CUSTOMER DISCRIMINATION IN MAJOR LEAGUE BASEBALL

by

Torben Andersen and Sumner J. La Croix*

Working Paper No. 89-14
July 14, 1989

* Department of Humanities and Social Science, Red Deer College, Red Deer, Alberta and Department of Economics, University of Hawaii, Honolulu, Hawaii, respectively. Helpful comments were received from Richard Hartman, Levis Kochin, Shelly Lundberg, Albert Black, Jr., and Keith Leffler. We are, of course, responsible for all remaining errors.

Racial discrimination in labor markets can occur whenever some employees, employers, or customers have preferences for discrimination against certain racial groups. Although Gary Becker noted these three sources of discrimination in his 1957 classic, *The Economics of Discrimination*, most researchers in this area have concentrated on identifying the existence rather than the source of discrimination in labor markets.[1] Why have researchers not attempted to identify the group with the discriminatory preferences in cases where labor market discrimination has been found to exist? One reason may be that the three types of discrimination often have implications which are observationally equivalent. For instance, discrimination by members of any one of the three groups will lead to lower wages being paid to the racial group that is the object of discriminatory preferences.

On the other hand, some implications of discrimination by the different groups are not identical. For example, if the firm's customers discriminate, discriminatory practices in hiring, compensation and promotion will increase firm profits. But if the discriminatory preferences originate with the firm's owner, those practices will reduce profits. Unfortunately, the difficulty in obtaining and interpreting data on firms' economic profits renders this approach (and many similar ones) virtually impossible to implement. Thus, the main problem seems to lie in empirically distinguishing, in a practical way, between the implications of the different sources of discrimination.

A number of researchers have examined major league baseball player salaries for evidence of racial discrimination. On the whole the findings are somewhat mixed, and in any case give little indication of the source of whatever discrimination has been found to exist. Also, results of these studies appear to be quite sensitive to the specification of functional forms, to sample sizes, and to the variables included in the regressions as control variables. Several authors have devised ingenious tests to determine whether baseball fans discriminate against black (or non-white) baseball players. Although some researchers have found significant discrimination by fans, others have been unable to detect any such discrimination.[2]

This paper attempts to test for discrimination by fans through examination of the market for a product closely related to major league baseball: bubble gum cards depicting baseball players. Prices paid by fans for baseball cards reflect the different performance records of the ballplayers depicted as well as other demand variables. Because the relevant performance variables differ completely between pitchers and non-pitchers, this paper's analysis is restricted to non-pitchers, also referred to as hitters. Unless fans have discriminatory preferences, a player's race should not be a significant determinant of the relative price of his baseball card. This test for discrimination is cleaner than others reported in the literature, as it eliminates the possibility that discrimination by baseball team owners and/or baseball players themselves will contaminate the test results.[3]

It is reasonable to assume that the preferences of baseball card collectors regarding players' racial origins should be quite representative of the larger group of all baseball fans (of whom collectors are a subset).[4] If so, we can, by examining customer discrimination in the baseball card market, draw inferences concerning customer discrimination in the market for viewing major league baseball games.

I. The Relative Price Structure of Baseball Cards

Baseball cards appeared at least as early as 1887 when they were used as inserts in cigarette packages. Cards were primarily used with tobacco products until about 1908 when caramel sellers also began to use them extensively. The cigarette issues dwindled and ultimately ceased before World War I. By 1933 several companies included baseball cards in bubble gum packages. From 1956 to 1980 the Topps Chewing Gum Company held a monopoly in the market for major nationally issued baseball card sets. Following an antitrust action, the monopoly was broken in 1981, and each year since then there have been at least two companies producing major national sets in competition with Topps.[5]

The card prices used in this study are resale market prices, since cards are not individually priced in the primary retail market.[6] The price of one ball player's baseball card relative to another player's card is obviously determined by the supply of and demand for the cards. A baseball card is a durable good yielding a stream of consumption services over time. Following Lancaster and Becker's approach to demand theory, a card viewer gets satisfaction not from the card per se, but from the attributes embodied in the card.[7] While the necessary information to estimate a complete Lancastrian consumer model does not exist, a hedonic price equation representing a reduced-form of the card supply and demand equations can be specified. Demand for a given card in period t, D_t, is determined by its price in period t, P_t, the (constant) per period discount rate, r, and the present and future levels of the relevant hedonic attributes. Let X_{it} represent the vector of attributes 1 through n in time periods 1 through z. Then we can write the demand function:

$$(1) \quad D_t = f(X_{11}, X_{21}, \ldots, X_{n1}, X_{12}, X_{22}, \ldots X_{n2}, \ldots X_{1z}, X_{2z}, \ldots, X_{nz}, P_t, r)$$

where $dD_t/dP_t < 0$ and $dD_t/dr < 0$.

A key objective of a hedonic model is to estimate the shadow prices of the utility-generating attributes. But hedonic theory does not indicate *a priori* what attributes will generate satisfaction for the consumer. From our experience in the card market and study of hobby periodicals, we assume that the attributes of a card which generate utility are: 1) physical condition or grade; 2) design of the card; 3) quality of the photograph; 4) popularity of the player; 5) scarcity of the card; and 6) whether it is a rookie card. We control for variation in physical condition by examining only prices for cards in the "mint" condition grade. Within a set, design of the card and quality of the photograph tend to be relatively constant and difficult to measure. In the analysis below we concentrate on potential determinants of a player's popularity.

The more popular a player is, the more demand there will be for his card. Fundamentally, popularity seems a function of proven productivity, potential productivity, fan familiarity and, possibly, race. Operationalizing these variables allows specification of a testable regression model.

A. Productivity

In baseball, perhaps more than in any other sport, selected individual performance statistics are widely regarded to reflect the quality of a player's performance with tolerable accuracy.[8] There is no doubt that on the whole players with statistics indicating high performance (i.e., productivity) are more popular with fans. They are simply judged better ballplayers. There is a good deal of literature on the issue of the theoretically ideal measure of baseball player productivity.[9] But since we are actually trying to model how baseball fans assess productivity, the issue of a theoretical ideal is less important in this study than in some others. Hence, we seek a comprehensive measure of productivity that reflects such dimensions of player performance as fans seem to find important. The measure should also be as independent as possible of the performance of teammates as well as the size and nature of the player's home ballpark.

Three measures of hitters' productivity were considered in our study. Batting average (BA) simply measures the frequency of a base hit per time at bat. This productivity measure is flawed, as it does not consider the power of the hitter or walks. Slugging average (SA) is more comprehensive as it incorporates both frequency and power of base hits, but SA does not measure the effects of walks and stolen bases on productivity. Offensive average (OA) measures power and base hit frequency and also incorporates the effects of walks and stolen bases on productivity.[10] Following Bennett and Flueck we employ the OA measure:[11]

> We recommend that major league baseball adopt OA as its relative evaluator of player offensive performance because of its simplicity in calculation and its strong relationship to run production. Of all estimators examined in this study OA provides the least disruption to traditional baseball practice, since it is only a slight modification of slugging percentage.

