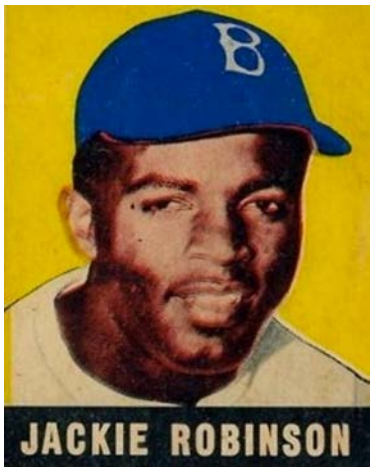
EJW
ECON JOURNAL WATCH
Scholarly Comments on Academic Economics

ECON JOURNAL WATCH 9(2)
May 2012: 122-140

Race, Ethnicity, and Baseball Card Prices: A Replication, Correction, and Extension of Hewitt, Muñoz, Oliver, and Regoli

David W. Findlay[1] and John M. Santos[2]

LINK TO ABSTRACT



Since Pascal and Rapping's (1972) seminal article, scholars have given much attention to racial and ethnic discrimination in professional sports. Because high-quality performance statistics are so accessible, professional sports is a great area for testing discrimination. Following Becker (1971), scholars consider discrimination by other players (employee discrimination), owners (employer discrimination), and fans (customer discrimination). Kahn (1991a, 2000) has provided excellent reviews of the literature on discrimination in professional sports.

Our paper investigates whether consumers in the secondary market for baseball cards[3] are willing to pay more for cards of

1. Colby College, Waterville, ME 04901.
2. Robert Morris University, Chicago, IL 60605. We thank two anonymous referees for helpful comments and suggestions.
3. All cards pictured are among those used in the analysis. The authors would like to thank The Topps Company for its assistance. Images of Topps baseball cards used courtesy of The Topps Company, Inc. For more information about The Topps Company, please see their website at www.topps.com.

players of a particular race or ethnicity, holding performance and card availability constant. The issue is important. Becker and others have argued that discrimination based on racial or ethnic preferences by employers is unlikely to persist in the long-run in competitive markets as the discriminating employers face higher productions costs than non-discriminating employers. Customer discrimination, however, can potentially persist under competitive conditions as employers are able to pass along the higher costs to customers willing to pay a premium to indulge their racial or ethnic preferences. Kahn (1991b) has demonstrated that customer discrimination, under certain conditions, can cause persistent wage differentials in competitive markets.

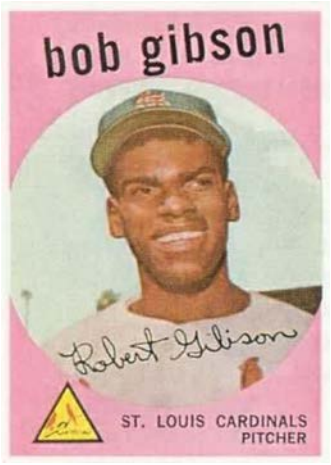
The results of prior empirical studies of the baseball card market and discrimination are mixed. Evidence that black or Hispanic players with equivalent performance statistics hold lower card prices is found by Nardinelli and Simon (1990), Andersen and La Croix (1991), Gabriel, Johnson, and Stanton (1999), and Fort and Gill (2000). In contrast, evidence that race and ethnicity have little effect on card prices is found by Gabriel, Johnson, and Stanton (1995) and McGarrity, Palmer, and Poitras (1999). Scahill (2005) obtains evidence of both racial and ethnic discrimination in the baseball card market, but he notes that the negative effects of player race and ethnicity on card prices disappear in the more recent years of his sample.

We add to this literature by correcting, replicating, and extending the approach of another paper, which we introduce in the next section. Our paper offers baseball lovers the chance to mix the pleasures of empirical economics with references to great heroes of the game. To readers generally, it offers improved results on the question of whether race or ethnicity affects how much card buyers are willing to pay to own the card of the hero.

Summary of Hewitt, Muñoz, Oliver, and Regoli (2005)

Hewitt, Muñoz, Oliver, and Regoli (2005) examine the effects of the race of the baseball player on the price of his rookie card. In creating their sample of players, Hewitt, Muñoz, Oliver, and Regoli (hereafter HMOR) include only those

players who were elected to the National Baseball Hall of Fame prior to 2004. As Jackie Robinson broke the race barrier in 1947, HMOR exclude players whose rookie cards were issued prior to 1948; they also exclude five Hispanic players. The result is a sample of 51 players where blacks account for 2 of the 19 pitchers and 16 of the 32 hitters.[4]



The dependent variable in their study is the price of each player's rookie card as reported in the April 2003 issue of *Beckett Baseball Card Monthly*.[5] HMOR hypothesize that card prices may be affected by player performance, card availability, and player race. HMOR (415) use the statistic Total Baseball Ranking (TBR) as their measure of player performance, and they report that their data source is the 2001 issue of *Total Baseball* (Thorn, Palmer, and Gershman 2001). For pitchers, Total Baseball Ranking equals the pitcher's Total Pitcher Index (TPI). For hitters, Total Baseball Ranking equals the hitter's Total Player Rating (TPR). Leaving the details aside,[6] we highlight only that the TPI data for a pitcher and the TPR data for a hitter are reported in Thorn, Palmer, and Gershman (2001) by season and by career.[7] HMOR use the career totals.

Throughout our paper, even in our extensions, we keep to the parsimonious spirit of the model specified by HMOR—the model is parsimonious in that it uses for each player a single performance statistic. In this sense we follow HMOR and do not investigate disaggregated measures of player performance.

