EJW
ECON JOURNAL WATCH
Scholarly Comments on Academic Economics

ECON JOURNAL WATCH 9(2)
May 2012: A1-A18

Supplement to "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

We add to the literature on discrimination in professional sports by correcting, replicating, and extending the approach of Hewitt, Muñoz, Oliver, and Regoli (2005). This supplement and the accompanying article (**link**) offer results on the important question of whether player race or ethnicity has a significant effect on card prices.[3] Along the way, we examine several other issues related to the determinants of prices in the secondary market for baseball cards.

Hewitt, Muñoz, Oliver, and Regoli (hereafter HMOR) examine the effects of the race of the baseball player on the price of his rookie card. HMOR hypothesize that card prices may be affected by player performance, card availability, and player race. Because they remove five Hispanics from their sample, HMOR do not explore the effects of player ethnicity on card prices. 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*. HMOR (415) use Total Baseball Ranking (TBR) as their measure of player performance and report that their data source is the

1. Colby College, Waterville, ME 04901.
2. Robert Morris University, Chicago, IL 60605. We thank two anonymous referees for helpful comments and suggestions.
3. The exposition of this supplement presumes that the reader has first read the accompanying article.

2001 issue of *Total Baseball* (Thorn, Palmer, and Gershman 2001).[4] HMOR use data obtained from the April 2003 Professional Sports Authenticator (PSA) Population Report to measure card availability. After controlling for their measures of player performance and card availability, HMOR find that player race has no statistically significant effect on card prices.

Throughout our extensions, we follow the parsimonious approach of the model specified by HMOR in the sense that, like HMOR, we use a single summary or, equivalently, aggregate statistic to measure player performance. We extend HMOR's analysis in a number of important directions such as allowing card availability and price to vary by card quality. We also investigate an alternative aggregate measure of player performance other than career TBR. However, a consideration of disaggregated measures of player performance is beyond the scope of the article and supplement.

The Measurement Errors in HMOR's Data

In the article that accompanies this supplement, we document numerous discrepancies between the performance data reported in HMOR and their cited source, *Total Baseball* (2001). To briefly review, the TBR values for only 13 of the 51 players in their sample are equal to the values found in the 2001 edition of *Total Baseball*, while 38 of the 51 performance data observations reported in HMOR differ from those found in their reported data source. In other words, for nearly 75 percent of the players in their sample, the performance data reported and used by HMOR differ from their cited source.

Our Table 1 includes all information reported in HMOR's Table 1 and additional data in the final three columns. 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*. We also constructed two additional variables in Table 1 to measure the magnitude of HMOR's performance data errors. The *Difference in %* column includes the percentage point difference between the HMOR performance data and the performance data reported in *Total Baseball*. And finally, we obtained a performance ranking of all 51 players using the HMOR data and the data we obtained from *Total Baseball* (2001). The *Difference in Rank* column demonstrates how player rankings change when HMOR's incorrect data are used.

4. Player TBR data are reported both by year and by career; HMOR use career totals.

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.

We initially believed that the discrepancies in Table 1 were due to transcription errors but eventually discovered, after much investigation, that the TBR values for the other 38 players are equal to the values found in the 1989 edition of *Total Baseball* (Thorn, Palmer, and Reuther 1989). 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. In fact, this is precisely what HMOR did for their single performance statistic.

There is an additional measurement error that results from HMOR's pooling of the TBR values from the 1989 and 2001 editions issue of *Total Baseball*. For the 38 players in their sample for which HMOR obtained data from the 1989 issue, they have performance data only through the 1988 season. There are ten players in their sample whose careers continued *after* 1988; for five of these players, HMOR only have performance data through 1988. Thus while HMOR estimated card price equations based on card selling prices in 2003 that reflect the secondary market's valuation of the entire careers of the retired players, they have used incomplete career performance totals for nearly 10 percent of their sample. This measurement error implies that for these five players, the explanatory variable, career performance, is incomplete in the data that HMOR used to estimate their price equations. For four of these five players (Brett, Fisk, Ryan, and Yount), the sum of the TBRs for the remaining years of their respective careers is positive. For the remaining player (Schmidt) who played just one year beyond 1988, the 1989 value of his TBR is negative.

