Pay, productivity and aging in Major League Baseball

Turner, Chad and Hakes, Jahn

31. July 2007

Pay, Productivity and Aging in Major League Baseball

Jahn K. Hakes
Albion College

and

Chad Turner
Nicholls State University

July 31, 2007

Contact information:

Hakes: jhakes@albion.edu. 517-629-0419.
Address: Dept. of Economics, Albion College, 611 E. Porter St., Albion, MI 49224.

Turner: chad.turner@nicholls.edu. 985-448-4194.
Address: Dept. of Finance & Economics, Nicholls State University, Thibodaux, LA 70310.

Abstract

Using panels of player pay and performance from Major League Baseball (MLB), we examine trends in player productivity and salaries as players age. Pooling players of all ability levels leads to a systematic bias in regression coefficients. After addressing this problem by dividing players into talent quintiles, we find that the best players peak about two years later than marginal players, and development and depreciation of ability appear to be more pronounced for players with the highest peak ability levels. Within-career variation, however, is less pronounced than between-player variation, and the talent level of players within a given quintile will typically remain lower than the talent level for rookies in the next higher quintile. Free agents are paid proportionately with their production at *all* ability levels, whereas young players' salaries are suppressed by similar amounts.

Keywords: Major League Baseball (MLB); career dynamics; player salaries and performance; quintile analysis

I. Introduction

The correspondence between salaries and worker productivity is a central issue in labor economics. Due to the abundance of performance and salary data for athletes in professional team sports, a great deal of empirical work has studied salary patterns in baseball and other sports, with sub-fields examining salary discrimination and the salary effects of arbitration and free agency, among other topics.[1] Although this path has been trod frequently, we believe there is still room in the literature for exploration yielding new insights.

Our contribution is to use quintile analysis to examine how productivity patterns vary between cohorts of players with similar ability and see how well the labor market accommodates ability variation in setting salaries. To illustrate the extent of the bias produced by traditional ordinary least squares (OLS), we run regressions estimating separate salary and performance paths for each talent quintile. This technique will reduce the tendency of pooled regressions of salary or productivity on experience to yield a "flatter" time profile than is actually the case.[2]

The primary advantage to studying productivity in MLB is the abundance of quantitative data on multiple components of individual performances. Sabermetricians, statistically-minded members of the Society for American Baseball Research (SABR), have produced a great number of performance metrics derived from these components in an attempt to quantify production on a single dimension. The peformance measure we utilize is OPS (on-base percentage plus slugging average). While some of the sabermetric alternatives are preferable to OPS in specialized

---

[1] In baseball, the majority of this research stems from the seminal work of Scully (1974). Kahn (1991) summarized the early racial discrimination literature, which has since expanded to include analyses of coaching discrimination in MLB (Singell [1991]), the NFL (Madden [2004]), and the NBA (Humphreys [2000], Kahn [2006]). Marburger (1994, 2004) looks at final offer arbitration in MLB, while Zimbalist (1992), Kahn (1993), Vrooman (1996) and Miller (2000), among many others, look at the salary effects of free agency in baseball.

[2] We note this technique is not the same as quantile regression, although both terms have a common element of placing greater analytical weight on particular centiles of a distribution.

scenarios, the more widespread acceptance of OPS is worth the potential very small cost in efficiency for our purposes.

In Section II, we discuss the econometric problems involved in career path estimations, and how quintile analysis mitigates some of those problems. Section III discusses the data set used and our empirical method. Section IV looks at player productivity over the span of a career and Section V presents career salary paths. Section VI offers findings on the crude relationship between pay and performance in each league, while Section VII summarizes our findings and details our plans for further research.

II. Variation of Individual Characteristics Over Time

II.A. Career Paths in Productivity

Ordinary least squares regression is a tool so familiar to empirical economists that its use to study relationships such as that between player salaries and years of experience is virtually reflexive. We do not wish to combat the intuition that regression analysis is an appropriate tool here – especially as we will use that technique for our own modeling – but rather to note an observational deficiency in the data that in the absence of correction would lead to biased estimators.

We can model the abilities of players in a top tier professional sports league as individuals in the extreme right-hand tail of the ability distribution for playing that sport. To obtain a roster spot, a player must have an ability level above "replacement level." A replacement level player has ability only marginally above that of the top player not contracted to play in the league, and typically is a benchwarmer or part-time fill-in at the top level, or a top player or prospect in a minor or secondary league.[3] Typically, pay and productivity are only recorded for those players who have been selected into the top tier league.

In the absence of entry barriers, a young player developing his talent would be able to hold a spot on a professional roster once his ability reached replacement level, and as he further developed he would continue to remain employed until diminishing skills (due to injuries or age) rendered him once again below replacement level. It has been solidly established that baseball hitting ability climbs until peaking at about age 27, and slowly trails off thereafter.[4] Adapting that methodology to other sports has led to similar findings, with minor variations of peak age, but in each case describing an age function that resembles an inverted-"U" and which is commonly estimated as a quadratic function. In equation (1), this would be noted as $\beta_1 > 0$, and $\beta_2 < 0$, where α is the vertical intercept, and ε_i is a stochastic error term for player *i*.

$$\text{Ability}_i = \alpha + \beta_1 \text{ age}_i + \beta_2 \text{ age}_i^2 + \varepsilon_i \qquad (1)$$

We can speak of a player's career path in productivity, or "ability path" as the locus of Ability_i level over the relevant range of ages, and if the correct function form for equation (1) is indeed quadratic, we can then discuss the shape of ability paths in terms of the parameters α, β_1, and β_2.

To date, we know of no studies searching for or finding systematic differences in the ability paths of star players (the highest peak ability), journeyman players (of middling ability), and marginal players (only slightly above replacement level). The typical practice is to pool all players when estimating equation (1), although Krautman (1993) did acknowledge the estimation problems posed by inherently different ability levels. The one notable exception is Schulz et al.

---

[3] Many players in baseball make frequent journeys back and forth between the top minor league level (AAA) and the "show," earning them the label of "4-A" players.

[4] See James (1982), Krohn (1983), Albert (1999), and Schulz et al. (1994) for estimation of production peaks and Fort (1992) and Horowitz and Zappe (1998) for salary estimation models.

(1994), who found the abilities of their group of elite players declined more slowly after the peak age than did those of non-elite players. Our paper extends the approach of Schulz et al. by using a more detailed segmentation of the player population, applying the technique to salaries, and considering the economic implications of the resulting estimates.

Pooled regressions implicitly assume that the only structural difference between a star player and a marginal player is a higher value for the intercept term, α, for the star. That assumption leads to the implicit implications that (a) player ability improves or declines by the same amount at a given age, regardless of ability cohort, (b) players of all ability levels hit their "peak" at the same age, (c) star players reach replacement level at an earlier age than marginal players, and (d) that star players will remain above replacement level longer than marginal players, and will be forced out at later ages.

Pooled empirical estimation not only leaves testable implications unexamined, but results in biased coefficients as well. The mechanism for selecting observations is based on whether ability is higher than the replacement level threshold. While all professional players will be in the sample at their peaks, yielding an average ability level at the peak age that represents the central tendency of all players, only the stars' abilities will be sufficiently high for their statistics to be observed at younger and older ages, leading to a positive observational bias that increases in magnitude as the distance (in years) from the peak age increases. This theoretical effect is illustrated in Figure 1, and shows that an estimated regression line from pooled player data will tend to underestimate the magnitudes of the β_1 and β_2 terms in Equation (1).