One defect of OA (which is shared by SA and BA) is that it totally omits defensive performance. This performance component is difficult to measure quantitatively and we have not included such measures in our econometric analysis. This is unlikely to constitute a major defect in our analysis, as the difficulty of constructing good defensive statistics means that defensive performance may play at most a minor role in determining a player's popularity for statistics oriented baseball fans.

Popularity, and hence card demand, may be nonlinearly related to productivity if additional productivity draws increasing increments of attention to the player. Given the particular attention paid to superstars by the mass media and the public, we include quadratic measures of productivity (OASQ) in our regression specifications.

Because baseball fans are very concerned with cumulative statistics and achievement, a player's popularity reflects his lifetime performance in the major leagues, not merely his most recent season.[12] A star player will continue to be very popular, and his card to hold its high value, even after his career has peaked. Accordingly, our model will include both lifetime OA and times at bat (ATBATS), a measure of career duration.

B. Potential Productivity

Just as the price of a share of common stock reflects the expected future performance of the firm, the price of a baseball card reflects the expected future performance of the player pictured. A young player who has hit many home runs early in his career will be more popular than a player who has accomplished the same feats but at a relatively late age. It will be perceived that the younger player is more likely to last long enough in the major leagues to, for example, break or challenge Hank Aaron's all-time home run record or hit enough home runs to ultimately be selected to the Hall of Fame. In anticipation of the future popularity (and high baseball card prices) that such achievements would bring, card buyers speculate and drive up current prices of promising young stars.

How, operationally, is potential productivity determined? First, past or proven productivity (OA discussed above) is an obvious predictor. Second, the player's expected longevity in the majors is related to his current age (AGE). Comparing two players with identical productivity, the younger player has greater potential to achieve the statistics needed to reach the Hall of Fame than does an older player. Potential productivity should be negatively related to AGE. Third, regardless of the player's age, if he is already retired, his potential productivity is zero. Therefore we introduce RETIRED, a dummy variable (1 if retired, 0 if active).

C. Fan Familiarity

The demand for a player's card depends on how familiar he is to baseball fans. Given two players with identical performance statistics and other characteristics, the more familiar player's card will be in higher demand. More productive players and more experienced players will be more familiar, but these variables are already captured in the productivity variables, OA and ATBATS.

Retired players not only lack potential future productivity (as discussed above), but also may be less familiar to fans if it has been a number of years since they last played baseball in the major leagues. To adjust for this effect we specify a variable (YRET) which is equal to the number of years since retirement.[13] A negative coefficient on this variable is expected.

Some retired players gain recognition by continuing in baseball as major league team managers. The number of years experience as a major league manager (MGR) may positively influence fan familiarity and thereby card demand. Given the rarity of black and latin managers, if MGR is a significant variable, then omitting it would upwardly bias estimates of race discrimination.

Certain characteristics of his team will also influence how well a player is known to fans. Players on teams located in very large cities may become well known to many more people than players in small cities in part due to the relatively heavy national media attention (including game telecasts) that teams in the largest cities receive. Fans also tend to prefer hometown players. For both of these reasons demand for a player's card should be positively related to the SMSA population (POP) of his team's home city. The more successful his team, the more exposure a player gets. Players on pennant contending teams are the subjects of more newspaper and magazine articles and are featured on more national network broadcasts than players on noncontending teams. The extreme form of this phenomenon occurs when a team reaches the Championship Series in its league and then the World Series. We measure team success by two variables: the number of games a team finishes behind the division winner (GBL), and the number of post-season games in which a player appears (POST). Because a player's familiarity to fans depends on his career history rather than just recent seasons, POP and GBL are measured as career averages and POST and MGR as career totals.

Additional popularity measures include LEAGUE (1 for American League players, 0 for National League players), and BOTHLGS (1 if player has played 10% of his career in a second league, 0 otherwise).

D. Rookie Card Status

Many card collectors consider some rookie cards to be particularly valuable. A rookie card is a card from the first year the player appeared in a major league baseball card set (which may not be his first year in the majors). For instance, a very promising young player's 1988 rookie card may have a market value several times as high as a 1988 card of a proven superstar, while at the same time the young player's 1989 card may have a value less than half that of the superstar's 1989 card. This price pattern cannot be explained by the demand determinants discussed above.[14] We specify a dummy variable (ROOKIE) for rookie cards, and since the price increment associated with rookie cards is clearly dependent on player productivity, also add a variable (ROOKOA) measuring this interaction between rookie status and productivity. A positive coefficient is expected on ROOKOA.

E. Playing Position

Among the different playing positions, substantially different offensive productivity expectations are held by team managers and by fans. This is because few players are equally adept at hitting and fielding, and because some positions are far more demanding in the defensive dimension. We specify dummy variables for CATCHER, FIRST, SECOND, THIRD, and SHORT; outfield is the omitted variable.

F. Race

Last, if baseball fans prefer to watch white ballplayers, they presumably also prefer to buy baseball cards of white players. Such a taste for discrimination raises the relative price of the cards of white players relative to nonwhites, *ceteris paribus*. Race is measured with two dummy variables, BLACK and LATIN. Latin players are defined as players born in Mexico, Central America, South America, or the Caribbean. Of the remaining players, North Americans, we determined race by inspecting baseball cards and other photographs. Especially useful was the magazine, *Ebony*, which annually featured a collection of photographs of blacks in the major leagues. Except for fringe players or midseason additions who might have been omitted from the annual feature, we considered *Ebony* to be authoritative.

II. The Econometric Analysis

This section presents regression results for the model of baseball card prices developed in Section I. Three samples of Topps baseball cards are analyzed. A common feature of the data is that within each set of cards there is a "common card" price which represents the minimum observed value in the sample. Since many of the cards in any given set bear the minimum price, the card price distribution is censored. We use Tobit analysis to ensure that regression estimates are unbiased.[15]

In order to test the general reliability of our model, and to see whether the degree of fan discrimination may differ for players from different eras, we selected a sample from the early 1960's and a sample from the late 1970's. The first sample is constructed by taking cards of hitters from the 1960 and 1961 Topps sets and evaluating them at 1982 prices. We began with the 1960 set because earlier sets contain too few nonwhite players to allow for meaningful analysis. However, since there are two common card prices within the 1960 set, presumably because of differences in quantities supplied for different "series" within the set, only part of the 1960 set is included in the sample. This limitation exists because with the Tobit regression technique we used, one limit value must be specified for the dependent variable. But since in 1982 those cards numbered 1-440 from the 1960 set and cards 1-522 from the 1961 set had the same common card price, we enlarge the sample by using cards from both sequences. Since the two sets are in adjacent years, we consider the players to be from the same cohort. In the case of a player who appears in both sets, only his 1960 card is used as an observation.

The second sample consists of hitters cards from 1977 evaluated at 1985 prices. This set is unique in that an equal number was produced of each card in this set. The 660 cards were printed as five sheets of 132, so each card appeared on only one sheet, and only once on that sheet. Therefore, supply variation by the manufacturer is not responsible for any price variations within the set. In addition, the 1977 Topps set does not have any significant competitors, while in the 1980s, several national sets of cards compete each year for consumer attention.