Another potentially serious measurement issue relates to the price series used by HMOR.[5] The authors state (416) that they use prices listed in *Beckett* (2003) that are "in near mint or mint condition." However, we discovered the prices they use are the HI prices reported by *Beckett*.[6] We also discovered that PSA reports price data for cards with different grades. So while HMOR create a card availability category (i.e., near mint or better) that aggregates or sums the availability of cards of different quality ratings, it is not obvious that the authors used the most appropriate price series for the aggregate card quality category they create. We return to this issue in our extensions section.

---

5. As we discuss in the article, there are several price data errors in HMOR. Gaylord Perry's 1962 Topps rookie card price is $80 (not $85) and Jackie Robinson's 1949 Leaf rookie card price is $1500 (not $1100). We also 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.

6. As we note in the article, 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.

A Replication and Correction of HMOR

The discovery of numerous data errors raises questions about the reliability of the estimated parameters in HMOR's price equations and, consequently, their key finding of no significant effect of player race on rookie card prices. 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," and they advise (97) that "in these cases, careful examination of the data and robustness estimation techniques can clearly improve the analysis."

To investigate the sensitivity of HMOR's findings, we estimate the same two specifications of HMOR's price equation:

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

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 and TPR for pitchers and 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 same two specifications of HMOR's price equation first using the data reported by HMOR and then using the corrected data reported in Table 1.

Table 2 displays estimates of the price equations (1) and (2) obtained using ordinary least squares regression. We obtain equations (2.1) and (2.2) when we use the reported data of HMOR. The results for equation (2.1) are identical to those reported by HMOR. While they do not report all of the estimates of the price equation that includes the interaction term, our results are identical to those HMOR discuss in their paper. HMOR, for example, report the R^2 statistic (0.27) only for equation (1). We report both the R^2 and adjusted R^2 statistics in Table 2.

TABLE 2. 2003 Card Price Equations

	Replications		Corrections		Alternative Model Specifications			
	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)	(2.6)	(2.7)	(2.8)
Dependent Variable	*Beckett*	*Beckett*	*Beckett*	*Beckett*	PSA8	PSA8	PSA8	PSA8
Availability Variable	# 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 PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher
Source of Availability Data	HMOR	HMOR	HMOR	HMOR	HMOR	HMOR	HMOR	HMOR
Independent Variables:								
Intercept	6.356*** (7.42)	6.125*** (6.82)	6.265*** (7.43)	6.101*** (6.93)	7.799*** (8.96)	7.678*** (8.42)	7.336*** (8.45)	7.481*** (8.30)
Availability	−0.404** (2.51)	−0.416** (2.57)	−0.411** (2.60)	−0.420** (2.63)	−0.592*** (3.62)	−0.600*** (3.62)	−0.543*** (3.44)	−0.541*** (3.40)
Performance	0.027*** (3.07)	0.035*** (2.76)	0.030*** (3.44)	0.036*** (2.96)	0.033*** (3.62)	0.037*** (2.96)		
Performance Per Season							0.746*** (4.25)	0.665*** (3.06)
Black	−0.124 (0.34)	0.485 (0.62)	−0.098 (0.27)	0.385 (0.49)	−0.089 (0.24)	0.269 (0.33)	−0.165 (0.46)	−0.670 (0.78)
Black*Performance		−0.015 (0.87)		−0.012 (0.68)		−0.009 (0.49)		
Black*Performance Per Season								0.236 (0.64)
Summary Statistics:								
R^2	0.27	0.29	0.30	0.31	0.38	0.39	0.43	0.44
Adjusted R^2	0.23	0.22	0.26	0.25	0.34	0.33	0.39	0.39
F	5.93	4.61	6.83	5.18	9.72	7.23	11.8	8.86
N	51	51	51	51	51	51	51	51

Notes: [a] *Beckett* equals the *Beckett* HI price. [b] PSA7 equals near mint. [c] PSA8 equals near mint-mint. [d] PSA9 equals mint. [e] The numbers in parentheses are the absolute values of the t-statistics. [f] * indicates $p < 0.10$; ** indicates $p < 0.05$; and *** indicates $p < 0.01$.