On the left-hand side of the relevant range of ages, this effect might be somewhat lessened in some leagues (particularly the NFL and NBA) by entry barriers ostensibly designed to keep players out of the league until they have attended college. Though the magnitude of the mismeasurement would be reduced as a result, the bias from aging players in the right tail would persist. If the bias is originating mostly from the right tail of the age distribution due to left-tail truncation, we would also expect estimated peak ages to be biased upward.

II.B. Career paths in salary

Estimation of player pay has followed a substantially different methodology than that used for player performance, with most models estimating salaries directly as a function of productivity or ability.[5] Our purpose here, however, is to establish time paths for individual player salaries as a function of experience (exp_i). Therefore we will formulate an empirical equation with a structure similar to the ability equation above.

$$\ln(\text{salary})_i = \alpha + \beta_1\, exp_i + \beta_2\, exp_i^2 + \boldsymbol{\gamma}'\mathbf{Pos_i} + \varepsilon_i \qquad (2)$$

Equation (2) has a right hand side analogous to the salary model, although it also includes a vector of control variables indicating player *i*'s defensive position and it uses the natural logarithm of salary as the dependent variable to preserve normality of the residual terms. In comparison with traditional salary models, this model is extremely sparse. It lacks not only the normal control variables for negotiating freedom, market size, player awards, etc., but also the variable which typically is the primary regressor – ability.[6]

The use of experience rather than age is necessitated by the nature of baseball's collective bargaining agreements. Aside from ability, the most important predictor of player salary is freedom to contract, which has been defined in the league agreements as a function of playing experience. Consequently, our salary fits are far more efficient using experience. While we

[5] As we are focusing upon improved measurement of the individual components, the pay to performance ratios we construct will be somewhat crude. However, we will compare our broad-brush results to those from adaptations of the canonical Scully (1974) model in Section VI.

[6] The division of players into ability quintiles allows a crude control for ability. We include indicator variables for position as a concession to identifying player MRP, as players with similar offensive productivity at more challenging positions are demonstrating ability along another dimension and should merit a salary premium.

begin our analysis by estimating fits for ability using age in order to illustrate our premise of replacement level observation bias, we will subsequently estimate performance equations using the experience measure in order to allow subsequent comparisons between pay and performance using a common time dimension.[7]

II.C. Correcting for observational bias

If we were to estimate equation (2) using a pool of all players, it would be subject to the observational bias we described in section II.A. The approach we use to reduce this bias is to divide our sample of players into ability cohorts that are more homogeneous than the pooled sample. As the variation in ability is reduced within each cohort, the resulting estimates are subject to less bias caused by differences in the age at which the players rose above or fell below replacement level. The estimated β_1 and β_2 terms should then be larger in magnitude, with estimates that consistently approach the true parameter value as the number of equally-sized cohorts approaches infinity.

Prior to estimating equation (2), however, we must ascertain whether the semilog quadratic model is appropriate. Two sources of potential difficulties are readily apparent. First, on the left side of the age range, reserve clauses and other restrictions on negotiating freedom for inexperienced players would be expected to lead to depressed salaries during the early stages of player careers. Indeed, measuring the extent of salary suppression is one of the main objectives of the pay-and-performance literature. Separate regressions for free agents and players under restricted bargaining may present significant improvements in the accuracy of estimates.[8]

---

[7] As could be predicted, the effect of the switch from age to experience in productivity equation (1) is that higher-ability players peak later in their careers, as they debut at younger ages, allowing more playing time prior to their peak age.

[8] Fort (1992) estimated separate parabolic arcs for salary trends of players of below mean and above mean age. Although our differing arcs "hinge" at the point of free agent eligibility rather than mean age, the econometric objective is the same.

Second, whereas player ability has been empirically determined to fade as a player ages past his peak, empirical evidence of player salaries falling as they age has not been as well documented. One exception to this has been Horowitz and Zappe (1998), who concluded both that "once the average player has put in his 9 years ... eroding skills result in lower pay," and that the effect was much smaller for former-star players. This may be because star players refuse to accept salary cuts as their contracts come up for renewal after a performance decline, leading to observational selection with a short lag, or that perhaps the marginal revenue product (MRP) of star players – the theoretical source of employer willingness-to-pay for the player's services – doesn't decline as ability does.[9] Upon observing the data we will assess the appropriateness of the functional form of equation (2).

III. Data

The Baseball Archive database, version 5.3, edited by Sean Lahman (www.baseball1.com) contains season-by-season data on player performance, salaries, and many other variables that would serve as useful controls in a structural analysis. As salary data is only available for the years 1985 to 2005, we set that as our sample period.

In order to compare productivity across a wide range of positions, we focus on hitting productivity, and have eliminated pitchers from our sample.[10] Our preferred measure of hitting ability is OPS (on base percentage plus slugging), which has been shown to be both simple to calculate and an accurate predictor of team output (wins). As OPS measures production per unit of playing time, whereas the chief sabermetric alternative, runs created, is a count statistic of

[9] The latter possibility allows for veteran fan favorites and stars that either retain their popularity and drawing power through the end of their careers, or watch their fame decline more slowly than their fading skills, as was the case with the end-of-career versions of Cal Ripken, Jr. and Willie Mays. Although the discussion is beyond the scope of this paper, there is also a large literature in personnel economics that studies reasons why salaries may vary from MRP under competition. Common theories of this sort include efficiency wages to combat shirking and incentives for career contracting.

[10] Defensive ability has proven both difficult to measure and, consequently, difficult to establish as a significant predictor of salaries except (weakly) through simple binary variables for defensive position.

production that increases with playing time, the use of OPS provides conservative estimates of within-career variation.

Although ability as measured by OPS is not subject to general price inflation or reserve clauses as salaries are, league-wide OPS figures do vary over the period of our sample. The annual league-wide OPS ranges from 0.707 in 1988 to 0.796 in 2000, so that indexing by annual average is necessary. Furthermore, there are large differences in OPS across defensive positions, as the average OPS for first basemen is more than 125 points higher than that for shortstops (0.827 to 0.700). Failure to account for the positional differences would mean that shortstops and catchers would be significantly over-represented in the lowest ability quintile, while first basemen would be under-represented. We have therefore indexed OPS to correct for between-year and between-position variations.

A player's peak ability is measured by the indexed OPS level of his third-best season of more than 130 at bats, including all seasons from 1985 to 2005, plus earlier seasons of players who were active prior to 1985.[11] For our regressions, the sample consists of all player-seasons where we have both salary data and performance over at least 130 at bats for those position players (non-pitchers) with an established peak ability measure. Using the reference levels for each player's peak ability, we established cut lines between players so equal numbers of player-seasons were represented in each of five quintiles.[12] Table 1 reports means of all measures used.

IV. Career productivity paths

[11] Indexed OPS controls for systematic changes in the league-wide OPS level at each position across seasons. The use of the third-best season is somewhat arbitrary, but serves two purposes. First, it avoids the overestimation of ability of players who did extremely well in injury-shortened seasons or a partial-season "cup of coffee" call up from the minors. Second, the cutoff removes very marginal players with fewer than three qualifying seasons from further consideration in the sample. To further limit the sample, players younger than age 29 in 2005 are removed due to the possibility that they might not have peaked yet, so that their inclusion might taint the cohorts.