HMOR (416) also hypothesize that the "value [price] of a baseball card is affected by how scarce or rare it is." The authors use data obtained from the April 2003 Professional Sports Authenticator (PSA) Population Report to measure

4. See HMOR (2005) for a detailed discussion of the criteria.

5. The April 2003 issue reports two prices for each card, a LO price and a HI price. A comparison of the data reported in HMOR with that found in *Beckett* (2003) reveals that HMOR use the HI price data as their dependent variable.

6. Thorn, Palmer, and Gershman (2001, 2501) define the Total Pitcher Index as "the sum of a pitcher's Pitching Runs—expressed as Ranking Runs, employing the same formula used to compute Relief Ranking Runs—Batting Runs (in the AL since 1973, zero), and Fielding Runs, all divided by the Runs per Win factor for that year (generally around 10, historically in the 9-11 range)." They define the Total Player Rating as "the sum of a player's Adjusted Batting Runs, Fielding Runs, and Base Stealing Runs, minus his positional adjustment, all divided by the Runs per Win factor for that year (generally around 10, historically in the 9-11 range)" (2501). For an even more detailed discussion of the construction of TPI and TPR, see Thorn, Palmer, and Gershman (2001).

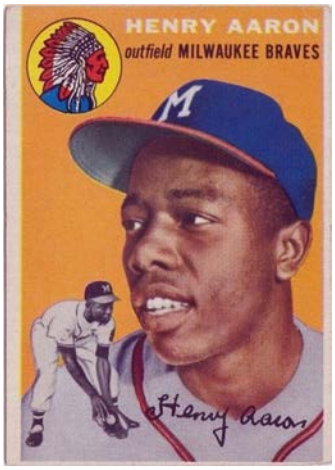
7. The career value of a pitcher's TPI is simply the sum of the pitcher's yearly TPI statistics over the course of his career. The career value of a hitter's TPR is calculated similarly.

card availability. Professional Sports Authenticator (**psacard.com**) provides information about card prices and documents the number and quality of the cards that it evaluates. HMOR (416) "constructed a measure of availability based on how many rookie cards of each player in near mint or better condition were reported [by PSA] to exist."[8]

HMOR hypothesize that player race may affect rookie card prices. A race dummy variable is constructed where Black equals one if a player was black and equals zero otherwise. As HMOR omitted the five Hispanic players from their sample, the reference group in their study is white players.

HMOR estimate two specifications of a baseball card price equation to determine whether player race affects the card prices of the 51 rookie cards. They find that card prices are higher for those players with greater career performance. The results also indicate that increases in card availability cause reductions in card prices. After controlling for their measures of player performance and card availability, HMOR find that player race has no statistically significant effect on card prices in either of their specifications and suggest (419) "the possibility that Black superstars or sports heroes are seen on the same plane as Whites. To put it differently, on some level, the evaluation of the performance of Black [Hall of Fame] members transcends racial consideration."

The Sources and Errors of HMOR's Data



Our Table 1 includes all information reported in HMOR's Table 1 and additional data in the final three columns. We applaud HMOR for reporting their data as it facilitates replication of their results. The first column in Table 1 lists the 51 players, Hank Aaron to Robin Yount. The second column includes the year of the card, the manufacturer (Bowman, Fleer Update, Leaf, and Topps), and card number. The third column includes the April 2003 HI price reported in *Beckett* (2003). The fourth column gives the PSA availability data as reported by HMOR. The fifth column includes the Total Pitcher Index or Total Player Rating reported by HMOR.

8. PSA reports the number of times it has evaluated each card for each quality category. These data are updated as PSA evaluates additional cards.

TABLE 1.

Name	Card	Price	Availability	HMOR Performance	*TB* (2001) Performance	Difference in %	Difference in Rank
Aaron, Hank	1954 T #128	1500	164	90.1	89.1	1.1	0
Banks, Ernie	1954 T #94	800	121	24.9	26.9	−7.4	−4
Bench, Johnny	1968 T #247	125	239	30.2	25.6	18	5
Berra, Yogi	1948 B #6	450	70	34.8	37.4	−7	−4
Brett, George	1975 T #228	80	1042	43.9	40	9.8	4
Brock, Lou	1962 T # 387	125	111	2	10.5	−81	−1
Campanella, Roy	1949 B #84	700	90	22.2	22.5	−1.3	−2
Carlton, Steve	1965 T #477	150	114	35.6	33.7	5.6	6
Carter, Gary	1975 T #620	15	74	30.1	30.1	0	−3
Drysdale, Don	1957 T #18	225	83	34.7	34.6	0.3	0
Fingers, Rollie	1969 T #597	40	109	22.5	22.5	0	0
Fisk, Carlton	1972 T #79	50	96	33.4	24.9	34.1	13
Ford, Whitey	1951 B #1	1400	35	39.2	39.2	0	2
Gibson, Bob	1959 T #514	200	84	46.3	43.7	5.9	4
Hunter, Jim	1965 T #526	80	136	6.9	6	15	1
Jackson, Reggie	1969 T #260	250	252	44	42.2	4.3	3
Jenkins, Fergie	1966 T #254	70	146	32.1	29.8	7.7	3
Kaline, Al	1954 T #201	600	119	45.9	45.2	1.5	1
Killebrew, Harmon	1955 T #124	250	159	32.8	27.3	20.1	6
Kiner, Ralph	1948 B #3	150	48	27	25.9	4.2	−1
Koufax, Sandy	1955 T #123	800	179	20	20.5	−2.4	0
Lemon, Bob	1949 B #238	200	13	35.2	38.4	−8.3	−4
Mantle, Mickey	1951 B #253	8500	53	77.4	77.4	0	0
Mathews, Eddie	1952 T #407	8000	15	52.2	52.2	0	0
Mays, Willie	1951 B #305	3000	56	95.9	95.9	0	0
McCovey, Willie	1960 T #316	125	64	38.1	37.3	2.1	4
Morgan, Joe	1965 T #16	60	29	63.9	54.8	16.6	0
Murray, Eddie	1978 T #36	80	1313	34.1	34.1	0	0
Musial, Stan	1948 B #36	800	49	70.1	71.5	−2	−1
Niekro, Phil	1964 T #541	80	37	38	33.8	12.4	9
Palmer, Jim	1966 T #126	100	200	36.4	34.9	4.3	2
Perry, Gaylord	1962 T #199	80 (85)	93	36.8	34.9	5.4	4
Puckett, Kirby	1984 F #93	100	1416	32.3	32.3	0	−1
Roberts, Robin	1949 B #46	250	39	25.9	25.9	0	−1
Robinson, Brooks	1957 T #328	350	214	23.3	20.1	15.9	4
Robinson, Frank	1957 T #35	200	162	71	67.6	5	1
Robinson, Jackie	1949 L #79	1500 (1100)	81	33.3	32	4.1	2
Ryan, Nolan	1968 T #177	600	398	14.2	20.7	−31.4	−3
Schmidt, Mike	1973 T #615	150	361	77.9	79.6	−2.1	0
Seaver, Tom	1967 T #581	500	264	51.2	48.7	5.1	1
Smith, Ozzie	1979 T #116	80	796	42.4	42.4	0	−1
Snider, Duke	1949 B #226	900	55	24.3	24.1	0.8	0
Spahn, Warren	1948 B #18	300	43	43.1	50.2	−14.1	−6
Stargell, Willie	1963 T #553	125	85	31.6	31.6	0	−1
Sutton, Don	1966 T #288	50	45	13.7	13.2	3.8	0
Wilhelm, Hoyt	1952 T #392	750	38	29.2	40.8	−28.4	−21
Williams, Billy	1961 T #141	60	215	30.1	30.1	0	−3
Winfield, Dave	1974 T #456	40	699	36.9	36.9	0	3
Wynn, Early	1949 B #110	125	59	18.2	17.1	6.4	1
Yastrzemski, Carl	1960 T #148	150	280	46.1	46.7	−1.3	0
Yount, Robin	1975 T #223	50	853	31.4	46	−31.7	−22