When we re-estimate the price equations (1) and (2) using the corrected performance and price data, we obtain equations (2.3) and (2.4) in Table 2. Aside from the slight increase in the R^2 and F statistics, the general pattern of results is similar to that reported in HMOR and replicated here. 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. The negative and significant estimated coefficient on the card availability variable indicates, as expected, that increases in card availability lower the selling price of a player's rookie card. As both Availability and Price are

measured in log form, the estimated coefficient on Availability in equation (2.3) implies that a 10 percent reduction in card availability raises card selling price by 4.1 percent. Using equation (2.4), the comparable prediction is a 4.2 percent rise in card price.

More importantly, none of the estimated coefficients on the race variables, Black or Black*Performance, in equations (2.3) and (2.4) is statistically significant indicating that, holding player performance and card availability constant, race has no effect on card prices. Thus the corrections in Table 2 strengthen our confidence in HMOR's finding of no statistical evidence of customer discrimination against blacks in the market for baseball cards.[7]

Extensions of HMOR

Even using the corrected data, HMOR's estimated price equation (2.3) only explains 30 percent of the observed variation in rookie card prices.[8] This sent us in search of alternative models with more explanatory power while simultaneously maintaining the parsimonious spirit of the original model specified by HMOR.

Matching Price and Card Quality

As we discuss in the article, PSA determines the authenticity and quality of cards submitted. That is, PSA will grade (i.e., categorize) cards 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. HMOR create an availability variable that 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.[9]

HMOR (416) state that they use prices listed in *Beckett* (2003) that are "in near mint or mint condition." However, as we previously noted, the prices they use are the HI prices reported by *Beckett*. We also discovered that PSA reports price data for cards with different grades. The average selling prices listed by PSA correspond

7. We also examined, for several reasons, whether our results were affected by the inclusion of Jackie Robinson in the HMOR sample. First, Jackie Robinson was an epic figure who transcends baseball. Second, unlike the vast majority of the players in the sample, he started his professional career in the Negro Baseball League. And finally, as we discovered, HMOR incorrectly used the Bowman rookie card price rather than the Leaf rookie card price. Therefore, HMOR may have mismatched Robinson's Bowman price with Robinson's Leaf availability data. 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.

8. Equation (2.3) in Table 2 of this supplement is the same as equation (3) in the article.

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. Thus 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.

While HMOR create a card availability category (i.e., cards rated near mint or better) that aggregates or sums cards of different quality ratings, the authors 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 data 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, which enables us to measure card availability by specific card quality ratings as reported by PSA. This in turn allows us to explore the relationship between card availability and card prices across card quality categories.

We re-estimate equations (1) and (2) using alternative price series beyond the *Beckett* HI price series chosen by HMOR. We first use the 2003 LO prices obtained from *Beckett* (2003). The results using this alternative measure of card prices are similar to those obtained using *Beckett*'s 2003 HI prices reported in equations (2.3) and (2.4), so we do not report them here. We then obtained three different PSA price series: PSA7 (prices of near mint cards), PSA8 (prices of near mint-mint cards), and PSA9 (prices of mint cards).[10] When we again estimate the basic price equations (1) and (2) using these three PSA price series, we obtain a similar pattern of results for the Performance, Black, and Black*Performance variables. However, several other results are noteworthy. The absolute values of the estimated coefficient and the t-ratio for the Availability variable both increase when any of the

9. The number of cards reported by PSA represents the number of times PSA has evaluated a player's particular card for each quality category. Each time an individual submits a card to PSA, PSA will update its population report. While the PSA data do measure the frequency with which each player's card has been evaluated by PSA, the data do not measure the number of cards that are "reported to *exist* [emphasis added]." 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. However, we must note that 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.

