the limited number of observations. As stated above, 39 Latin players were on Major League rosters from 1920 to 1939. The 100-at bat constraint limited the number of Latin players in the regression to ten. Many Latin players were shuffled out of line-ups or summarily dismissed before they could reach one hundred at-bats. Alternatively, it could mean that discrimination against Latin players did not take the form of lower salaries.

An additional home run would garner a 1.5% pay raise. An additional run batted in (RBI) would garner a .4% pay raise. The fear, as stated above, is that the interrelation of performance variables like HOME RUN and RBI would add little explanatory power and present a problem of multicollinearity. It appears the size of Haupert's database mitigated this effect.

BATTING AVERAGE appears to have the highest magnitude of every variable. Even though BATTING AVERAGE is not statistically significant, it still has some explanatory power. Not including BATTING AVERAGE in this regression would have sizably decreased R-Squared and Adjusted R-Squared. The BATTING AVERAGE coefficient, however, has no practical significance with its current interpretation. The coefficient on BATTING AVERAGE shows a 1 point increase in batting average yields a 46% premium on a given salary. But a one point increase in batting average is impossible. A batting average has never and probably will never exceed .5 for a player with over 100 at-bats. The maximum batting average in Haupert's database is Rogers Hornsby, who hit .424 in 1924. Dividing the coefficient by 1000 would adjust the coefficient to the correct change in batting average: .001. If Hornsby hit .434, or ten points higher, he would garner a .46% salary premium.

COLLEGE was positive and statistically significant. A college player was paid 8% more than a non-college player with identical productivity. This is consistent with the claims of Harvey, Douthit, and Olbrecht. Further, the COLLEGE coefficient is the same magnitude as Olbrecht's finding. College attendance appears to be rewarded at similar rates across
decades.

Age and experience variables were all statistically significant as Brown et al., Fort, Scully, and Olbrecht found. AGE and EXPERIENCE were positive, but AGE-SQUARED was negative. This indicates that salary rises as a player ages, but the salary rises at a decreasing rate. This is consistent with Mincer's work and the age-earnings profile (Becker 1964, 230-235).

As Rottenberg predicted, players in the New York market were paid 13% more than players from other teams. The maximum in Haupert's database is Babe Ruth's salaries from 1922 through 1933. New York players like Edd Roush, Lou Gehrig, Robert O'Farrell, and Frank Frisch all ranked in the top 70 salaries (Baseball-Reference 2014). This estimation is consistent with Rottenberg's assertion.

VII. Conclusions and Further Research

The coefficient on LATIN appears to dispel the notion that salary discrimination was present in the 1920's and 1930's. Although there is documented evidence of racial discrimination against Latinos, the discrimination probably did not take the form of lower salaries as the LATIN coefficient is indistinguishable from zero. Most of Cambria's Cuban recruits did not appear in this estimation because of the 100 at-bat constraint. Most Latin players were shuffled in and out of line-ups, never reaching 100 at-bats in a season. The only players included in the regression were the players that “made good.” Latin players who did not appear to be the next “$25 a week” star were summarily dismissed (Considine 1940, 19).

The results are consistent with human capital theory. Every variable besides LATIN, CATCHER, and BATTING AVERAGE yielded the expected sign and was statistically significant. College players were paid 8% higher than their non-college counterparts, the equivalent of hitting an extra six home runs. It also appears that players were paid on a systematic, productivity-related basis (Brown et. al. 2012, 459).

Further research could use a different choice of performance variables. Career slugging percentage, according to Scully, has more explanatory power than batting average. “Career slugging average is a better measure than that of the previous season because hitting fluctuates from season to season. A career average contains more information about ability than performance in a single year” (Scully 1989, 156).
The most documented evidence of discrimination was discrimination from fans (Revel and Munoz 2009, 2). Further research could focus on revenues for the teams that employed Latin talent. Teams like the Washington Senators were known as havens for Latin talent (Considine 1940, 19). Those teams could have experienced lower attendance and lower gate receipts as a result.

References