[12] The cutoff levels of indexed OPS are 0.958, 1.023, 1.073, and 1.153. The indexed values are relative to averages which condition for year and defensive position. The use of player-seasons as the unit of measure, coupled with the longer career lengths of high-ability players, means that there are fewer players in the top (5th) quintile than in lower ones. There are 126 players in 5th quintile and 285 in the 1st quintile.

IV.A. Productivity paths with respect to Age

We will estimate player development and deterioration alternatively by age and by years of experience. Conditioning with respect to age is the intuitive way to analyze the validity of the four implicitly assumed characteristics of career performance paths as presented in Section II.A. and to observe selection effects early and late in careers. However, changes with respect to experience are necessary to make comparisons to salary. Table 2 shows the effect of age on each ability quintile of our MLB sample using the estimated parameters for regression equation (1). In equation (1), each player-season is weighted by plate appearances to avoid bias caused by part-time players, especially in the extreme tails of the age distribution. Panel A of Table 2 reports the results from three model specifications that allow testing of Implication (a). Model 1 naively pools all player seasons without any controls for ability quintile. Model 2 adds differential intercept terms to the Model 1 configuration. The associated Chow F-statistic shows that the improved explanatory power of the model, from 0.2 to 40.0 percent of the variation in ability, is statistically significant. Model 3 tests whether adding differential slope and quadratic terms to Model 2 further improves the fit. The F-statistic of 1.5 confirms that the very small increase in goodness-of-fit is not statistically significant.

The rejection of the null hypothesis in Model 2 confirms that pooling of the ability quintiles is inappropriate, even though the evidence is not strong enough to disprove Implication (a) – that players develop at the same rate. Even so, the fitted equations reported in Panel B of Table 2 for each ability quintile show a clear trend where each successive quintile has a higher slope at younger ages, and a larger magnitude in the quadratic term. The additional curvature means that the best baseball players will exhibit more within-career variation than marginal players, who never rise far above replacement level and lose their jobs when they return to that level.

The Peak Age column in Panel B of Table 2 shows a general trend whereby the relatively poor MLB players in the bottom three quintiles begin to fade sooner. Whereas players in quintiles 1 through 3 peak at ages 25.6 to 26.8, players in quintile 4 peak at 27.5 years of age, and players in quintile 5 peak at age 28.2. The fact that true superstars peak somewhat later than journeymen do is important to know for those attempting to forecast the value of mid-career players on the free-agent market, especially when careers last an average of less than seven years.[13] The result represents a refutation of Implication (b) .

The final result we show in Tables 2 returns to the concept of observational bias. The Chow F-statistics in Model 2 confirmed that pooling of all five ability quintiles was inappropriate. In Section II.A, we predicted that the effect of the pooling was to "flatten" the estimated productivity paths. The bottom row of Panel B in Table 2 shows the estimated pooled regression alongside the five fitted quintile equations. The parameter estimates on the age and age^2 variables in both leagues have lower magnitudes in the pooled regression than for *any* individual quintile in MLB. In addition to this flattening, we see that the effect of the biased slope coefficients is that the estimated peak age is biased upward by 1.5 years above the median peak age of the constituent quintiles, and above the point estimate for any one of them.

Figure 2 illustrates the fitted productivity paths for the five quintiles and the flatter regression line for the pooled sample, as well as the contingent mean values of indexed OPS by age. Each fitted regression line is discontinued at the point when there are fewer than ten players

---

[13] So as not to overstate this finding, it should be noted that the estimated raw OPS levels for top players at ages 26 through 33 are all within 15 points of one another, and given the stochastic error in year-to-year performance these fits could quite easily be thought of as statistically equivalent (the peak resembling a plateau more than a point). Also, the Chow test in Model 3 fails to reject the null that players in different quintiles have statistically equivalent aging patterns aside from having differential quintile intercepts.

remaining in the subsample at that age level. The entry ages for players in all quintiles are similar, but the high level of ability of upper quintile players as they debut lends support to Implication (c), that star players reach replacement level while younger.[14]

It can also be seen in Figure 2 that the players with lower peak ability retire at much younger ages, so that the remaining players force the conditional mean OPS upward for the upper range of ages. That the upward observational bias is mostly occuring in the right tail of the age distribution is consistent with the overestimate of peak age in Table 2. Even though Figure 2 shows that higher ability players retire before they return to replacement level, the production paths extended enough years to verify Implication (d), that star players remain above replacement level longer than marginal players.

IV.B. Productivity Paths with respect to Experience

To anticipate the comparison of productivity to salaries in section V, we fit equation (1) to the data using experience rather than age. Although we saw evidence of sample truncation in the left end of the age distribution, some highly talented players debuted while relatively young, so we would expect somewhat different results from this set of regressions. Expanding upon the notion in footnote 7, the switch from age to experience can be thought of as a relative rightward shift of the higher quintile productivity paths in Figure 1. In addition to delaying the peak productivity points for the highest ability quintiles in terms of experience, the shift would mean that we would expect rates of player development at a given age to differ between quintiles. The best players would be expected to develop faster at low levels of experience and to have ability

[14] The apparent similarity in age of MLB debut across quintiles may be slightly exaggerated in Figure 2. While players in lower quintiles might be brought up as part-time players while young, and play their way into a regular job over two or three seasons, this is less common for top prospects who are projected to become stars. Because free agent eligibility is determined by time on a major league roster rather than playing time, having a young player sit on the major league bench has an opportunity cost of future productivity rents. This opportunity cost will be higher for players who are far above replacement level, so teams are reluctant to promote them until there is a starting position open for them. So while all quintiles are being called up at similar ages, the early observations for the top quintiles tend to represent much more playing time, and have more weight in the Table 2 regressions.

deteriorate more slowly at post-peak levels of experience *even when Implications (a) and (b) are true.* One testable implication of this is that we would expect the Chow F-statistics for differential slopes to be statistically significant, even when they were not when conditioning by age, as was the case in Table 2.

We again weight each player-season observation by the number of plate appearances and index to control for seasonal and positional variation in OPS. The coefficients are reported in Table 3. As predicted, the Chow test in Model 3 of Panel A rejects the null hypothesis of identical slopes at the 99 percent confidence level. Panel B of Table 3 presents the fitted equations, which clearly show that the higher ability quintiles are exhibiting more curvature from within-career variation and have a higher initial slope in the pre-peak years of experience, as we observed in the regressions using age reported in Table 2.

Unlike the models using age in Table 2 where the intercept had no economic interpretation, the intercepts in Panel B of Table 3 estimate the mean indexed OPS for players in that quintile at the time of their debut. The Chow test in Model 2 of Panel A serves to confirm what we saw in Figure 2, that higher quintile baseball players are not debuting as soon as they reach replacement level, but have higher ability (that is, statistically significant differential intercepts) as they enter the league. Even though the sample doesn't observe the highest ability players at replacement level, Implication (c) would be supported by any pre-debut development function that does not impose a discrete "jump" that disproportionately benefits future stars.[15]

[15] Implications (c) and (d) could both be formally tested using forecasting techniques that project the estimated regression equation outside the data range. All of those techniques, in one way or another, factor the extent of data extrapolation into their forecasting errors and confidence intervals, and in this manner would be conducive to formal hypothesis testing even at age and experience levels where the are no observations. We defer from using this approach, however, as the nature of our project is inherently empirical and we wish to avoid the assumption, necessary for extrapolation or forecasting, that the structural form of productivity growth or decline is unchanged outside our data range.