10. We obtained the three different price series from the April 2003 PSA *Sports Market Report*. We purchased this particular issue to ensure that the date of the price data would be consistent with the date (April 2003) of the availability data used by HMOR and included in our Table 1.

three PSA price series are used. Additionally, both R^2 statistics and the F-ratio are higher when we replace the *Beckett* HI price with any of the three PSA price series.

We find that HMOR's aggregated card availability category has the most explanatory power (measured by either R^2 statistic) when paired with the PSA8 price series—the price of cards that PSA grades as near mint-mint quality. We obtain equations (2.5) and (2.6) in Table 2 when the dependent variable is the PSA8 price. Comparing equations (2.3) and (2.5), the percent of total variation in card prices explained by HMOR's model increases from 30 to 38 percent (26 to 34 percent using the adjusted R^2 statistic) when prices are measured using the PSA8 series instead of the *Beckett* HI series.[11] We observe a similar increase in explanatory power when we compare equations (2.4) and (2.6) which include the interaction term (Black*Performance).

For both equations (2.5) and (2.6), the Availability variable is now significant at the .01 level (compared to the .05 level for when we use the Becket HI price series). Furthermore, the estimate of the effect of card availability on card prices is larger when price is measured using the PSA8 price series. Comparing the estimated coefficients on Availability in equations (2.3) and (2.5), we now predict that a 10 percent reduction in Availability causes a 5.9 percent increase in card prices in contrast to a 4.1 percent increase for the *Beckett* HI price measure used by HMOR. Comparing equations (2.4) and (2.6) which include the interaction term, we now estimate that the same 10 percent reduction in Availability causes a 6.0 percent increase in card prices when using the PSA8 price series compared to a 4.2 percent increase for the *Beckett* HI price measure used by HMOR.[12]

Our results suggest that one's use of the card availability data reported by PSA requires an analysis of the sensitivity of any results to the selection of the price data series. Our initial results suggest that the quality of the cards that are included in the availability data should, whenever possible, be matched with the corresponding price data series for that specific card quality. In the case of HMOR, we find that the aggregate availability measure they create fits any of the three PSA price series better than the *Beckett* HI price series that they chose. And the best fit apparently is the PSA8 price series which corresponds to cards of near mint-mint quality.

---

11. Though we do not report the estimated price equation in Table 2, we obtain an $R^2 = .37$ when we re-estimate equation (2.3) using the PSA7 price series (corresponding to cards graded near mint quality) instead of the *Beckett* HI price series used by HMOR.

12. For the remainder of our extensions, we focus most of our discussion on the models without interaction terms, both because the results are nearly identical for the equations with and without interaction terms and because the addition of the interaction term(s) leaves the adjusted R^2 either unaltered or slightly lower, implying that any gain in explanatory power as measured by the unadjusted R^2 is offset by a loss of degrees of freedom in the sample.

We have so far used an aggregate measure of card availability as defined and reported by HMOR. However, for the 2010 data, we were able to obtain real-time availability data (in April 2010) by card quality levels as reported by PSA. So we are able to create card availability categories, both by aggregated card quality levels (a la HMOR) and for disaggregated (i.e., specific) card quality levels. Given that the availability data reported by PSA may serve as a proxy for the number of cards of a given quality level that exist in the secondary market for rookie cards, we exploit the availability data set we construct to estimate separate card price equations for different categories of card quality.