The third sample consists of "star" hitters chosen from various years card sets, all evaluated at 1985 prices. A disadvantage of both the 1960-61 and 1977 samples is that many of the observations are of common cards, so that a large proportion of the dependent variable values are at the limit price in the Tobit analysis. In effect, there may be too little price variation within each sample to permit racial discrimination present to be detected, especially if it is small. A sample of star players was constructed in order to obtain prices above common card prices. Selectivity bias would obviously emerge if cards were selected on the basis of price alone. The sample was constructed by excluding all players with less than 1000 career base hits in the major leagues by 1987. This criterion combines both longevity and performance and does not seem to racially bias the selection of players for the sample. Using nine Topps cards sets from between 1967 and 1983, 347 of the 409 players with 1000 base hits who were active between 1952 and 1987 are included in the sample. Virtually all of the excluded players ended their careers in the early 1950s and consequently were not active in 1967, the first card set in our sample.[16]

Regressions are estimated using all variables specified in Section I. Insignificant variables are then eliminated if their exclusion does not significantly reduce the overall fit of the equation (as measured by a likelihood ratio test) and if their exclusion does not significantly alter the coefficients for any race variables. The resulting "core" equations are reported when we analyze each of the three samples below. Unlike OLS regression coefficients, Tobit regression coefficients cannot be interpreted simply as the partial derivatives of the dependent variable with respect to the independent variables. A Tobit function is inherently nonlinear: the impact on the dependent variable of a given change in one independent variable is a function of the values of all independent variables at that point. For certain key variables we report transformed Tobit coefficients which represent partial derivatives of the expected card price functions evaluated at the sample means for all regressors in the equations (see Table 10).

A. <u>The 1960-61 Hitters Sample</u>

The means and standard deviations for variables in the 1960-61 sample are exhibited in Table 1. Significance (t-tests) are performed to determine whether the means for blacks and latins differ from corresponding means for whites.[17] The 1960-61 sample consists of 308 hitters: 246 whites, 36 blacks, and 26 latins. Blacks have more experience than whites and are more productive than whites whether productivity is measured by OA, SA, or BA. Performance levels of latins are not significantly different from whites other than that latins have higher batting averages and more career at bats. Black hitters are more likely to be outfielders than whites are. The low frequency of blacks and latins reflects the slow pace of racial integration in major league baseball throughout the late 1940s and 1950s. The slow pace of integration may also account for part of the difference in performance since in the early days of integration, the blacks joining the major-leagues were undoubtedly the best black players.

Tobit regressions on linear and log prices are reported in Table 2; while (for the linear model) the transformed coefficients that reflect partial derivatives are reported in Table 10.[18] Dummy variables (HIGH and SET61) added to adjust for possible differences in demand and supply between the 1960 and 1961 cards were not significantly different from zero.[19] In this sample card price is related to OA in the predicted exponential fashion. The coefficient on the linear term OA is negative, and the coefficient on the quadratic term OASQ is positive. Both are significant at the 1% level. When the derivative dP/dOA is evaluated for the linear price equation at the OA sample mean of 0.448, it is

TABLE 10

PREDICTED PRICE IMPACT OF
CHANGES IN SELECTED INDEPENDENT VARIABLES
IN LINEAR MODELS

SAMPLE	1960-61	1977	1977	1977	STAR	STAR	STAR	STAR
DEPENDENT	P82	P85	P85	P85	NORMP	NORMP	NORMP	NORMP
ATBATS	0.004**	0.004**	0.004**	0.004**	0.001**	0.001**	0.001**	0.001**
OA	-1028**	87.2*	89.0	-485.8**	-237.4**	-237.7**	-298.7**	-370.0**
OASQ	1195**	1.08	-10.2	621.0**	284.5**	285.0**	352.5**	425.7**
BLACK	-5.40	-3.20*	3.03	-243.0**	-0.990	0.253	11.88*	-29.06
BLACKAB	--	--	-0.001	--	--	-0.2E-3	--	--
BLACKOA	--	--	--	1004**	--	--	-25.4**	138.0
BLACKOASQ	--	--	--	-1043**	--	--	--	-161.5
LATIN	2.48	-3.48	8.10	-242.9	-0.456	0.056	-3.47	-177.7**
LATINAB	--	--	-0.002**	--	--	-0.1E-3	--	--
LATINOA	--	--	--	1103	--	--	--	-742.9**
LATINOASQ	--	--	--	-1247	--	--	--	-768.1**
F(Z) - $P_{min}*f(Z)/\sigma$	0.059	0.102	0.088	0.108	0.454	0.453	0.455	0.449

TABLE 1

SUMMARY STATISTICS BY RACE
1960-61 HITTERS -- 1982 PRICES

VARIABLE MEANS	WHITES n = 246	BLACKS n = 36	LATINS n = 26
P82	60.1	178.1	68.3
	(215.4)	(399.1)	(163.1)
ATBATS	3061	4756 **	4389 **
	(2455)	(3393)	(2863)
OA	0.443	0.487 **	0.442
	(0.065)	(0.059)	(0.050)
SA	0.379	0.421 **	0.383
	(0.062)	(0.061)	(0.049)
BA	0.250	0.269 **	0.265 **
	(0.028)	(0.021)	(0.024)
POST	4.96	9.19	6.62
	(10.97)	(13.05)	(8.07)
GBL	19.3	15.9 **	18.4
	(6.80)	(6.49)	(8.30)
LEAGUE	0.552	0.394 *	0.415
	(0.417)	(0.424)	(0.413)
POP	4004	4054	3659
	(1798)	(2025)	(1484)
MGR	0.66	0.11 **	0.00 **
	(2.27)	(0.67)	(0.00)
YRET	16.4	12.9 **	12.3 **
	(3.9)	(5.5)	(5.5)
AGE	50.9	49.0 *	48.5 **
	(4.3)	(4.4)	(3.9)
ROOKIE	0.150	0.222	0.269
	(0.358)	(0.421)	(0.452)
OUTFIELD	0.333	0.583 **	0.385
	(0.472)	(0.500)	(0.496)

Standard deviations are in parentheses below means.
Significantly different from WHITES at 0.01 level = **
at 0.05 level = *

TABLE 2

TOBIT CARD PRICE REGRESSIONS:
1960-61 HITTERS -- 1982 PRICES -- CORE MODELS

DEPENDENT	P82		LOGP82	
INDEPENDENT VARIABLES	REGRESSION COEFFICIENT	T-RATIO	REGRESSION COEFFICIENT	T-RATIO
BLACK	-91.59	-1.07	-0.160	-0.65
LATIN	42.10	0.43	0.201	0.72
ATBATS	0.067	4.15 **	0.3E-3	6.37 **
OA	-17418	-4.48 **	-42.42	-3.41 **
OASQ	20261	4.99 **	50.41	3.91 **
MGR	35.19	3.90 **	0.147	5.38 **
YRET	-16.07	-2.16 *	-0.031	-1.44
ROOKIE	-1911	-2.59 **	-5.13	-2.41 *
ROOKOA	4470	2.98 **	11.58	2.65 **
LEAGUE	76.35	1.17	0.143	0.77
POST	7.82	3.91 **	0.030	4.97 **
SET61	-148.5	-1.73	0.024	0.10
HIGH	-10.68	-0.08	0.135	0.37
CATCHER	68.67	0.76	0.207	0.80
FIRST	-18.39	-0.24	0.044	0.19
SECOND	-37.12	-0.37	-0.111	-0.38
SHORT	45.53	0.44	0.253	0.87
THIRD	-159.3	-1.38	-0.230	-0.74
CONSTANT	3263	3.44 **	9.98	3.22 **