Panel B of Table 3 reports the experience level of peak performance for each ability quintile cohort and for the pooled sample using the estimated equations. While the lower three quintiles peak early, with slow deterioration beginning after only two or three years of experience, players in quintiles 4 and 5 continue developing until peaking after 5.3 and 7.5 years of experience, respectively. While we previously observed an overestimated peak with the pooled regressions using age, the estimated peak from the pooled regression is vastly overestimated, at nearly 9 years of experience, due to earlier retirements of lower ability players. The reason for the increased distortion here, of course, is that because all players are in the population at zero years of experience, the pooled sample regression in Table 3 perfectly exemplifies the extreme where all the observational bias effects are in the right hand tail.

Figure 3 shows that high ability players in baseball can be identified quite early in their career, and these players continue to show high ability until quite late in their careers. While it is true that older players' abilities do fade, as a general rule they retire before they fall into mediocrity, and they are still having very productive years after nearly all the quintile 1 players in their "rookie class" have retired.

V. Career Salary Paths

Table 4 shows annual arithmetic mean and geometric mean salaries in MLB each season from 1985 to 2005. Arithmetic means have increased from about $448,000 to over $3.3 million in that 21 year span, while geometric means have increased from about $267,000 to over $1.57 million.[16] Use of the geometric mean as the measure of central tendency mimics the traditional preference of log-salaries to dollar salaries in labor market regression analysis to counteract the right-skewness of the salary distribution.

[16] The overall upward trend in salaries is not monotonic, interrupted briefly by collusion, expansion, and other short-term influences. There were also particularly sharp rises in 1990 to 1992 and from 1998 to 2001.

To control for inflation and other institutional changes, we index a player's salary relative to the geometric mean salary of players who are free agent eligible in MLB that season. Players with six or fewer years of experience were omitted from the calculation due to their limited ability to contract, which suppresses their salaries below competitive market levels.[17] Table 4 also shows the means for free agent position players, weighted by plate appearances. The arithmetic mean for this group increased over time from $777,000 to $6.08 million, while the geometric mean increased from $682,000 to $4.08 million. As applying natural logarithms to a simple ratio index presents mathematical problems, we have adjusted the formula using a scaling technique to assist computation of a statistic that is log-normally distributed, as are the salaries themselves.[18]

To estimate the best-fitting regression line to these trends, we tried several structural break points and functional forms. The intuitive structural break between 6 and 7 years of experience is used to indicate the approximate date of passage into free agency. Before and after the onset of free agency, separate quadratic regressions estimate the parameters of the quadratic function, with each player-season in the sample weighted by the number of plate appearances. Although there are some "rough seams" between years 6 and 7, the overall fit is good.

The regression results in Panel A of Table 5 show that experience alone explains over half of the indexed log-salary ratio (hereafter, "adjusted salary") for players in their first six seasons, as their salaries increase with ability to contract. The Chow test F-statistic in Model 2

---

[17] In the MLB collective bargaining agreement, experience is measured in days of service on a team's roster, but data on service time is not publicly available. We estimate experience crudely by subtracting the player's debut year from the season year. We set the cutoff at seven years rather than six both due to the frequent practice of short end-of-season callups for young players, which will lead to an early debut but little service time, and as this cutoff maximizes goodness of fit in our later regression models.

[18] To solve the mechanical problem of below-average salaries yielding negative ln(ratio) results, we adjusted to formula to represent percent salaries. Thus, lnnsal = 1 + ln(salary/gmsal)/ln(100), where gmsal is the league-wide geometric mean salary for free agents. The function calibrates the statistic so that a player with the geometric mean salary has an lnnsal of 1.00, a player with double the mean salary will have lnnsal of ln(200)/ln(100) = 1.151, and a player with half the mean salary will have lnnsal of ln(50)/ln(100) = 0.849.

refutes the null that players in all quintiles debut at the same salary level. A look at the intercepts reported in Panel B of Table 5, however, does not show a clear ordinal pattern between the quintiles. The low intercepts for the estimated adjusted salary paths for the quintiles of players as shown in Figure 4 are tightly packed, with rookie players receive only about 8 percent of the geometric mean for free agents.[19]

The Chow test in Model 3 indicates that the quintiles have dissimilar slopes, and the slope coefficients reported in Panel B show that salary increases are directly related to ability and tend to be log-linear across time, as the quadratic terms are of low magnitude and statistically insignificant. Figure 4 shows that the salaries of players with different ability levels are clearly separate by seasons 4 through 6, as negotiating power increases through final-offer arbitration and the desire of some teams to negotiate multi-year contracts for emerging stars prior to the onset of the player's free agent eligibility.[20] The effect of the salary increases can be seen in the rightmost column of Panel B, which show that by the end of the fifth year of experience, player salaries are correctly rank ordered by MRP, although only the very best players are expected to have salaries as high as the free agent geometric mean.

Panel C of Table 5 shows that baseball player salaries are relatively stable once a player reaches free agency. The slope coefficients yield downward trending estimated adjusted salaries as players reach their final seasons, but the standard errors are so large that the coefficients are

[19] The mean adjusted salary (*lnnsal*) for players with 0 years of experience is about 0.45. Inverting the formula, the simple ratio of salary to the geometric mean salary is exp(*lnnsal*·ln(100))/100. For *lnnsal*, this equates to about 0.0794, or about 8 percent of the geometric mean salary.

[20] The mean adjusted salaries conditional upon years of experience confirm that the trend over seasons 1-6 is rising steadily, not in discontinuous jumps as would be true if freedom to contract increased discretely after the third and sixth season. One reason for this comes from possible measurement error in our estimation of arbitration and free agent eligibilty. Another would be that teams will offer some players multi-year contract extensions before the player attains eligibility status, and a portion of the player's expected salary increase at that future date can then be collected at the time of the contract extension.

not statistically significant. The 1.186 adjusted salary ratio for top players at their salary peak represents a salary 2.35 times the geometric mean salary for all free agents in the league.

The R^2 values in Models 4, 5, and 6 in Panel A show that it is the quintile differential intercepts that explain the lion's share of salary variation, so that the between-player variation is more significant than within-career variation in salaries. Even so, Figure 4 shows that expected salaries fall for players in all quintiles as they near retirement, a pattern that is disguised by the observational bias in the right tail of experience. An analyst studying the naive pooled sample model indicated by the thick trend line might conclude that end-of-career salaries decline far more slowly than they actually do.

VI. Pay and Productivity

Due to the multiple adjustments and instances of indexing we have applied to the raw data to assist our analysis of the age trends in performance and in pay, it is unclear what the "efficient" ratio of our constructed estimators would be in an ideally functioning labor market.[21] For this reason, our discussion of pay and performance will be of a relatively heuristic nature. That said, we wish to place our research in the context of the existing literature.

Previous attempts to compare player productivity to compensation have followed the seminal theoretical work of Scully (1974). Using a two-stage model, team revenue is shown to be primarily determined by team wins, establishing that the production of wins is the player's marginal physical product, and that the revenue accruing to the team from a player's performance statistics determine his marginal revenue product. In its final form, the second-stage equation usually takes the form of a semi-log model, as in equation (4).