For reference, equations (3.1) and (3.2) in Table 3 are the results when we re-estimate HMOR's basic model (equations (1) and (2)) using their pairing of price and availability measures (*Beckett* HI price paired with the count of cards graded PSA7 or higher) for the 2010 data. We find that HMOR's model fits the 2010 data better than the 2003 data. Comparing the R^2 statistics, equation (3.1) explains 49 percent of the variation in rookie card prices as compared to 30 percent explained by equation (2.3) for the 2003 data. As was the case for the 2003 data, we find the HMOR's availability measure (cards rated PSA7 or higher) has more explanatory power when paired with any of the PSA prices than when matched with the *Beckett* HI price.[13] And we again find that HMOR's availability measure has the most explanatory power (measured by either R^2 statistic) when paired with the PSA8 price series: the R^2 statistic rises from .49 to .61 when we replace the *Beckett* HI price with the PSA8 price. This consistent finding across the 2003 and 2010 data suggests that the best price measure to pair with HMOR's availability measure (all cards of near mint or better quality) is the PSA8 price (cards of near mint-mint quality).[14]

However, when we disaggregate card availability by specific card quality and pair with the corresponding quality-specific PSA price series, the percent of variation in rookie card prices explained (as measured by R^2) rises even more—to 66 percent for the PSA8 price-PSA8 availability pairing and to 69 percent for the PSA9 price-PSA9 availability pairing. Equations (3.3) and (3.4) in Table 3 are the regression results when we re-estimate equations (1) and (2) using the PSA9 price-PSA9 availability pairing.

13. Table 4 reports the unadjusted R^2 values for most of the model specifications (without interaction terms) that we estimate including many price equations for which space constraints do not allow us to report full regression results in Tables 2 and 3.

14. We could not estimate PSA7 card price equations because there did not exist a PSA7 card price for every player in our sample.

TABLE 3. 2010 Card Price Equations

	Blacks and Whites Only				Blacks, Whites and Hispanics			
	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)	(3.6)	(3.7)	(3.8)
Dependent Variable	*Beckett*	*Beckett*	PSA9	PSA9	*Beckett*	*Beckett*	PSA9	PSA9
Availability Variable	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA9	# of Cards Rated PSA9	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA7 or Higher	# of Cards Rated PSA9	# of Cards Rated PSA9
Source of Availability Data	PSA	PSA	PSA	PSA	PSA	PSA	PSA	PSA
Independent Variables:								
Intercept	9.250*** (9.58)	9.142*** (8.99)	9.859*** (20.35)	9.872*** (18.11)	9.211*** (9.68)	9.212*** (9.21)	9.839 *** (20.8)	9.915*** (18.5)
Availability	−0.738*** (5.32)	−0.737*** (5.26)	−0.895*** (8.79)	−0.894*** (8.69)	−0.744*** (5.42)	−0.748*** (5.43)	−0.906*** (9.08)	−0.907*** (9.01)
Performance	0.027*** (3.31)	0.029** (2.65)	0.027*** (3.63)	0.026** (2.59)	0.029*** (3.68)	0.029** (2.68)	0.028*** (3.97)	0.026** (2.62)
Black	−0.027 (0.08)	0.210 (0.30)	0.070 (0.24)	0.039 (0.06)	−0.037 (0.12)	0.214 (0.31)	0.067 (0.23)	0.048 (0.07)
Black*Performance		−0.006 (0.38)		0.001 (0.05)		−0.006 (0.38)		0.001 (0.05)
Hispanic					0.401 (0.79)	−0.722 (0.64)	0.445 (0.95)	−0.570 (0.54)
Hispanic*Performance						0.044 (1.16)		0.039 (1.10)
Summary Statistics:								
R^2	0.49	0.49	0.69	0.69	0.49	0.51	0.69	0.69
Adjusted R^2	0.45	0.44	0.67	0.66	0.45	0.45	0.66	0.66
F	14.89	11.00	34.25	25.14	12.19	8.37	28.0	18.6
N	51	51	51	51	56	56	56	56

Notes: [a] *Beckett* equals the *Beckett* HI price. [b] PSA7 equals near mint. [c] PSA8 equals near mint-mint. [d] PSA9 equals mint. [e] The numbers in parentheses are the absolute values of the t-statistics. [f] * indicates $p < 0.10$; ** indicates $p < 0.05$; and *** indicates $p < 0.01$.