	P82	LOGP82
P_{min} -	21.00	3.10
PREDICTED PROB $P > P_{min}$ GIVEN AVE X(I) -	0.070	0.142
OBSERVED FREQ. $P > P_{min}$ -	0.240	0.240
AT MEANS OF X(I), (P) -	29.24	3.109
SQ. CORRELATION BETWEEN OBS. AND EXP. VALUES -	0.761	0.833
STD ERROR OF ESTIMATE -	266.41	0.814
LOG-LIKELIHOOD FUNCTION -	-547.9	-133.3
MCFADDEN'S R SQUARED -	0.181	0.512

TOTAL OBSERVATIONS - 308 LIMIT - 234 NON-LIMIT - 74

ASYMPTOTIC T-RATIO SIGNIFICANCE LEVELS: 0.01 ** 0.05 *

found that an increase of 0.100 in OA will raise expected price by 4.34 cents. From an initial OA value of 0.500 the same 0.100 increase will raise expected price by 16.77 cents. On the other hand, at OA values slightly below the mean OA level, the parameters imply that OA increases will cause price to fall. Substituting SA or BA in place of OA as the measure of productivity changes the overall results little.

Experience has a powerful positive impact on price. An increase of 1000 ATBATS raises price by 3.95 cents in the linear equation.[20] Post-season games and managerial experience have significant and positive coefficients. The coefficient on years since retirement is significantly negative. The ROOKIE dummy variable is significantly negative in both equations, while the interactive variable ROOKOA is significantly positive. This suggests that cards of rookies are highly valued only when the player has a successful career. None of the coefficients on playing position dummies are significant. This suggests that card collectors place little emphasis on defensive contributions and that players' offensive contributions are evaluated in relation to all hitters rather than in relation only to hitters playing the same position.

The coefficient on BLACK in both equations is negative, but is insignificant with a t-ratio of -1.07 in the linear model and -0.65 in the log model. The coefficients on LATIN are positive but are also insignificant. Likelihood ratio tests indicate that the two race variables are not jointly significant. We also examine the possibility that race effects vary according to players' productivity and experience levels, as was found in an earlier salary study by Scully (1974a). Interactive variables between BLACK and LATIN with ATBATS or OA are added to the specification, but are not found to be either individually or jointly significant.[21] In sum, we cannot reject the null hypothesis that BLACK and LATIN have no effect on prices.

B. The 1977 Hitters Sample

This sample (see Table 3) consists of 345 hitters: 218 whites, 91 blacks, and 36 latins. Prices are significantly lower for latins than for whites. The blacks have significantly more experience than whites and are significantly more productive whether productivity is measured by OA, SA, or BA. Compared to white hitters, blacks and latins are significantly more likely to be outfielders. Very few of the differences between latins and whites are significant.

Tobit regressions for linear and log specifications of 1985 prices are reported in Table 4. In the linear model both OA and OASQ have positive coefficients, but neither is significant. In the log model the linear term is significantly positive and outweighs a small and insignificant negative coefficient on OASQ. Overall, the influence of productivity is positive as predicted. ATBATS is significantly positive in both regressions. An increase of 1000 ATBATS increases price by 3.97 cents in the linear model. The dummy variable specifying experience in both the American and National leagues is negative and significant. This result is unexpected, as players who have played in both leagues should be familiar to fans in a larger number of cities. YRET has a small positive, but insignificant coefficient. AGE decreases price, as expected, and is highly significant in both equations. Holding other variables constant, card price drops by 1.19 cents in 1982 for each additional year of age. Since the coefficient on age outweighs the YRET coefficient, the combined effect of these two variables remains negative for retired players, consistent not only with the potential productivity hypothesis, but also the hypothesis that familiarity begins eroding when a player retires. Like the 1960-61 sample, dummy variables for positions are insignificant.

TABLE 3

SUMMARY STATISTICS BY RACE
1977 HITTERS -- 1985 PRICES

VARIABLE MEANS	WHITES n = 218	BLACKS n = 91	LATINS n = 36
P85	30.2	36.2	12.6 **
	(71.9)	(63.3)	(7.6)
ATBATS	3492	4833 **	4019
	(2289)	(2231)	(2474)
OA	0.433	0.499 **	0.434
	(0.063)	(0.095)	(0.059)
SA	0.366	0.407 **	0.364
	(0.059)	(0.053)	(0.056)
BA	0.252	0.272 **	0.259
	(0.024)	(0.021)	(0.026)
POST	8.6	14.7 **	8.6
	(12.6)	(13.8)	(12.6)
GBL	15.0	12.8 **	15.3
	(6.1)	(5.7)	(5.8)
POP	3714	3548	3406
	(1833)	(1517)	(1250)
YRET	2.72	2.34	2.92
	(2.80)	(2.69)	(2.59)
AGE	36.9	37.3	37.6
	(3.6)	(3.9)	(4.1)
ROOKIE	0.101	0.055	0.083
	(0.302)	(0.229)	(0.280)
OUTFIELD	0.243	0.681 **	0.417 *
	(0.430)	(0.469)	(0.500)

Standard deviations are in parentheses below means.
Significantly different from WHITES at 0.01 level = **
at 0.05 level = *

TABLE 4

TOBIT CARD PRICE REGRESSIONS:
1977 HITTERS -- 1985 PRICES -- CORE MODELS

DEPENDENT	P85		LOGP85	
INDEPENDENT VARIABLES	REGRESSION COEFFICIENT	T-RATIO	REGRESSION COEFFICIENT	T-RATIO
BLACK	-31.35	-2.23 *	-0.236	-1.64
LATIN	-34.12	-1.65	-0.328	-1.55
ATBATS	0.039	8.57 **	0.5E-3	10.28 **
OA	855.1	0.85	28.12	2.35 *
OASQ	10.60	0.01	-16.58	-1.38
YRET	3.02	0.88	0.024	0.67
AGE	-11.69	-4.93 **	-0.145	-5.81 **
ROOKIE	278.4	1.70	2.15	1.06
ROOKOA	-457.8	-1.28	-2.32	-0.53
BOTHLGS	-31.27	-2.65 **	-0.324	-2.66 **
GBL	-0.877	-0.65	-0.025	-1.75
POST	1.15	2.13 *	0.007	1.27
CATCHER	32.91	1.73	0.613	3.10 **
FIRST	-4.97	-0.29	0.050	0.28
SECOND	-3.29	-0.14	-0.033	-0.14
SHORT	22.70	0.91	0.433	1.67
THIRD	-11.77	-0.60	-0.097	-0.48
CONSTANT	-167.5	-0.61	-3.96	-1.23

	P85	LOGP85
P_{min} =	9.00	2.20
PREDICTED PROB $P > P_{min}$ GIVEN AVE X(I) =	0.129	0.207
OBSERVED FREQ. $P > P_{min}$ =	0.386	0.386
AT MEANS OF X(I), E(P) =	13.44	2.286
SQ. CORRELATION BETWEEN OBS. AND EXP. VALUES =	0.641	0.798
STD ERROR OF ESTIMATE =	68.980	0.731
LOG-LIKELIHOOD FUNCTION =	-780.3	-186.8
MCFADDEN'S R SQUARED =	0.170	0.523

TOTAL OBSERVATIONS = 345 LIMIT = 212 NON-LIMIT = 133

ASYMPTOTIC T-RATIO SIGNIFICANCE LEVELS: 0.01 ** 0.05 *

The coefficient on BLACK is significantly negative in the linear model, and indicates that at the sample means a black player's card is valued 3.20 cents lower than a white's. BLACK is insignificantly negative in the log model where the t-ratio is -1.64. The coefficients on LATIN are also negative, but are insignificant in both specifications.[22] To examine whether race effects vary with productivity or experience, interactive variables between BLACK and LATIN with ATBATS and then with OA and OASQ are separately added to each regression as shown in Table 5.