Notes: [a] B, F, L, and T refer to Bowman, Fleer Update, Leaf, and Topps cards, respectively. [b] We give the correct prices for Perry and J. Robinson (80 and 1500) and include in parentheses the incorrect price data reported in HMOR for those players (85 and 1100). [c] *Difference in %* is the percentage point difference between the HMOR performance data and the *Total Baseball* (2001) performance data. [d] *Difference in Rank* measures the extent to which the player's relative performance ranking changes when performance data reported in HMOR are used.

A casual inspection of the performance data reported by HMOR for pitchers indicates that TPI ranges from a high of 51.2 for Tom Seaver to a low of 6.9 for Jim Hunter. For hitters, the TPR ranges from a high of 95.9 for Willie Mays to a low of 2.0 for Lou Brock. We offer these casual observations because, when we first read the HMOR article in early 2010, it was the low performance value for Lou Brock that caught our eye. Brock is the only hitter in HMOR's table whose TPR statistic is less than 20; Roy Campanella has the next highest TPR statistic, 22.2.

Despite our initial ignorance about the TPI and TPR statistics, we decided to compare Brock's TPR figure reported by HMOR with that found in Thorn, Palmer, and Gershman (2001)—the same source of performance data reportedly used by HMOR.[9] Little did we know at the time that this would start a process where we would discover a large number of discrepancies and puzzles. We immediately discovered that Brock's career TPR, as found in HMOR's data source, equals 10.5 and not 2.0 as HMOR reported. We then decided to obtain the performance data for the remaining 50 players using the same source as HMOR.[10] For all 51 players in the sample, we consulted every section of the data source (as outlined in footnote 10) to record and to verify the performance data. In so doing, we discovered numerous differences between the data found in HMOR and those in Thorn, Palmer, and Gershman (2001): 38 of the 51 performance data observations in HMOR differ from those found in the 2001 edition of *Total Baseball*—HMOR's reported data source. In other words, nearly 75 percent of the performance data observations used by HMOR were apparently incorrect. We include in column 6 in Table 1 (see *TB (2001) Performance*) the correct performance data as found in the 2001 edition of *Total Baseball*.[11]

We constructed two additional variables in Table 1 to measure the magnitude of these performance data errors. The seventh column (*Difference in %*) includes the percentage point difference between the HMOR performance data and the

9. See pages 415, 416, and 425 in HMOR where HMOR discuss their performance data source.

10. The TPI data can be obtained in Thorn, Palmer, and Gershman (2001) by examining each pitcher's career statistics in the Pitcher Register section (pp. 1307-1957) or on the following pages: (1) p. 2317; (2) pp. 2317-2318; (3) pp. 2319-2320; and (4) pp. 2320-2321. The TPR data can be obtained from the same source by examining each hitter's career statistics in the Player Register section (pp. 553-1306) or on the following pages: (1) pp. 2303-2304; (2) pp. 2304-2305; (3) pp. 2319-2320; and (4) pp. 2320-2321.

11. We did discover one discrepancy in the 2001 issue of *Total Baseball* for pitcher Bob Lemon. His Total Baseball Ranking is reported to be 37.9 (see p. 2320) while his Total Pitcher Index is reported to be 38.4 (see pages 1587, 2317, and 2318). We use his TPI value of 38.4 in our paper.

performance data reported in *Total Baseball*. These errors range, in absolute value, from 1.1% for Hank Aaron to 81.0% for Lou Brock. The mean absolute value of these data errors measured in percentage terms is 11.3%. We also obtained a performance ranking of all 51 players using the HMOR data and the data we obtained from *Total Baseball* (2001). The last column (*Difference in Rank*) reports the extent to which each player's ranking changes when the data from HMOR are used. For example, the rankings of Carlton Fisk, Hoyt Wilhelm, and Robin Yount change rather dramatically.