TABLE 4. 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). [f] Replication of HMOR model in Row (1) corresponds to equation (2.1) in Table 2; HMOR Model with Corrected Data in Row (2) corresponds to equation (2.3) in Table 2. [g] Column (3) of Alternate Model 1 in row (3) corresponds to Equation (2.5) in Table 2; columns (5) and (8) of Alternate Model 1 in row (3) correspond to equations (3.1) and (3.3) in Table 3. [h] Columns (3) and (8) of Alternate Model 2 in row (4) correspond to equations (3.5) and (3.7) in Table 3. [i] Column (8) of Alternate Model 3 in row (5) corresponds to equation (2.7) in Table 2.

As was the case with the 2003 data, the estimated effect of Availability on card prices is greater when price is measured using the PSA series. Using the 2010 data and HMOR's availability measure, equation (3.1) implies that a 10 percent reduction in availability causes a 7.4 percent increase in price when the *Beckett* HI price series is used. However, when we match the PSA8 price series with HMOR's availability measure, the predicted price response to the same 10 percent reduction

in availability rises to 9.8 percent.[15] And when we disaggregate card quality and match to the corresponding quality-specific PSA prices, we estimate price responses of similar magnitudes to changes in card availability. Using the PSA9 card price-availability pairing in equation (3.3), we predict that a 10 percent reduction in availability of PSA9 cards raises PSA9 card prices by almost 9 percent. And for the PSA8 card price-availability pairing, we estimate that the same 10 percent reduction in availability of PSA8 cards raises PSA8 card prices by 9.7 percent.[16]

Our results for the 2010 data, for which we are able to measure card availability both by aggregated and disaggregated quality categories, suggest that the choice of price series is nontrivial, depending on the quality of the card availability category one uses or creates. We find that when we appropriately match the PSA availability measure to the corresponding PSA price series, we can predict—across PSA card quality categories—that a 10 percent reduction in availability in the secondary market causes an increase in card prices in the range of 9 to 10 percent. Furthermore, when we carefully match price to card quality and availability, we obtain price equations with more explanatory power.

And finally, with the use of the more recent 2010 data, we can examine whether the effects of player race change over time. The statistically insignificant coefficients on all Black variables indicate that player race continues to have no effect on card prices in the more recent period.

Does Ethnicity Matter?

HMOR did not explore whether their finding on race extends to ethnicity, instead choosing to omit the five Hispanics from their sample.[17] For the 2010 data for which we have access to availability data, we added the five Hispanic players to our sample, raising the sample size of players from 51 to 56. We then created a dummy variable that equals one if the observed player is Hispanic and equals zero otherwise. The reference group in our sample is white players. We estimated this modified equation for the alternative price series-card availability pairings previously discussed. Equations (3.5) and (3.6) are the results when we use the card price-card availability pairing used by HMOR while Equations (3.7) and

15. When we re-estimate equation (3.1) again using HMOR's availability measure but after replacing the *Beckett* HI price series with the PSA8 price series corresponding to cards of near mint-mint quality, we obtain an estimated coefficient on (log of) availability of −.9752 with an absolute value of the t-ratio equal to 7.06.

16. Though we do not report in Table 3 the estimated price equations for the two cases of PSA8 price-PSA7 or higher availability combination and the PSA8 price-PSA8 availability pairing, the full regression results for these two equations are available from the authors upon request.

17. The five Hispanic players are Luis Aparicio, Rod Carew, Roberto Clemente, Juan Marichal, and Tony Perez.

(3.8) are the comparable results when we use the PSA9 card price-PSA9 availability pairing.