In the linear equation LATINAB is significantly negative, and implies that for a 1000 ATBATS increase in experience a latin player's card would rise in price by 1.8 cents, compared to 3.9 for a white player's card. BLACKAB is negative but the t-ratio is only -1.80. In the log model both LATINAB and BLACKAB are negative but individually and jointly not significant at the 5% level.

In the equations augmented with interactive race-productivity variables, the BLACKOA and BLACKOASQ coefficients are significant while the LATINOA and LATINOASQ variables are insignificant at the 5% level. Results are reported in Table 6. Furthermore, with this respecification the OA and OASQ coefficients become significant and have the expected signs. Using the transformed Tobit coefficients shown in Table 10 we calculate expected card prices for white and black players for various OA levels. For OA equal to 0.500 (the sample OA mean for blacks is 0.499), the black/white price ratio is 0.80. There is discrimination against blacks at all OA levels, and the discrimination is greatest at higher levels of OA. This indicates that among the above average blacks, the better the player is, the more discrimination he will suffer. The general trends consistent with this estimate are illustrated in Figure 1. Latin

TABLE 5

TOBIT CARD PRICE REGRESSIONS:
1977 HITTERS -- 1985 PRICES -- RACE INTERACTED WITH ATBATS

DEPENDENT	P85		LOGP85	
INDEPENDENT VARIABLES	REGRESSION COEFFICIENT	T-RATIO	REGRESSION COEFFICIENT	T-RATIO
BLACK	34.48	0.90	0.040	0.10
LATIN	92.06	1.90	0.626	1.20
ATBATS	0.044	8.67 **	0.5E-3	9.89 **
BLACKAB	-0.011	-1.80	-0.5E-4	-0.72
LATINAB	-0.023	-2.75 **	-0.2E-3	-1.94
OA	1011	0.99	29.83	2.46 *
OASQ	-116.4	-0.11	-18.15	-1.49
YRET	1.29	0.38	0.013	0.36
AGE	-10.57	-4.50 **	-0.139	-5.55 **
ROOKIE	304.6	1.78	2.07	0.98
ROOKOA	-511.8	-1.38	-2.13	-0.47
BOTHLGS	-32.17	-2.76 **	-0.325	-2.69 **
GBL	-0.610	-0.47	-0.023	-1.67
POST	1.12	2.13 *	0.007	1.27
CATCHER	42.23	2.19 *	0.674	3.36 **
FIRST	1.21	0.07	0.095	0.54
SECOND	2.48	0.11	-0.006	-0.03
SHORT	32.30	1.30	0.509	1.95
THIRD	-10.44	-0.53	-0.080	-0.39
CONSTANT	-283.9	-1.00	-4.81	-1.48

	P85	LOGP85
P_{min} =	9.00	2.20
PREDICTED PROB $P > P_{min}$ GIVEN AVE X(I) =	0.114	0.198
OBSERVED FREQ. $P > P_{min}$ =	0.386	0.386
AT MEANS OF X(I), E(P) =	12.73	2.280
SQ. CORRELATION BETWEEN OBS. AND EXP. VALUES =	0.656	0.800
STD ERROR OF ESTIMATE =	67.275	0.723
LOG-LIKELIHOOD FUNCTION =	-776.1	-185.0
MCFADDEN'S R SQUARED =	0.175	0.527

TOTAL OBSERVATIONS = 345 LIMIT = 212 NON-LIMIT = 133

ASYMPTOTIC T-RATIO SIGNIFICANCE LEVELS: 0.01 ** 0.05 *

TABLE 6

TOBIT CARD PRICE REGRESSIONS:
1977 HITTERS -- 1985 PRICES -- RACE INTERACTED WITH OA AND OASQ

DEPENDENT	P85		LOGP85	
INDEPENDENT VARIABLES	REGRESSION COEFFICIENT	T-RATIO	REGRESSION COEFFICIENT	T-RATIO
BLACK	-2250	-2.93 **	-28.75	-4.15 **
LATIN	-2249	-1.32	-20.62	-1.71
ATBATS	0.038	8.59 **	0.5E-3	8.75 **
OA	-4498	-2.81 **	-23.04	-2.98 **
OASQ	5750	3.38 **	36.57	3.59 **
BLACKOA	9299	3.07 **	112.5	4.21 **
BLACKOASQ	-9658	-3.22 **	-110.3	-4.24 **
LATINOA	10211	1.41	92.6	1.73
LATINOASQ	-11546	-1.51	-103.9	-1.76
YRET	2.71	0.82	0.030	-0.20
AGE	-10.62	-4.53 **	-0.141	-6.30 **
ROOKIE	-178.0	-0.73	-3.76	-0.99
ROOKOA	565.2	1.06	10.90	1.17
BOTHLGS	-35.74	-3.17 **	-0.374	-3.19 **
GBL	-0.038	-0.03	-0.021	-2.43 *
POST	1.16	2.29 *	0.007	0.45
CATCHER	38.3	2.09 *	0.625	0.95
FIRST	-6.80	-0.42	-0.035	-0.25
SECOND	-0.672	-0.03	0.021	0.39
SHORT	9.51	0.40	0.234	0.94
THIRD	-13.8	-0.73	-0.094	-0.13
CONSTANT	1023	2.66 **	8.11	3.68 **

	P85	LOGP85
P_{min}	= 9.00	2.20
PREDICTED PROB $P > P_{min}$ GIVEN AVE X(I)	= 0.139	0.229
OBSERVED FREQ. $P > P_{min}$	= 0.386	0.386
AT MEANS OF X(I), E(P)	= 13.58	2.292
SQ. CORRELATION BETWEEN OBS. AND EXP. VALUES	= 0.681	0.815
STD ERROR OF ESTIMATE	= 64.977	0.689
LOG-LIKELIHOOD FUNCTION	= -771.6	-178.2
MCFADDEN'S R SQUARED	= 0.179	0.545