Given the data discrepancies between HMOR and *Total Baseball* (2001), we decided to check the remaining data reported by HMOR. To verify the accuracy of the price data, we cross-referenced the card information reported by HMOR with the price data found in *Beckett* (2003), the same data source used by HMOR. We discovered two price data errors: (1) Gaylord Perry's 1962 Topps rookie card price is $80 (not $85); and (2) Jackie Robinson's 1949 Leaf rookie card price is $1500 (not $1100).[12] We then attempted to check the availability data. To do so, we contacted PSA and learned that they do not archive the data nor keep old issues of their Population Report; therefore, it is impossible to verify the accuracy of the availability data reported in HMOR.[13]

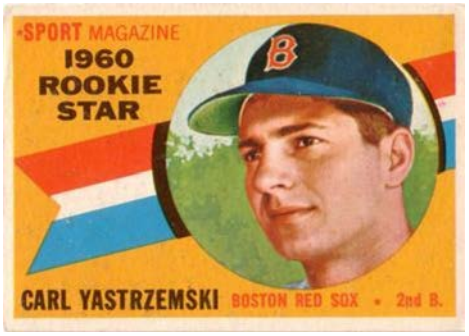
Our discovery of the data errors made us wonder whether HMOR's results are affected by these errors. Could the effects of race change? This is not a matter of mere curiosity for it speaks to the issue of racial discrimination.

In addition, we were perplexed by HMOR's explicit decision to remove Hispanic players from their sample. That decision causes an already small sample to be nearly 10 percent smaller than it otherwise would be. Second, the removal could potentially impact the estimated effects of race on card prices. Finally, by removing Hispanics from the sample, the authors could not examine the effects of player ethnicity on card prices. The latter implication was the most perplexing to us given that one would think that researchers examining customer discrimination would include a group of players, in this case Hispanic players, for whom such dis-

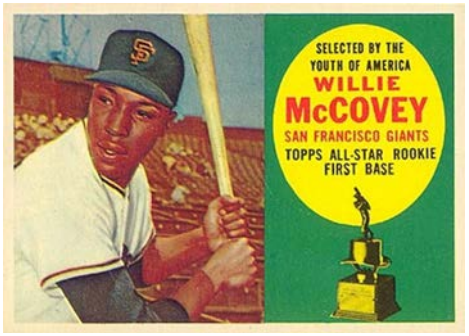
12. We discuss in more detail below what we eventually learned about the price data error for Jackie Robinson's rookie card.

13. While not data errors, we also discovered several card information errors in HMOR. Johnny Bench's card number is 247 (not 447). Steve Carlton's rookie card was issued in 1965 (not 1975). Al Kaline's card number is 201 (not 210). And finally, Willie Mays' card number is 305 (not 395).

crimination has already been found to exist in several previous studies of the baseball card market.



At this point in 2010, we believed that the errors were simply the result of transcription errors. We decided to write a paper that corrects, replicates, and extends the analysis of HMOR. The obvious outlet for this paper was the *Journal of Sport & Social Issues*, the same journal that published HMOR's 2005 paper. We completed the paper during the summer of 2011 and submitted it to the *Journal of Sport & Social Issues* on September 1, 2011. On September 20, 2011, we received an email from the Editor in Chief of the journal informing us that "Unfortunately, at this time, this article is not a good fit with the journal and I have decided against sending it out for review."[14]



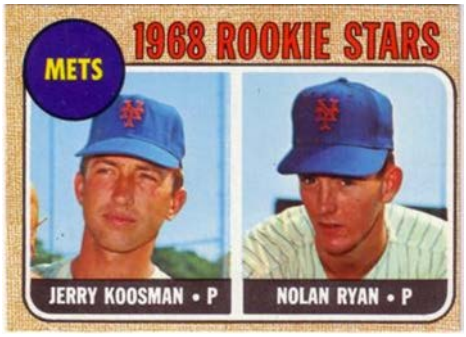
We immediately submitted the paper to *Econ Journal Watch*, and an editor there subsequently asked us to investigate the data errors in more detail. As we investigated, we recalled that HMOR cite Bill Deane's essay, "Awards and Honors"; the Deane essay was published in the 1989 issue of *Total Baseball*. We never imagined that HMOR would have used different editions of *Total Baseball* as their source of the performance data. During our initial investigation of the TBR data, we learned that *Total Baseball* revises the TBR statistics over time (see pages 539-542 of the 2001 edition). One, therefore, cannot pool or "mix" TBR data from different editions.

Without any other explanation for the data errors, we decided to obtain the performance data for all 51 players from the only other edition of *Total Baseball* cited by HMOR.[15] We discovered the following. The TBR values for only 13 players

14. We have had our papers rejected for publication in the past and anticipate that future papers will be rejected for publication as well. However, we were surprised and a bit perplexed that a paper that documents such a large number of data errors, in addition to exploring a number of extensions, of a previously published article in the same journal would not even be sent out for review.

15. We also obtained the 1991 and 1995 editions of *Total Baseball* to determine if HMOR may have relied on these other editions for their performance data. None of the TBR observations reported in HMOR equals the values found in these two editions of *Total Baseball*.

in HMOR's sample are equal to the values found in *Total Baseball* (2001). The TBR values for 38 players are equal to the values found in *Total Baseball* (1989).[16] For two of the players in the sample (George Brett and Frank Robinson), the TBR values used by HMOR are not equal to any of the values published in the 1989, 1991, 1995, and 2001 editions of *Total Baseball*. We believe that these two errors are transcription errors because the values used by HMOR are closest to the values reported in *Total Baseball* (1989). We thus discovered that HMOR used two different sources for their performance data and, therefore, pooled performance data that are revised over time.