[21] If one were to assume that the efficient salary for a player of position-specific mean ability is the geometric mean salary of all free agents, and that deviations of ability from the position-specific mean should be rewarded with log-linear increases in salary, then a ratio of 1.00 could be considered efficient. It is not at all clear to us that either of those conditions should necessarily hold.

$$\ln(\text{salary})_i = \alpha + \beta \text{ ability}_i + \delta' X_i + \varepsilon_i \quad (4)$$

In equation (4), X_i represents a vector of exogenous control variables which influence salary, such as the (log) population of the player's host city, whether the player has been named the league MVP or an All-Star, and typically contains a pair of dummy variables to distinguish the levels of negotiating freedom.

By changing the methodological approach to one that is comparing time paths of pay and performance with respect to changes in a mutual covariate, experience, the inclusion of control variables is not necessary. While it is true that the structural causes of changes in the ratio may go undetected with this method, the magnitude of the change in ratio will still be correctly measured. For example, suppose that the two main components of change in the ratio were increased freedom to negotiate contracts and migration of high ability players from small markets early in their careers to large markets mid-career. While we will remain unable to decompose our overall salary ratio increase into estimates for what proportion of the change in relative salaries is due to each effect, the fitted salary ratios should nonetheless correctly measure the sum total of all the partial effects.

For the purpose of constructing a pay-to-productivity ratio for MLB, we use the adjusted salary formula discussed earlier in Section V. The adjusted salaries are approximately normally distributed and exhibit symmetry between players receiving n times the average salary and players receiving 1/n-th the average salary.[22] Dividing this measure by the indexed performance measure discussed in Section IV results in the statistics seen in Figure 5. The key pattern here is that during the years of limited negotiating power, the ratio is quite low, and that beginning with

[22] The only significant departure from normality is the spike that occurs at the truncation in the left tail caused by the league minimum salary.

the advent of free agent eligibility, the ratio becomes very close to 1.[23] Moreover, this is true for players in *all* quintiles, and is such that the ratio stays stable and does not experience a significant drop as the player ages.[24] This indicates that as abilities slowly decline, salaries decline at a proportionate rate. The increased variation in the data points in the right tail is likely due to reduced sample sizes in each conditional mean.

Although our methodology is very different from that in the traditional literature, the ratio analysis both confirms earlier findings regarding the underpayment of inexperienced players, and adds the novel finding that MLB free agents are paid approximately their MRP not just *on average*, but *throughout* the talent distribution.

Another source of possible insight into pay and performance issues is to inspect the peak levels of each reported in Tables 3 and 5. For each quintile, adjusted salary peaked at least 1.8 years after productivity peaked, with an average lag of 2.8 years. We do not claim this is evidence of labor market inefficiency, but merely present this as an empirical finding so that future theories of baseball labor market efficiency might incorporate it.

VII. Summary of findings and future plans

Quintile data permit exploration for differing rates of skill development and deterioration for players of different levels of peak ability, and comparison of the variation in a player's ability across his career to the variation of talent within the league. Moreover, regressions using quintile analysis reduce observational bias caused by careers of star players who enter the league before and/or remain in the league after lesser-able players reach replacement level.

---

[23] As mentioned previously, however, given the multiple manipulations of the two statistics being compared, we cannot claim that a ratio of 1.00 represents market efficiency, but only that relative player salary is increasing in the ratio.

[24] The contingent mean ratios for the lowest quintile of players vary more during years 6 through 11 than the ratios for the other quintiles, due to increased standard error from small sample sizes. But as the ratios are only varying between 0.83 and 0.93, this says more about the stability of the ratios for the other quintiles than about erratic pay ratios for lower ability players.

The data in MLB show some, but not all, of the characteristics that would be consistent of the theory of "nested" ability quintiles proposed in Section II.A. Due to partial truncation of the left tail of the sample, young MLB players within a given ability quintile will debut at a productivity level similar to or above the peak ability level of players in the next lower quintile. In the right tail of the age distribution, star players tend to retire before their skills fall to the level of peaking journeymen, although mid-quintile players attempt to linger longer. In combination, these results strongly suggest, in accordance with the idea of nesting, that stars reach replacement level ability at younger ages than less-talented players, and remain above replacement level ability well after lower-peaking players have been forced into retirement.

It is not true, however, that the productivity paths of the various ability quintiles only differ through a fixed differential intercept. The data show evidence that higher ability players develop faster than lower peaking players. Higher quintile players show substantially more within-career variation. MLB players in the top two quintiles appear to peak about two years later than players in the lowest three quintiles.

Even with star players retiring before they completely regress to replacement level and the partial left tail truncation caused by rent-maximization decisions of MLB owners, the relatively extended careers of players in higher quintiles result in biased slope coefficient estimates from pooled sample regressions on age or experience. As the observational bias is predominately in the right tail for age (and completely so for experience), estimated rates of development and deterioration are systematically underestimated and peak ages and experience levels are overestimated by between 2 and 5 years in pooled productivity regressions on age.

So as not to overstate the importance of these results, the variation in an individual player's ability due to development and deterioration at the MLB level (standard deviation of about 30 points) is of far smaller magnitude than between-player variation (standard deviation of over 112 points). Use of the quintile methodology along with the calculation of adjusted salary to indexed productivity ratios permit a suggestive examination of underpayment of players with reduced negotiating power due to league collective bargaining agreements. As in previous literature, we find that salaries of young players are suppressed below those of similarly talented older peers. A novel result, however, is that we show that salaries of young players of high ability increase more rapidly than salaries of more marginal young players throughout their first years in the league, despite the theoretical ability of teams to exert monopsony power. The negotiating flexibility provided by multi-year contracts allows players to obtain a portion of the rents that would otherwise be extracted from them in their years prior to free agent eligibility, and this results in a smoothing of the career salary path. Interestingly, at all levels of ability, salary levels peak at least 1.8 years after hitting productivity peaks in baseball.

Being able to correctly estimate rates of productivity increase and decline allows competitively negotiated salaries to more closely approximate subsequent productivity, helps managers to make more cost-effective personnel decisions, and assists the calculations of rents in imperfectly competitive labor markets to inform industry-wide policy or regulatory decisions.

In our future research, we will conduct a similar analysis using player productivity and salary data from the NBA, in which imperfect competition in the labor market persists even for veteran players. In an extension of our MLB research, we estimate separate peak ages for distinct skills within baseball and measure the improvements in productivity forecasts that result from replacing the current one-dimensional OPS model. We will also extend our analysis to identify systematic aging trends in productivity of baseball pitchers and test whether the rates of player deterioration are structurally stable across decades.

ACKNOWLEDGEMENTS:

The authors wish to thank Glenn Knowles and session attendees at the March 2007 Midwestern Economics Association meetings for their helpful comments on an earlier draft of this paper.

REFERENCES

Albert, Jim. "Bridging Different Eras in Sports: Comment," *Journal of the American Statistical Association*. 94 (Sept. 1999): 677-680.

Fort, Rodney. "Pay and Performance: Is the Field of Dreams Barren?" in Paul M. Sommers, ed. *Diamonds Are Forever: The Business of Baseball,* 1992 (Washington DC: Brookings Institute Press).

Hakes, Jahn K. and Raymond D. Sauer. 2005. "Effects and Valuation of Fielding in Major League Baseball: A Play-by-Play Analysis" Clemson University mimeo.