When we compare equations (3.5), (3.6), (3.7) and (3.8) that include Hispanics to the corresponding equations (3.1), (3.2), (3.3) and (3.4) that exclude Hispanics, we find very little change in the magnitude or the statistical significance of the estimated coefficients on performance and card availability. As was the case for the sample without Hispanics, the R^2 statistics and the estimated coefficient of the availability variable are higher when we use the PSA9 price-PSA9 availability pairing rather than the *Beckett* HI price-PSA7 or higher availability combination. Though we do not report the equations in Table 3, the pattern of results is very similar to equations (3.7) and (3.8) for the PSA8 price-PSA8 availability pairing.

Additionally, race remains statistically insignificant across all price-availability pairings. Just as importantly, we find that there is no statistically significant effect of ethnicity on card prices. When we allow the effects of race and ethnicity to vary with player performance in equations (3.6) and (3.8), the Black and Black*Performance interaction variables continue to be insignificant. While the estimated coefficient on the Hispanic*Performance interaction variable is larger than its standard error, the t-statistic falls below critical values at conventional significance levels.[18]

Although statistically insignificant, the coefficients on the Hispanic dummy variable in equations (3.5) and (3.6) are positive. We suspected that the unique status of Roberto Clemente might be influencing the sign and the magnitude of the Hispanic coefficients: Clemente's career was cut tragically short when he died in a plane accident on December 31, 1972. We decided to re-estimate the equations after removing Clemente from the sample and we found that the estimated coefficients on the Hispanic dummy variable become negative yet remain statistically insignificant (with t-statistics close to zero). For example, when we re-estimate equation (3.5) for the sample that excludes Clemente, the estimated coefficient on the Hispanic dummy variable now equals −0.009 with a t-statistic equal to −0.02. In general, our corrections and extensions show that neither race nor ethnicity has a statistically significant effect on card prices. These findings are robust to alternative price-availability pairings.

18. We also construct a dummy variable, Nonwhite, that equals one if a player was Black or Hispanic. When we replace the black and Hispanic variables with Nonwhite, we find that the estimated coefficients on Nonwhite and on the Nonwhite*Performance interaction variable are never statistically significant. All other results are virtually identical to those reported in Table 3.

Alternative Model Specifications: Per Season Performance and Year of Rookie Card

So far, we have investigated whether the effect of availability on card prices varies by card quality. We consider two additional model specification issues here. 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. But perhaps the underlying demand for these cards depends on how efficiently the player achieved his career totals. As the examples we discuss in the article illustrate, some hitters and pitchers in our sample achieved approximately the same career TPR and TPI, respectively. However, we observe clear differences in career performance when we measure performance on a per season basis. So we ask: Do fans focus on players' career totals or on players' career per season averages?

To investigate this issue, we create a per season performance measure, either pitcher's TPI per season (TPI/seasons played) or hitter's TPR per season (TPR/seasons played), and re-estimate the previously discussed equations after replacing the career performance totals (TPR and TPI) with the corresponding per season measures. Whether performance is measured by career totals or by per season averages appears to matter for the 2003 data. The percent of variation in card price explained (as measured by either R^2 statistic) by HMOR's model rises modestly for each of the price series that we pair with HMOR's availability data when lifetime player performance is measured on a per season basis.