There is an additional implication from HMOR's use of the 1989 issue of *Total Baseball*. By obtaining data from this earlier source, HMOR have performance data only through the 1988 season. There are ten players in their sample whose careers continued *after* 1988. Of these ten players, HMOR use the 1989 issue for five of them, namely, George Brett (who retired in 1993), Carlton Fisk (1993), Nolan Ryan (1993), Mike Schmidt (1989), and Robin Yount (1993). For these five players, career performance is incomplete in the data that HMOR used to estimate their price equations.

Also, the Jackie Robinson rookie card in HMOR is the 1949 Leaf card (#79). HMOR list the price of that card as $1100. We discovered that Jackie Robinson had another rookie card issued in 1949 by Bowman (#50). As reported in *Beckett* (2003), the price of the Leaf card is $1500 and the price the Bowman card is $1100. We believe that HMOR incorrectly used the Bowman price rather than the Leaf price.[17]

16. As it turns out, the career total TBR observations for two players in the sample (Willie Stargell and Robin Roberts) are the same in the 1989 and 2001 editions of *Total Baseball*. It is a coincidence that the career totals for Stargell and Roberts are identical across the two editions because their annual observations are not equal in the two editions of *Total Baseball*. This, of course, implies that the revised calculations used to construct the more recent TBR values caused the career total TBR statistics to change for the other 49 players in the sample.

17. The incorrect use of the Bowman price has a potentially additional implication for their results. HMOR also obtain availability data for each card in their sample. It is possible that HMOR included in their study the availability data for the Leaf card rather than the availability data for the Bowman card. As noted previously, because PSA does not archive the data, we could not determine whether HMOR used the availability data for Jackie Robinson's Leaf or Bowman card.

A Replication and Correction of HMOR

We estimate the same two specifications of HMOR's price equation:

(1) $Price = a_0 + a_1 Availability + a_2 Performance + a_3 Black$
(2) $Price = b_0 + b_1 Availability + b_2 Performance + b_3 Black + b_4 Black*Performance$

where Price is the natural log of the rookie card price; Availability is the natural log of the card availability obtained from PSA; Performance is the TPI/TPR for pitchers/hitters, respectively; and Black is a dummy variable equal to one if the player is black and zero otherwise. For equation (1), any effect of race on card prices is assumed to be independent of player performance. For equation (2), the inclusion of the interaction term, Black*Performance, allows for any race effects to vary by player performance. We estimate the equations first using the data reported by HMOR and then using the corrected data reported in Table 1.

We are able to replicate HMOR's results when we use the data reported by HMOR. This is a result in and of itself given, as noted by Winfree (2010, 48), the "difficulties in replicating results in sports economics." Winfree further explains "that there can be many roadblocks to replicating the results of any empirical study" (ibid.). We again applaud HMOR for removing one of those roadblocks by reporting their data. Furthermore, had HMOR never reported their data, we likely would not have discovered their data and measurement errors.

When using the correct performance and price data, none of the estimated coefficients on the race variables, Black or Black*Performance, is statistically significant. Our results, using the correct data, imply that player race has no effect on card prices—a result consistent with that originally reported by HMOR.

Using least squares regression to find the equation of best fit, we obtain the following estimate of equation (1) when we use the correct data reported in Table 1[18]:

(3) $Price = 6.265 - 0.411 Availability + 0.030 Performance - 0.098 Black$
(7.43) (2.60) (3.44) (0.27) $R^2 = 0.30$

We show in the supplement to this paper [**link**] that the estimated coefficient for the Black dummy variable is not statistically significant even at the 0.10 level. Specifically, the p-value for the t-statistic on the estimated coefficient on the Black

18. We include in parentheses the absolute value of the corresponding t-statistic for each estimated coefficient.

dummy variable in equation (3) indicates that one could only conclude that the Black coefficient is significantly less than 0 at a .393 significance level. Thus we find a lack of evidence that a player's race adversely affects the price of his rookie card, holding constant the effects of player performance and card availability. Although our results indicate that player race has no statistically significant effect on baseball card prices, we are mindful of Ziliak and McCloskey (2004, 334) who note that "statistical significance, to put it shortly, is neither necessary nor sufficient for a finding to be *economically* important." The estimated coefficient on the Black dummy variable indicates that the price of a black player's rookie card, all else fixed, is 9.3% lower than that of an otherwise identical white player.[19]

We also examined whether our results were affected by the inclusion of Jackie Robinson in the HMOR sample. Jackie Robinson was an epic figure who transcends baseball. Furthermore, unlike the vast majority of the players in the sample, he started his professional career in the Negro Baseball League. And finally, as we discussed in footnote 17, HMOR's incorrect use of Jackie Robinson's Bowman rookie card price may have resulted in HMOR incorrectly using the availability data for the Leaf card rather than the availability data for the Bowman card. All of our previous results are basically the same after we remove Jackie Robinson from the sample. Specifically, the size and significance of all parameters are nearly identical to those obtained when Jackie Robinson is included in the sample.

The negative and significant estimated coefficient on card availability in equation (3) indicates, as expected, that increases in card availability lower the selling price of a player's rookie card. Of particular interest to economists is the magnitude of the estimated coefficient on availability. Since both Availability and Price are measured in log form, we can interpret our numerical estimate of the Availability coefficient as implying that a 1 percent reduction in card availability raises card prices by about 0.41 percent or, equivalently, that a 10 percent reduction in card availability raises card selling price by approximately 4.1 percent. The positive and significant estimated coefficient on Performance indicates, not surprisingly, that the baseball cards of players with higher career performance levels sell for a higher price in the secondary market.