Horowitz, Ira, and Christopher Zappe. "Thanks for the Memories: Baseball Veterans' End-of-Career Salaries," *Managerial and Decision Economics,* 19 (1998): 377-382.

Humphries, Brad R. "Equal Pay on the Hardwood: The Earnings Gap Between Male and Female NCAA Division I Basketball Coaches," *Journal of Sports Economics,* 1 (2000): 299-307.

James, Bill. *The Bill James Baseball Abstract*. 1982. (New York: Ballantine Books).

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

Kahn, Lawrence M. "Free Agency, Long-Term Contracts and Compensation in Major League Baseball: Estimates from Panel Data," *The Review of Economics and Statistics,* 75 (1993): 157-164.

Kahn, Lawrence M. "Race, Performance, Pay, and Retention Among National Basketball Association Head Coaches." *Journal of Sports Economics*, 7 (2006): 119-149.

Krautman, Anthony C. "Shirking of Stochastic Productivity in Major League Baseball: Reply" *Southern Economic Journal*, 60 (July 1993): 241-243.

Krohn, Gregory A. "Measuring the Experience-Productivity Relationship: The Case of Major League Baseball." *Journal of Business and Economic Statistics*, 1 (Oct. 1983): 273-279.

Madden, J. F. "Differences in the Success of NFL Coaches by Race, 1990-2002: Evidence of the last hire, first fire." *Journal of Sports Economics*, 5 (2004): 6-19.

Marburger, Daniel R. "Bargaining Power and the Structure of Salaries in Major League Baseball," *Managerial and Decision Economics*, 15 (1994): 433-441.

Marburger, Daniel R. "Arbitrator Compromise in Final Offer Arbitration: Evidence From Major League Baseball, *Economic Inquiry*, 42, (January 2004): 60-68.

Miller, Phillip A. "A Comparison of Free Agent and Arbitration-Eligible Salary Structures in Major League Baseball," S*outhern Economic Journal,* 67 (2000): 87-104.

Schulz, R., D. Musa, J. Staszewski, and R. S. Siegler. "The relationship between age and major league baseball performance: Implications for development," *Psychology and Aging*, 9 (1994): 274-286.

Scully, Gerald. "Pay and Performance in Major League Baseball." *American Economic Review*, 64 (1974): 917-930.

Singell, Larry D., Jr. "Baseball-Specific Human Capital: Why Good but Not Great Players Coach in the Major Leagues," *Southern Economic Journal*, 58 (1991): 77-86.

Vrooman, John. "The baseball players' labor market reconsidered," *Southern Economic Journal,* 63 (1996): 339-360

Zimbalist, Andrew. *Baseball and Billions*, 2002, (New York: Basic Books).

Table 1: Summary statistics for Major League Baseball (MLB) analysis, overall and by talent quintile, 1985-2005.

Variable	q1	q2	q3	q4	q5	Pooled
OPS	0.666	0.708	0.737	0.763	0.846	0.744
indexed OPS	0.890	0.948	0.990	1.027	1.123	0.996
Age	28.2	29.0	29.5	30.1	30.2	29.4
Experience	4.85	5.97	6.62	7.69	8.56	6.74
1B d.v.	0.104	0.097	0.134	0.088	0.170	0.119
2B d.v.	0.113	0.136	0.134	0.132	0.091	0.121
3B d.v.	0.110	0.117	0.127	0.149	0.113	0.123
catcher d.v.	0.148	0.116	0.163	0.172	0.131	0.146
outfield d.v	0.400	0.401	0.313	0.388	0.401	0.381
shortstop d.v.	0.125	0.133	0.130	0.071	0.093	0.110
Year	1995.4	1995.2	1994.8	1993.6	1994.6	1994.7
plate appearances	325.1	400.5	434.2	477.7	534.8	434.5
at bats	292.4	359.2	387.3	423.8	464.8	385.5
runs created	37.1	51.4	60.7	71.1	98.0	63.7
on-base pct.	0.306	0.322	0.332	0.337	0.365	0.332
slugging pct.	0.359	0.386	0.406	0.426	0.481	0.412
salary (millions)	0.57	1.11	1.57	1.93	3.80	1.79
ln(salary)	12.75	13.29	13.57	13.84	14.44	13.58
indexed salary	0.193	0.356	0.502	0.710	1.169	0.586
adj. salary (lnnsal)	0.622	0.742	0.813	0.899	1.007	0.817
ratio of lnnsal/nOPS	0.713	0.793	0.830	0.884	0.903	0.825
N (player-seasons)	1204	1222	1204	1211	1209	6050
N (players)	285	184	161	154	126	910

Note: Quintile sizes are not exactly equal because the career length of marginal players might force some player-seasons across percentile boundaries.

Table 2: Estimated indexed OPS (iOPS), by age and ability quintile, 1985-2005.
Panel A: Regression diagnostics

Coefficient	Model 1	Model 2	Model 3
Intercept	0.810** (0.079)	0.267** (0.062)	0.626** (0.229)
Age	0.015** (0.005)	0.047** (0.004)	0.022 (0.016)
Age^2	-0.0003** (0.0001)	-0.0009** (0.0001)	-0.0004 (0.0003)
Quintile d.v.'s	No	Yes	Yes
Quintile d.v. X age	No	No	Yes
Quintile d.v. X age^2	No	No	Yes
observations	6050	6050	6050
R^2	0.002	0.400	0.401
F-statistic (Chow)		999.9**	1.5

* - Significant at 90% confidence level; ** - Significant at 99% confidence level
Standard errors in parentheses.

Panel B: Fitted equations and peaks, by quintile

Quintile	Intercept	β_1 (Age)	β_2 (Age^2)	Peak Age	Peak iOPS
1	0.626	0.022	-0.0004	25.6	0.906
2	0.437	0.040	-0.0007	26.8	0.969
3	0.512	0.039	-0.0007	26.1	1.017
4	0.456	0.044	-0.0008	27.5	1.056
5	0.311	0.060	-0.0011	28.2	1.163
Pooled	0.810	0.015	-0.0003	28.3	1.024

Table 3: Estimated indexed OPS (iOPS), by experience and ability quintile, 1985-2005.
Panel A: Regression diagnostics

Coefficient	Model 1	Model 2	Model 3
Intercept	0.981** (0.005)	0.887** (0.005)	0.904** (0.010)
Experience	0.011** (0.001)	0.007** (0.0006)	0.001 (0.004)
Experience^2	-0.0006** (0.0001)	-0.0007** (0.0001)	-0.0004 (0.0003)
Quintile d.v.'s	No	Yes	Yes
Quintile d.v. X Exp.	No	No	Yes
Quintile d.v. X Exp.^{2}	No	No	Yes
observations	6050	6050	6050
R^2	0.011	0.396	0.401
F-statistic (Chow)		964.0**	6.4**

* - Significant at 90% confidence level; ** - Significant at 99% confidence level
Standard errors in parentheses.