TOTAL OBSERVATIONS = 345 LIMIT = 212 NON-LIMIT = 133

ASYMPTOTIC T-RATIO SIGNIFICANCE LEVELS: 0.01 ** 0.05 *

FIGURE 1

Expected Prices vs. OA
Whites, Blacks, Latins--1977 Hitters

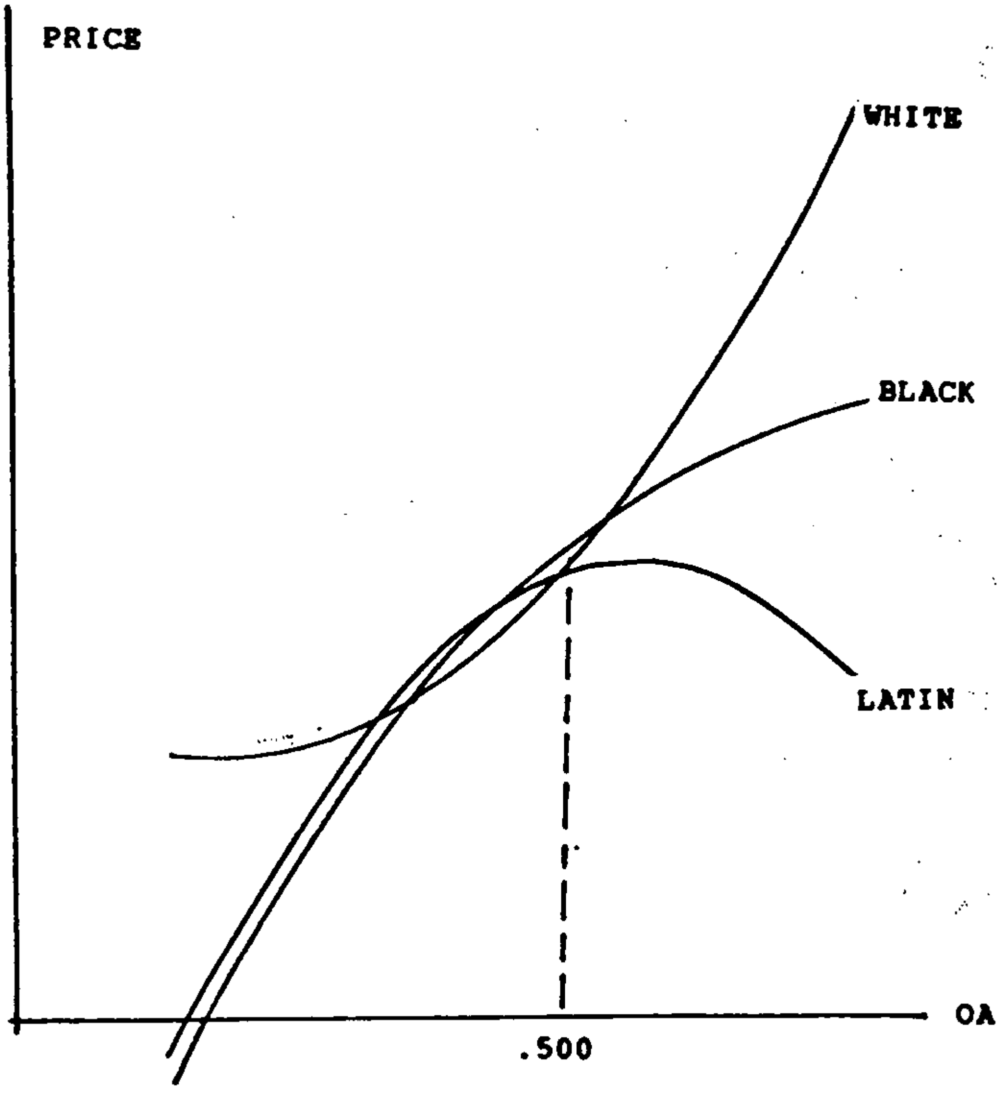

TABLE 7

SUMMARY STATISTICS BY RACE
STAR HITTERS -- 1985 PRICES

VARIABLE MEANS	WHITES n = 194	BLACKS n = 106	LATINS n = 47
NORMP	4.46	4.75	3.27
	(9.57)	(8.39)	(7.00)
ATBATS	5476	5536	5721
	(1749)	(2037)	(1801)
OA	0.477	0.507 **	0.457 *
	(0.056)	(0.049)	(0.053)
SA	0.410	0.427 **	0.392 *
	(0.054)	(0.052)	(0.051)
BA	0.269	0.278 **	0.275 *
	(0.017)	(0.015)	(0.021)
POST	14.1	14.9	11.5
	(15.8)	(13.0)	(10.4)
MGR	0.39	0.09 *	0.00 **
	(1.55)	(0.80)	(0.00)
RETIRED	0.63	0.48 *	0.68
	(0.48)	(0.50)	(0.47)
AGE	43.1	38.4 **	41.2
	(10.2)	(7.9)	(8.0)
OUTFIELD	0.330	0.700 **	0.447
	(0.471)	(0.473)	(0.503)

Standard deviations are in parentheses below means.
Significantly different from WHITES at 0.01 level = **
at 0.05 level = *

TABLE 8

TOBIT CARD PRICE REGRESSIONS:
STAR HITTERS -- RACE AND RACE INTERACTED WITH ATBATS

DEPENDENT	NORMP		NORMP	
INDEPENDENT VARIABLES	REGRESSION COEFFICIENT	T-RATIO	REGRESSION COEFFICIENT	T-RATIO
BLACK	-2.181	-1.94	0.559	0.19
LATIN	-1.005	-0.68	0.123	0.03
ATBATS	0.003	7.72 **	0.003	6.83 **
BLACKAB	---	---	-0.5E-3	-0.99
LATINAB	---	---	-0.2E-3	-0.27
OA	-522.9	-4.62 **	-524.4	-4.56 **
OASQ	626.7	5.52 **	628.9	5.45 **
MGR	-0.169	-0.50	-0.190	-0.56
RETIRED	-3.77	-1.24	-4.07	-1.33
AGE	-0.99	-5.44 **	-1.00	-5.47 **
POST	0.179	5.33 **	0.178	5.31 **
CATCHER	3.88	2.20 *	4.03	2.27 *
FIRST	-1.49	-1.11	-1.47	-1.09
SECOND	-0.09	-0.05	-0.042	-0.04
SHORT	0.804	0.43	0.843	0.84
THIRD	-1.94	-1.19	-1.99	-1.99 *
S60	23.95	4.80 **	24.47	4.86 **
S64	21.61	4.62 **	22.13	4.69 **
S68	23.25	5.68 **	23.80	5.72 **
S71	16.02	3.94 **	16.67	4.04 **
S74	8.29	2.34 *	8.80	2.44 *
S77	8.92	2.58 **	9.35	2.68 **
S80	3.77	1.12	4.26	1.25
CONSTANT	123.5	4.29 **	122.8	4.19 **

P_{min}	- 1.00	1.00
PREDICTED PROB $P > P_{min}$ GIVEN AVE X(I)	- 0.510	0.509
OBSERVED FREQ. $P > P_{min}$	- 0.640	0.640
AT MEANS OF X(I), E(P)	- 3.926	3.917
SQ. CORRELATION BETWEEN OBS. AND EXP. VALUES	- 0.662	0.661
STD ERROR OF ESTIMATE	- 7.107	7.101
LOG-LIKELIHOOD FUNCTION	- -798.5	-798.0
MCFADDEN'S R SQUARED	- 0.161	0.161
TOTAL OBSERVATIONS	- 347	347
LIMIT	- 125	125
NON-LIMIT	- 222	222

ASYMPTOTIC T-RATIO SIGNIFICANCE LEVELS: 0.01 ** 0.05 *

terms on OA imply that the derivative dNORMP/dOA is positive and rising for all OA above 0.417. The overwhelming majority of players in the sample are above this level. At bats and experience in post-season games both significantly increase price. Age significantly decreases price consistent with the potential productivity hypothesis, while the effect of retirement is insignificantly negative. Set dummies are almost all significantly positive, suggesting that relative prices within sets are affected by the scarcity of the set overall. Namely, star card prices tend to be higher multiples of common card prices in the older more scarce sets. BLACK and LATIN coefficients are negative, but only BLACK is close to significance at the 5% level with a t-ratio of -1.94.