At the end of equation (3), we report a well-known sample statistic, R^2, which measures the percent of the sample variation in the dependent variable explained by an estimated equation. When we re-estimate HMOR's equation (1) after cor-

19. As noted by Halvorsen and Palmquist (1980), the appropriate interpretation of the coefficient of a dummy variable in a semilogarithmic regression equation can be calculated as $100(e^{\beta_d} - 1)$ where β_d represents the estimated coefficient of the dummy variable.

recting their reported 2003 data, our corrections raise the percent of variation in card prices explained by HMOR's basic model from 27 percent to 30 percent.[20]

Extensions of HMOR

HMOR's basic model explains only 30 percent of the observed variation in rookie card prices using correct data. We sought a model of card prices that explains or predicts a greater percent.

Matching Price and Card Quality

HMOR state (416) that they use prices listed in *Beckett* (2003) that are "in near mint or mint condition."[21] A related issue is HMOR's use of the availability data reported by Professional Sports Authenticator. HMOR note (416) that their "measure of availability [is] based on how many rookie cards of each player in near mint or better condition were reported to exist." However, we discovered that PSA determines the authenticity and quality of cards submitted. That is, PSA will grade (i.e., categorize) cards, for example, as excellent, near mint, near mint-mint, mint, and gem mint—each of which will receive the corresponding numerical score of 6, 7, 8, 9, and 10. Thus, the availability data used by HMOR represent the total number of cards rated by PSA as near mint or better. Their availability variable, therefore, is the sum of or, equivalently, an aggregate measure of all PSA-rated cards that are near mint, near mint-mint, mint, and gem mint.

Upon further investigation, we discovered that PSA also reports price data for cards with different grades. As noted in PSA's *Sports Market Report* (2003, 36), PSA's prices "represent average dealer selling prices for PSA-graded cards. […] The prices that appear in Sports Market Report are for sportscards that have been graded and sold, […]." The average selling prices listed by PSA, therefore,

20. When we re-estimated equation (1) using the data reported by HMOR, we were able to replicate their results and, therefore, obtained the same value of the R^2 statistic.

21. As we discovered, however, the prices they use are the HI prices reported by *Beckett*. To obtain prices of near mint or mint condition cards, one would have to use the "Price Guide Percentage by Grade" table found on p. 15 of the April 2003 issue to adjust the HI prices by quality. It is also important to note that these price guide percentages do vary over time.

correspond to cards that the rating company evaluates to be a specific quality (e.g., mint) and, therefore, do *not* represent an average price of, say, near mint or better cards. This information implies that the choice of the price data depends crucially on the type of PSA availability data one uses or creates. And given that different prices exist for different quality cards, the question naturally arises as to the extent to which the choice of price data series affects any model's ability to explain card prices.[22]

At this point, we realized that HMOR do not investigate what is the appropriate price series for the aggregate card quality category they create. Additionally, by restricting their analysis to cards of this aggregate quality category, neither do they explore whether the magnitude of the availability effect on card prices varies by card quality. Since PSA does not archive availability data, we are unable to retrieve card availability by specific card quality for the 2003 data used by HMOR. However, for 2010, we obtained availability and card quality data in real-time (April 2010) from PSA. This allows us to explore the relationship between card availability and card prices across specific card quality categories.

2003 Data: Which Price Series Does HMOR's Availability Measure Best Explain?

Table 2 displays the percent of variation in rookie card prices explained by alternative specifications of the card price equation for the corrected 2003 data and for our update using 2010 data.[23] The first two rows show the percent of card prices explained by HMOR's basic model when we use HMOR's data and when we use the correct data, respectively. Row 3 shows the gains in percent of variation of card prices explained when we consider alternative price series and card availability pairings beyond HMOR's pairing (*Beckett* HI-PSA7 Plus).[24] Specifically, in columns 2, 3, and 4 for Alternate Model 1, we match for the 2003 data HMOR's availability measure with three different PSA price series. Judged by the percent of sample variation in card prices explained, HMOR's availability measure explains all three PSA price series better than the *Beckett* price series chosen by the authors. Additionally, as we show in the supplement to this article, the statistical significance

22. While one could argue that the PSA data serve as a proxy for the quantities of each card that exist in the secondary market for baseball cards, we must note the following. It is possible that the *same* card has been submitted to PSA by, say, five different individuals. PSA would then report a 5 for that card even though the same card has been submitted five different times.

23. As we discussed earlier in the paper, the percent of variation in rookie baseball card prices explained by each model specification is measured by the well-known sample statistic, R^2 (also called the multiple coefficient of determination).

24. PSA7 Plus equals the number of cards rated near mint or better.

of the availability effect on card prices also increases when any of the three PSA price series are matched with HMOR's availability measure.

TABLE 2. Percent of Baseball Card Prices Explained

	Price Series	*Beckett*	PSA7	PSA8	PSA9	*Beckett*	*Beckett*	PSA8	PSA9
	Year of Data	2003	2003	2003	2003	2010	2010	2010	2010
	Availability Measure	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA8	# of Cards Rated PSA8	# of Cards Rated PSA9
	Source of Availability Data	HMOR	HMOR	HMOR	HMOR	PSA	PSA	PSA	PSA
Row/ Model	Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Replication of HMOR with HMOR Data [without Hispanics]		27%							
(2) HMOR Model with Corrected Data [without Hispanics]		30%							
(3) Alternate Model 1: Different Price and Availability Variables and Career Total Performance [without Hispanics]			37%	38%	33%	49%	52%	66%	69%
(4) Alternate Model 2: Different Price and Availability Variables and Career Total Performance [with Hispanics]						49%	53%	66%	69%
(5) Alternate Model 3: Different Price and Availability Variables and Per Season Performance [without Hispanics]		36%	40%	43%	34%	50%	53%	64%	68%
(6) Alternate Model 4: Different Price and Availability Variables and Per Season Performance [with Hispanics]						50%	53%	64%	68%

Notes: [a] *Beckett* equals the *Beckett* HI price. [b] PSA7 equals near mint. [c] PSA8 equals near mint-mint. [d] PSA9 equals mint. [e] To be precise, this table reports the percent of the variation in the natural log of baseball rookie card prices explained by each model as measured by the unadjusted multiple coefficient of determination (i.e., the R^2 statistic).