Panel B: Fitted equations and peaks, by quintile

Quintile	Intercept	β_1 (Exp)	β_2 (Exp^2)	Peak Exp	Peak iOPS
1	0.904	0.001	-0.0004	1.9	0.905
2	0.969	0.001	-0.0004	1.4	0.970
3	1.012	0.003	-0.0006	2.7	1.017
4	1.034	0.008	-0.0008	5.3	1.055
5	1.096	0.018	-0.0012	7.5	1.165
Pooled	0.981	0.011	-0.0006	8.9	1.028

Table 4: Average Major League Baseball (MLB) salaries, 1985-2005

Year	Arithmetic Mean for all players ($millions)	Geometric Mean for all players ($millions)	Arithmetic Mean for free agents* ($millions)	Geometric Mean for free agents* ($millions)
1985	0.494	0.372	0.777	0.682
1986	0.448	0.280	0.837	0.715
1987	0.453	0.267	0.876	0.716
1988	0.474	0.288	0.978	0.804
1989	0.527	0.307	1.050	0.857
1990	0.572	0.341	1.171	1.001
1991	0.896	0.498	1.832	1.523
1992	1.075	0.553	2.179	1.760
1993	1.052	0.469	2.519	1.925
1994	1.143	0.517	2.473	1.805
1995	1.153	0.450	2.823	1.825
1996	1.228	0.511	2.908	2.028
1997	1.451	0.638	3.314	2.404
1998	1.505	0.666	3.627	2.573
1999	1.643	0.734	4.004	2.950
2000	2.198	1.021	4.511	3.278
2001	2.508	1.093	5.590	4.081
2002	2.741	1.213	5.754	4.008
2003	2.967	1.327	6.135	3.940
2004	2.892	1.282	6.004	3.743
2005	3.335	1.572	6.078	4.080

Mean salaries for free agents in the rightmost two columns have been weighted by plate appearances and do not include salaries of pitchers.

Table 5: Estimated adjusted MLB salaries (iSal), by experience and talent quintile, 1985-2005.

Panel A. Regressions

Coefficient	Model 1 years 1-6	Model 2 years 1-6	Model 3 years 1-6		Model 4 years 7+	Model 5 years 7+	Model 6 years 7+
Intercept	0.458** (0.014)	0.322** (0.019)	0.418** (0.024)		0.839** (0.041)	0.682** (0.042)	0.644** (0.190)
Experience	0.072** (0.007)	0.072** (0.006)	0.045** (0.012)		0.043** (0.007)	0.029** (0.006)	0.027 (0.038)
$Experience^2$	0.0038** (0.0010)	0.0032** (0.0009)	0.0028 (0.0018)		-0.0018** (0.0003)	-0.0017** (0.0002)	-0.0009 (0.0019)
Quintile dv	No	Yes	Yes		No	Yes	Yes
dv X Exp.	No	No	Yes		No	No	Yes
dv X Exp^2	No	No	Yes		No	No	Yes
N	3211	3211	3211		2839	2839	2839
R^2	0.511	0.631	0.653		0.067	0.442	0.447
F-stat (Chow)		43.0**	25.5**			78.5**	3.0**

* - Significant at 90% confidence level; ** - Significant at 99% confidence level; ^{N} – Not statistically significant at 90% confidence level Regressions included d.v. to control for defensive position, and for defensive position at each quintile in Models 2 and 3. Standard errors in parentheses.

Panel B: Fitted equations for years 1-6, by quintile

Quintile	Intercept	β_1 (Exp)	β_2 (Exp^2)	ln(iSal) in year 5
1	0.418	0.045**	0.0028^N	0.712
2	0.461	0.053**	0.0048^N	0.844
3	0.432	0.075**	0.0028^N	0.879
4	0.399	0.115**	-0.0002^N	0.968
5	0.463	0.140**	-0.0025^N	1.101
Pooled	0.458	0.072**	0.0038**	0.912

Panel C: Fitted equations for years 7 and beyond, by quintile

Quintile	Intercept	β_1 (Exp)	β_2 (Exp^2)	Peak Exp	Peak ln(iSal)
1	0.644	0.027^N	-0.0009^N	14.3^N	0.834
2	0.902	0.019^N	-0.0013^N	7.7^N	0.976
3	1.035	0.012^N	-0.0013^N	4.7^N	1.064
4	1.060	0.019^N	-0.0013^N	7.1^N	1.126
5	0.961	0.045^N	-0.0023^N	9.9^N	1.186
Pooled	0.839	0.043**	-0.0018**	11.7^N	1.089

Figure 1: Theoretical illustration of observational bias.