In a regression incorporating an interactive variable between race and experience (ATBATS) the coefficients on BLACK, LATIN, BLACKAB and LATINAB were neither individually nor jointly significant (see Table 8). Two specifications including interactive variables between race and productivity are reported in Table 9. First, we add BLACKOA and LATINOA to the core specification. The BLACK coefficient is significantly positive and the BLACKOA coefficient significantly negative. Both the LATIN and the LATINOA coefficient are insignificant. Beyond OA equal to 0.468, prices for blacks are lower. Since 0.507 is the mean OA among blacks, the large majority of blacks fall in the range where the net effect on price is negative. The second specification uses both linear and quadratic terms for OA interacted with BLACK and LATIN. None of the variables for blacks are significant in this specification, suggesting that the linear OA specification is more suitable. However, for latins this specification does work better. LATIN, LATINOA, and LATINOASQ are all highly significant. However, at the OA mean for latins (0.457), latin card prices are substantially higher than white card prices, and it is only at OA levels (between 0.500 and 0.550) achieved

TABLE 9

TOBIT CARD PRICE REGRESSIONS:
STAR HITTERS -- RACE INTERACTED WITH OA AND WITH OA AND OASQ

DEPENDENT	NORMP		NORMP	
INDEPENDENT VARIABLES	REGRESSION COEFFICIENT	T-RATIO	REGRESSION COEFFICIENT	T-RATIO
BLACK	26.13	2.42 *	-64.73	-0.91
LATIN	-7.64	-0.61	-395.8	-2.97 **
ATBATS	0.003	7.99 **	0.003	8.11 **
OA	-657.1	-5.37 **	-824.2	-5.99 **
OASQ	775.6	6.19 **	948.3	6.69 **
BLACKOA	-55.88	-2.62 **	307.4	1.11
BLACKOASQ	---	---	-359.8	-1.34
LATINOA	14.36	0.55	1655	2.97 **
LATINOASQ	---	---	-1711	-2.97 **
MGR	-0.111	-0.33	-0.065	-0.20
RETIRED	-4.04	-1.33	-3.43	-1.14
AGE	-1.04	-5.71 **	-1.14	-6.05 **
POST	0.167	5.04 **	0.165	5.07 **
CATCHER	4.07	2.34 *	4.47	2.59 **
FIRST	-1.42	-1.07	-1.28	-0.99
SECOND	-0.228	-0.14	0.128	-0.08
SHORT	0.682	0.37	0.455	0.25
THIRD	-2.10	-1.31	-1.80	-1.13
S60	25.16	5.02 **	26.17	5.26 **
S64	22.68	4.83 **	23.02	4.96 **
S68	23.47	5.75 **	23.57	5.83 **
S71	17.22	4.22 **	17.85	4.38 **
S74	8.82	2.48 *	8.81	2.50 *
S77	9.54	2.76 **	9.66	2.81 **
S80	4.91	1.45	4.32	1.29
CONSTANT	154.4	5.04 **	196.1	5.72 **

P_{min} =	1.00	1.00
PREDICTED PROB $P > P_{min}$ GIVEN AVE X(I) =	0.512	0.507
OBSERVED FREQ. $P > P_{min}$ =	0.640	0.640
AT MEANS OF X(I), E(P) =	3.892	3.798
SQ. CORRELATION BETWEEN OBS. AND EXP. VALUES =	0.677	0.691
STD ERROR OF ESTIMATE =	6.994	6.859
LOG-LIKELIHOOD FUNCTION =	-794.6	-789.0
MCFADDEN'S R SQUARED =	0.165	0.171
TOTAL OBSERVATIONS =	347	347
LIMIT =	125	125
NON-LIMIT =	222	222

ASYMPTOTIC T-RATIO SIGNIFICANCE LEVELS: 0.01 ** 0.05 *

by very few latins in the sample that latin prices drop below white prices at a given OA level. Therefore although the discrimination variables are significant for latins, at the levels of experience and productivity representative of most latin stars, discrimination, if anything, may slightly favor latins.

III. CONCLUSION

Our analysis has shown that black hitters from our 1960-61 sample are not discriminated against, but that black hitters from the more recent sample (1977) and from the star sample are discriminated against by baseball card buyers. In both the 1977 and the star samples, card prices respond less to increased productivity for blacks than for whites. As a consequence, prices for blacks are below prices for whites for the majority of black hitters, and the price gap is largest for the best black hitters. The results also reveal little, if any, discrimination against latin hitters. Coefficients on latin variables in the 1977 sample are insignificant with the one exception that the same increase in playing experience causes prices to rise less for latins than for whites. Discrimination against latin hitters could not be concluded from the 1960-61 or star samples. Our results provide support, therefore, for the proposition that discrimination by baseball fans against black hitters does exist. Our evidence also indicates that discrimination is more intense against black players who have played more recently, and especially, more intense against the better black players.

REFERENCES

Andersen, Torben. "Race Discrimination by Major League Baseball Fans." Unpublished Phd dissertation, University of Washington, 1988.

Becker, Gary. "A Theory of the Allocation of Time." *Economic Journal* 75 (September 1965): 493-517.

____________. *The Economics of Discrimination*, 2nd ed. (Chicago: Univ. of Chicago Press, 1971).

Beckett, James, and Eckes, Dennis V. *The Sport Americana Baseball Price Guide*, Numbers 2-7 (Lakewood, Ohio: Edgewater, 1980-1985).

Bennett, Jay M., and Flueck, John A. "An Evaluation of Major League Baseball Offensive Performance Models." *The American Statistician* 37 (February 1983): 76-82.

Borjas, George J. "The Politics of Employment Discrimination in the Federal Bureaucracy." *The Journal of Law and Economics* 25 (October 1982): 271-99.

Canadian Almanac and Directory. (Toronto: Clopp Clark Pitman, 1984-1987).

Christiano, Kevin. "Salaries and Race in Professional Baseball: Discrimination 10 Years Later." *Sociology of Sport Journal* 5 (1988): 136-49.

Gwartney, James, and Haworth, Charles. "Employers' Costs and Discrimination: The Case of Baseball." *Journal of Political Economy* 82 (July/August 1974): 873-81.

Hill, James R.; Madura, Jeff; and Zuber, Richard A. "The Short Run Demand for Major League Baseball." *Atlantic Economic Journal* 1982: 31-55.