2010 Data: Which Price Series-Card Availability Pairing Best Explains Card Prices?

Until this point, we have used HMOR's aggregate measure of card availability. For the 2010 data, we are able to create card availability categories, both by aggregated card quality levels (a la HMOR) and for disaggregated card

quality levels. We take advantage of our availability data set to estimate separate card price equations for different categories of card quality and to determine which price-availability pairing best explains variations in rookie card prices. We find that HMOR's price equation with their aggregate availability category explains 49 percent of the variation in 2010 rookie card prices as compared to 30 percent for the corrected 2003 data, indicating that HMOR's model fits the 2010 data better than the 2003 data. However, when we disaggregate card availability by specific card quality and match to the corresponding quality-specific PSA price series, the percent of rookie card prices explained rises sharply; see columns (7) and (8) for Alternate Model 1 in Table 2.

What Is the Responsiveness of Card Prices to Changes in Card Availability?

Our estimate of the magnitude of the effect of availability varies across the different card quality categories reported in Table 2.[25] For the 2003 availability data reported by HMOR, we find that a 10 percent reduction in availability of cards rated near mint or higher causes a 4.1 percent increase in card price when we measure price with the *Beckett* price series used by HMOR. In contrast, for the 2010 data, we find that when we appropriately match the aggregated or disaggregated PSA availability measure to the corresponding PSA price series, we can predict that a 10 percent reduction in availability in the secondary market causes about a 9 to 10 percent increase in market price.

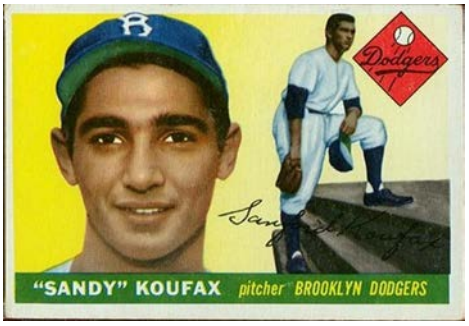
The Curious Omission of Hispanics: Does Ethnicity Matter?

HMOR did not explore whether their important finding on race extended to Hispanic ethnicity in the matter of Luis Aparicio, Rod Carew, Roberto Clemente, Juan Marichal, and Tony Perez. For the 2010 data for which we have access to availability data, we added Aparicio, Carew, Clemente, Marichal, and Perez to our sample, raising the sample size of players from 51 to 56. We then created an additional dummy variable that equals one if the observed player is

25. In the supplement to this article, we report and discuss the estimated equations behind the results in Table 2.

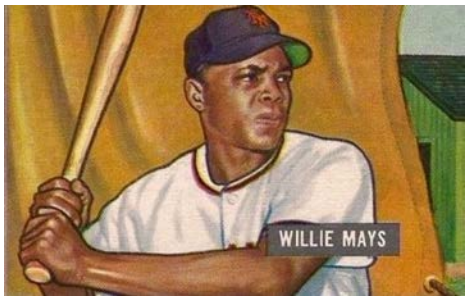
Hispanic and equals zero otherwise. The reference group in our sample is white players. As indicated in columns 5, 6, 7, and 8 for Alternate Models 1 and 2, the percent of variation in card prices explained is not greatly affected one way or the other when we add the five Hispanic players. Moreover, as the regression results we report in the supplement to this article show, we find very little change in the size or significance of the effects of player performance and card availability on card prices. Additionally, race remains statistically insignificant in all equations. Finally, we also find that there is no statistically significant effect of ethnicity on card prices. In general, our corrections and extensions show that neither race nor ethnicity has a statistically significant effect on card prices, and that the findings are robust to alternative pairings of price and availability measures.

Lifetime Player Performance: Career Totals or Average Per Season?



Our analysis to this point is based on the performance statistic used by HMOR, the *sum* of each player's TPI or TPR over the course of the player's career. An issue arises as to whether the underlying demand for these cards depends on how efficiently the player achieved his career totals. Jim Palmer and Gaylord Perry both have the same career TPI (34.9), but Palmer did it in 19 seasons while Perry did it in 22. In fact, Don Drysdale accumulated a slightly lower TPI (34.6) in just 14 seasons. Johnny Bench and Carlton Fisk had similar TPR statistics (25.6 and 24.9, respectively) with different career lengths—17 seasons for Bench and 24 seasons for Fisk. And finally, Kirby Puckett's career TPR statistic (32.3) was only slightly higher than Willie Stargell's (31.6). However, Puckett played in 9 fewer seasons than Stargell (12 versus 21). To examine this issue, we introduce a per season performance measure, either pitcher's TPI per season (TPI/seasons played) or hitter's TPR per season (TPR/seasons played). As row 5 (Alternate Model 3) of Table 2 shows for the 2003 data, the percent of variation in card price rises modestly when lifetime player performance is measured on a per season basis. For the 2010 data, however, the percent of card prices explained (with or without Hispanics) remains largely unchanged when we replace the career performance variable with the career per season average variable. Thus, we do not have a consistent answer across data sets as to whether baseball card market participants focus more on career or on per season achievement.

Willie Mays, Mickey Mantle, and the Guy Next-door



In this study we follow HMOR's approach of using a single performance statistic. Our study finds that race and ethnicity of baseball greats do not matter to the prices of rookie cards. As noted at the outset, other studies have found differently and others similarly. The test is an important one. Prices communicate information about market valuations. Combining prices with data on performance and card quality and availability might also tell us something about values and race or ethnicity.