For example, when we re-estimate equations (2.3) and (2.4) in Table 2 after replacing the career performance totals with the corresponding per season measures, R^2 rises from .30 to .36 for the equation without any interaction term, and from .31 to .36 when we include the interaction term. And when we pair the PSA8 price series with HMOR's availability variable and re-estimate equations (2.5) and (2.6) using the per season performance measure, we obtain equations (2.7) and (2.8) in Table 2. Comparing equations (2.5) and (2.6) to equations (2.7) and (2.8), we again observe a modest increase in explanatory power when performance is measured on a per season basis, with R^2 rising from .38 to .43 for the equation without any interaction term, and from .39 to .44 when we include the interaction term. We also note that the t-ratios on the career per season performance variables in equations (2.7) and (2.8) are somewhat larger than the corresponding t-ratios on the career total performance variables (2.5) and (2.6). The results for 2003, therefore, provide some evidence for the hypothesis that baseball card market participants focus on career per season performance rather than on career totals.[19]

For the 2010 data, however, it does not appear to matter much whether performance is measured by career totals or by per season averages. For example, when we re-estimate equation (3.1) in Table 3 after replacing the career perfor-

mance totals with the corresponding per season measures, R^2 rises from .49 to .50, but when we re-estimate equation (3.3) in Table 3 after replacing the career performance totals with the corresponding per season measures, R^2 falls from .69 to .68. Similarly, when we re-estimate the other equations in Table 3 using per season averages over the player careers rather than career totals, in each case, both R^2 statistics remain largely unchanged.[20] Thus, we do not have a consistent answer across data sets as to whether baseball card market participants focus more on career totals or on per season career averages.

We close with a finding that, while not surprising, both authors of this supplement nonetheless found interesting. Using the 2010 data set that includes the Hispanic players, we introduce a variable, Card Year, to see if the year in which the player's rookie card was issued matters for card prices. In addition to including this new variable, we also include it as an interaction term with the Black, Hispanic, and Nonwhite dummy variables. While none of the race, ethnicity and race/ethnicity interaction coefficients is statistically significant, the estimated coefficient on the Card Year variable is negative and significant in all equations. This result indicates that, holding other factors such as performance and card availability constant, older rookie cards command a higher price in the secondary market.[21]

Conclusions

Using HMOR's parsimonious approach of measuring player performance with a single summary statistic, we first replicate their results and then verify that their central finding of no significant effect of player race on baseball card prices holds after we correct both the measurement errors in their performance variable

19. We also re-estimate equations (2.5) and (2.6) after adding the number of seasons as a separate explanatory variable. In the two price equations that include both career performance totals and the number of seasons played to achieve the totals, the estimated coefficients on the seasons variable are negative and marginally significant (the absolute values of the t-ratios are equal to 1.74 and 1.65) and similar increases in R^2 are obtained. Although the estimated coefficients on the seasons variable are only marginally significant, these results also suggest that rookie card buyers may focus more on how efficiently a player achieves a given level of performance rather than focus solely on career totals. While we do not report the two price equations that include both performance totals and the number of seasons, the results are available from the authors upon request.

20. Table 4 reports the (unadjusted) R^2 values for all of the specifications of the basic price equation (without interaction terms) that we estimate in which career performance is measured on a per season basis.

21. While not the primary focus of our study, we examined several final issues. First, the rookie cards of more recently inducted Hall of Fame players have lower prices. And second, rookie card prices do not appear to be affected by the identity of the firm that produced the cards. In all of these equations that take into account the effects of the year in which the player was inducted to the National Baseball Hall of Fame and the identity of the card's producer, race and ethnicity have no effect on rookie card prices.

and their price data errors. Our 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. In short, our results paint a hopeful picture of a market in which card prices reflect market fundamentals, independent of a player's race or ethnicity.

Additionally, when we extend HMOR's model by allowing price and availability to vary with card quality, our results indicate that the choice of price series depends nontrivially on the card availability measure one uses or creates. Using HMOR's aggregate availability measure we find, for both the 2003 and 2010 data, better fits with any of the PSA price series than with the *Beckett* HI price series that HMOR chose. And when we are able to measure card availability by disaggregated quality categories and correctly match to the corresponding quality-specific price, we obtain equations that better fit the data. More generally, when we carefully match price to card quality and availability, we obtain price equations with more explanatory power.

Appendix

An Excel file (**link**) contains data used in the analysis.

References

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

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.

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

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.

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.

Go to Article homepage