Hill, James R., and Spellman, William. "Pay Discrimination in Baseball: Data from the Seventies." *Industrial Relations* 23 (1984): 103-112.

Judge, George G., et al. *Introduction to the Theory and Practice of Econometrics*. (New York: Wiley, 1982).

McDonald, John F., and Moffitt, Robert A. "The Uses of Tobit Analysis." *Review of Economics and Statistics* 62 (1980): 318-21.

Maddala, G.S. *Limited Dependent and Qualitative Variables in Econometrics* (New York: Cambridge University Press, 1983).

Pascal, Anthony H. and Rapping, Leonard A. "The Economics of Racial Discrmination in Organized Baseball." In A.H. Pascal, ed., *Racial Discrimination in Economic Life* (Lexington, Mass.: Heath, 1972).

Raimondo, Henry J. "Free Agents' Impact on the Labor Market for Baseball Players." *Journal of Labor Research* 4 (1983): 183-93.

Reichler, Joseph, ed. *The Baseball Encyclopedia* (New York: MacMillan, 1985).

Scully, Gerald. "Discrimination: The Case of Baseball," *Government and the Sports Business*. Brookings, Washington, D.C., 1974: 221-74.

__________. "Economic Discrimination in Professional Sports." *Law and Contemporary Problems* (Winter-Spring 1973): 67-84.

__________. "Pay and Performance in Major League Baseball." *American Economic Review* (December 1974): 915-30.

Sommers, Paul M., and Quinton, Noel. "Pay and Performance in Major League Baseball: The Case of the First Family of Free Agents." *The Journal of Human Resources* (Summer 1982): 426-36.

State and Metropolitan Area Data Book (Washington, D.C.: U.S.G.P.O., 1982).

Summers, G.W.; Peters, W.S.; and Armstrong, C.P. *Basic Statistics in Business and Economics*, 2nd ed. (Belmont: Wadsworth, 1977).

Tobin, James. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26 (1958): 24-36.

White, Kenneth J. "A General Computer Program for Econometric Methods - SHAZAM." *Econometrica* 46 (1978): 239-40.

NOTES

1. See Becker (1971): 14-17. One recent study that attempts to isolate the relationship between customer preference and wage discrimination is Borjas (1982).

2. Among prominent early studies of salary discrimination Pascal and Rapping (1972) find no discrimination, while Scully (1974a) finds significant discrimination against blacks, especially more productive blacks. More recent studies reveal only weak signs of salary discrimination in contemporary baseball (see for instance Raimondo (1983), Hill and Spellman (1984), and Christiano (1988).

3. Scully (1974a) and Hill, Madura and Zuber (1982) find that attendance at major league games is decreased when the home team's starting pitcher is black, taking into account productivity differences between black and white pitchers. Scully (1973) finds that total revenue for major league baseball teams is reduced by employing more blacks. On the other hand, Gwartney and Haworth (1974) find that adding blacks raises a team's home attendance, holding constant the effects on team winning performance. Sommers and Quinton (1982) find that the percentage of black players on a team does not significantly affect team revenue.

4. We are not arguing that baseball card collectors are a randomly drawn subset of baseball fans in general. Card collectors clearly seem to be intramarginal fans. Yet we can think of no reason to believe that the indication of racial preferences displayed by the card collectors' behavior should not be an unbiased estimator of baseball fans' racial preferences.

5. Beckett and Eckes (1983): 2-6.

6. Prices are taken from Beckett and Eckes, *Sports American Baseball Card Price Guide*. This annually updated guide provides a market price for every individual card ever produced for a major set. Another reason to use prices from the *Guide* is that it is widely used by collectors; many buy and sell offers in classified ads in the hobby periodicals cite prices that are tied to the *Guide*.

7. See Becker (1965) and Lancaster (1966) for analyses of hedonic demand models.

8. In other sports, individual performance statistics can be very deceiving. A basketball player may lead the league in average points per game yet not interact well with teammates. His contribution to team success could be quite low, despite impressive statistics. Since hitting (and to a lesser degree pitching) involve little interaction with other players, individual performance statistics in these areas more closely reflect productivity than do individual statistics in other team sports which stress "team" play.

9. See Bennett and Flueck (1983).

10. BA = (base hits)/(at bats). SA = (total bases gained on base hits)/(at bats). OA = (total bases gained on base hits + walks + stolen bases)/(at bats + walks).

11. Bennett and Flueck (1983): 82.

12. It is quite likely that recent seasons have more impact on popularity than seasons farther in the past. This suggests some system of lagged weights, but without any *a priori* knowledge about the nature of these weights we prefer to arbitrarily assign weights to every season.

13. Given the close relationship between YRET and RETIRED, each empirical model contains only one of these two variables.

14. The fascination with rookie cards, however difficult to rationalize, does have parallels in other collecting hobbies, e.g., the relatively high demand for first editions.

15. See Tobin (1958), Judge (1982): 526-8, and Maddala (1983): 149-58 for general accounts of Tobit regression techniques.

16. Two contemporary players meeting the requirements were excluded. Ryne Sandberg and Wade Boggs were omitted since their cards were rookie cards. No rookie cards are included in this sample.

17. To test for differences in means between whites and blacks (or latins) we calculate the statistic Z:

$$Z = A/S_A$$

where $A = X_w - X_b$, X indicates a sample mean, and S_A is the standard error of the difference estimated from the sample variance as:

$$S_A = ((S_w^2/n_w) + (S_b^2/n_b))^{.5}$$

If either $n < 30$, we use a small sample procedure that assumes equal variances for the two populations. Then:

$$S_A = S((1/n_w) + (1/n_b))^{.5}$$

where S^2 is a pooled estimate of the common population variance from the two sample estimates:

$$S^2 = ((n_w - 1)S_w^2 + (n_b - 1)S_b^2)/(n_w + n_b - 2)$$

See Summers, Peters, and Armstrong (1977): 310-2.

18. This convention of reporting transformed coefficients evaluated at the sample means is suggested by McDonald and Moffitt (1980). They also provide the derivative of the transformation:

$$dE(PRICE)/dX_i = b_i\ [F(z) - P_{min} * f(z)/\sigma]$$

wehre b_i is the Tobit regression coefficient on variable X_i, F(z) is the cumulative normal distribution function evaluated at z, f(z) is the unit normal density function, P_{min} is the limit value of the dependent variable PRICE, and σ is the standard deviation of the error term. F(z) is, in the SHAZAM output, "the predicted probability of PRICE > P_{min} given average X(I)." Given F(z), z and then f(z) can be determined from statistical tables. The standard error of the estimate is used for σ.

19. SET61 equals one for 1961 cards and zero for 1960 cards. It measures potential differences in fan demand for cards in the two sets. HIGH is a dummy variable assigned to the relatively more scarce cards in the 1960-61 sample. It captures whatever effects supply variation may have on card prices.

20. A full-time player achieves 500-600 at bats per season.

21. These specifications are not reported.

22. When SA replaces OA in the regressions, the coefficient magnitudes and significance levels drop for both BLACK and LATIN. This may be because blacks and latins average significantly more stolen bases per at bat than whites, outweighing the slight advantage of whites in receiving walks; stolen bases and walks are omitted in SA but included in OA.