Our discovery of measurement errors in HMOR's data set together with the omission of Hispanics from their sample raised concerns about the sensitivity of their finding of the absence of customer discrimination in the market for baseball cards. He and McGarrity (2004, 89) caution that "when estimating models with small data sets, empirical research runs the risk that errors may be responsible for the conclusions that scholars draw." They conclude (97) that "in these cases, careful examination of the data and robustness estimation techniques can clearly improve the analysis." Our corrections and extensions paint a picture of a market in which card prices reflect career performance, independent of player race or ethnicity.

We know how great Willie Mays and Mickey Mantle were. Race does not appear to matter in this market. Perhaps race matters more when the individual's merit is less well known and people rely on their notions, often faulty, about how merit and race (or ethnicity) co-vary.

Appendix

A supplement to this paper (**link**) estimates various specifications of the card price equations used by Hewitt, Muñoz, Oliver, and Regoli (2005). An Excel file (**link**) contains data used in the analysis.

References

Andersen, Torben and Sumner J. La Croix. 1991. Customer Racial Discrimination in Major League Baseball. *Economic Inquiry* 29(4): 665-677.

Becker, Gary S. 1971. *The Economics of Discrimination*. Chicago: University of Chicago Press.

Beckett Baseball. 2010. 10(3). Issue 50.

Beckett Baseball Card Monthly. 2003. 20(4). Issue 217.

Fort, Rodney and Andrew Gill. 2000. Race and Ethnicity Assessment in Baseball Card Markets. *Journal of Sports Economics* 1(1): 21-38.

Gabriel, Paul E., Curtis D. Johnson, and Timothy J. Stanton. 1995. An Examination of Customer Racial Discrimination in the Market for Baseball Memorabilia. *Journal of Business* 68(2): 215-230.

Gabriel, Paul E., Curtis D. Johnson, and Timothy J. Stanton. 1999. Customer Racial Discrimination for Baseball Memorabilia. *Applied Economics* 31(11): 1331-1335.

Halvorsen, Robert and Raymond Palmquist. 1980. The Interpretation of Dummy Variables in Semilogarithmic Equations. *American Economic Review* 70(3): 474-475.

He, Ling and Joseph P. McGarrity. 2004. Data Errors in Small Data Sets Can Determine Empirical Findings. *Atlantic Economic Journal* 32(2): 89-99.

Hewitt, John D., Robert Muñoz, Jr., William L. Oliver, and Robert M. Regoli. 2005. Race, Performance, and Baseball Card Values. *Journal of Sport & Social Issues* 29(4): 411-425.

Kahn, Lawrence M. 1991a. Discrimination in Professional Sports: A Survey of the Literature. *Industrial and Labor Relations Review* 44(3): 395-418.

Kahn, Lawrence M. 1991b. Customer Discrimination and Affirmative Action. *Economic Inquiry* 29(3): 551-571.

Kahn, Lawrence M. 2000. The Sports Business as a Labor Market Laboratory. *Journal of Economic Perspectives* 14(3): 75-94.

McGarrity, Joseph, Harvey D. Palmer, and Marc Poitras. 1999. Consumer Racial Discrimination: A Reassessment of the Market for Baseball Cards. *Journal of Labor Research* 20(2): 247-258.

Nardinelli, Clark and Curtis Simon. 1990. Customer Racial Discrimination in the Market for Memorabilia: The Case of Baseball. *Quarterly Journal of Economics* 105(3): 575-595.

Pascal, Anthony H. and Leonard A. Rapping. 1972. The Economics of Racial Discrimination in Organized Baseball. In *Racial Discrimination in Economic Life*, ed. Anthony H. Pascal. Lexington, Mass.: D. C. Heath and Company.

Professional Sports Authenticator. 2003. *Sports Market Report*. Issue 104.

Scahill, Edward M. 2005. A Reinvestigation of Racial Discrimination and Baseball Cards. *Eastern Economic Journal* 31(4): 537-550.

Thorn, John, Pete Palmer, and David Reuther. 1989. *Total Baseball*. New York: Warner Books.

Thorn, John, Pete Palmer, and Michael Gershman. 2001. *Total Baseball*. Kingston, N.Y.: Total Sports Publishing.

Winfree, Jason A. 2010. Issues with Replicating Results in Sports Economics. *Journal of Sports Economics* 11(1): 48-59.

Ziliak, Stephen T. and Deirdre N. McCloskey. 2004. Size Matters: The Standard Error of Regressions in the *American Economic Review*. *Econ Journal Watch* 1(2): 331-358. **Link**

About the Authors

David W. Findlay is the Pugh Family Professor of Economics at Colby College where he has been a faculty member since 1985. He received his B.A. with Honours from Acadia University (Canada) and his Ph.D. in Economics from Purdue University. Two of his previously published articles examine the effects of player race and ethnicity on election to the National Baseball Hall of Fame. A native of New England, he is a lifelong fan of the Red Sox. His email address is dwfindla@colby.edu.

John M. Santos is Professor of Economics in the School of Business Administration and Morris Graduate School of Management at Robert Morris University in Chicago, Illinois, where he has been a faculty member since 1997. He teaches a wide range of economics courses including labor markets and gender. He earned his B.A. from Knox College and his Ph.D. in Economics from the University of Illinois at Urbana-Champaign. His current research focuses on discrimination in sports. A native Chicagoan, he grew up on the city's south side and still lives and dies with his White Sox. His email is jsantos@robertmorris.edu.

Robert Muñoz's reply to this article

Go to Archive of Comments section
Go to May 2012 issue

Discuss this article at Journaltalk:
http://journaltalk.net/articles/5760