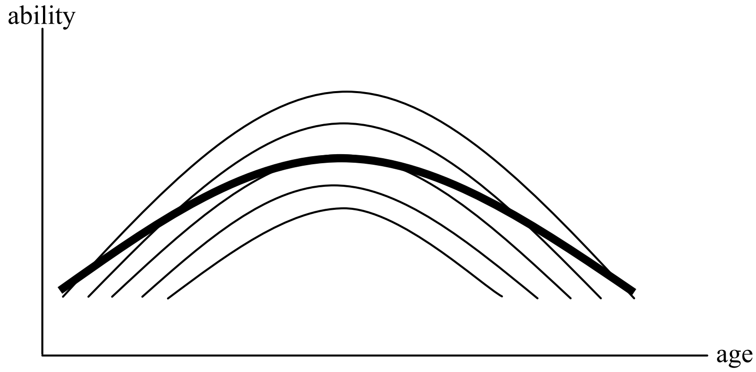
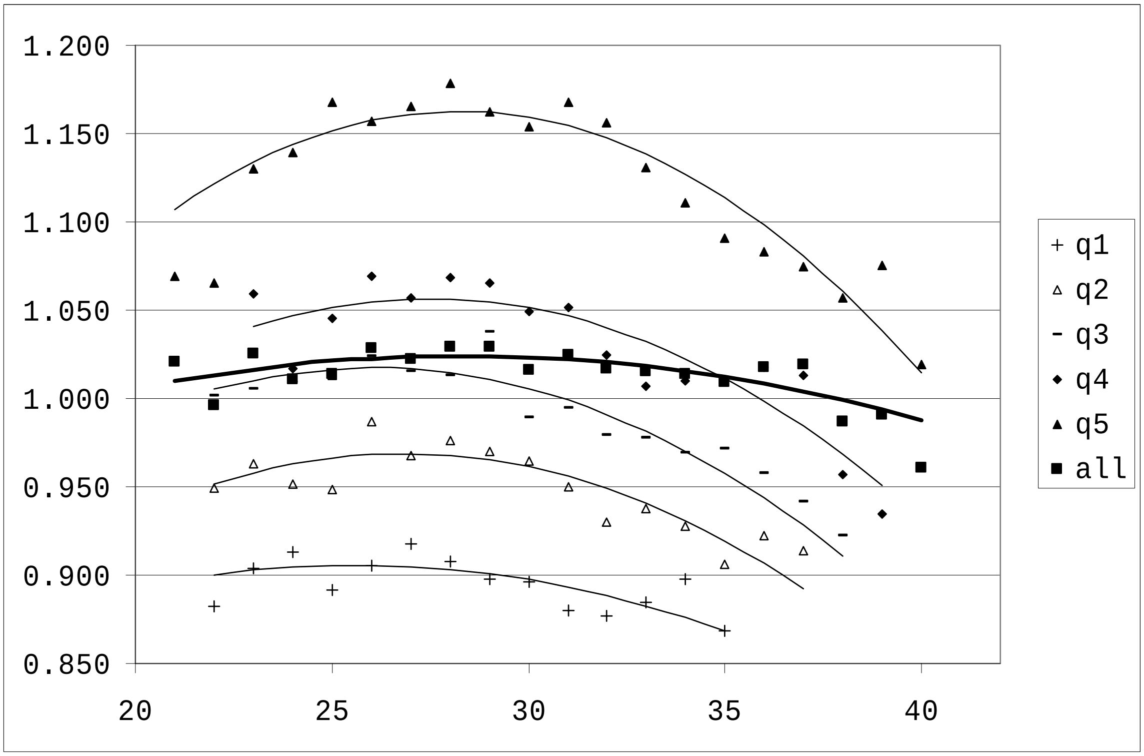
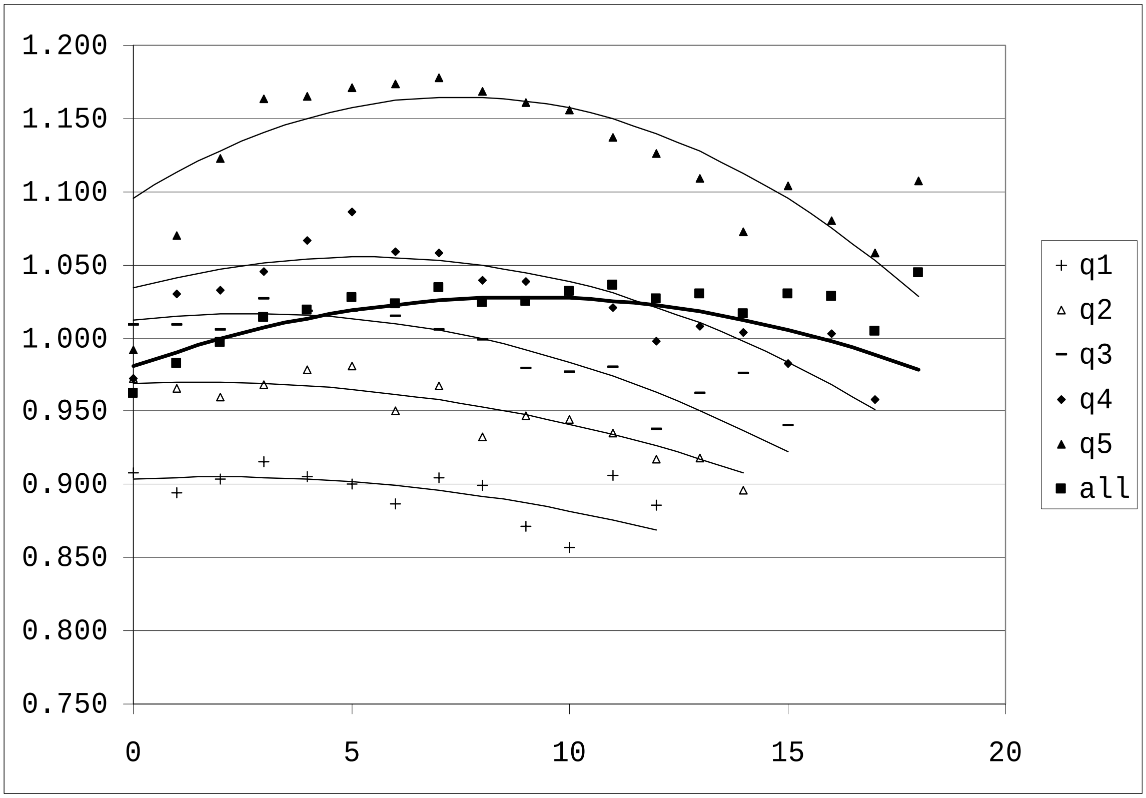


Figure 2: Estimated indexed OPS (iOPS) by age and ability quintile



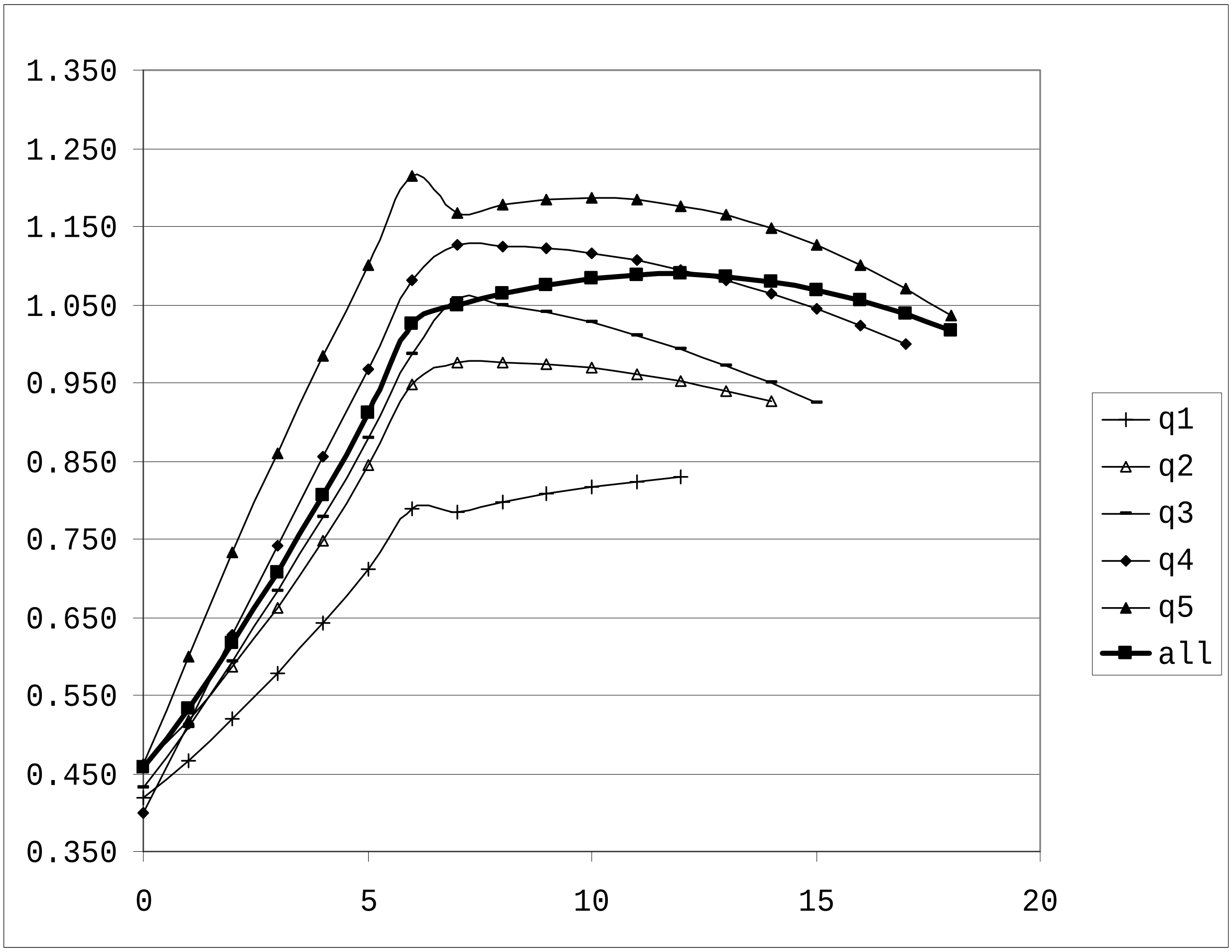
Age (in years) is on horizontal axis; OPS indexed by year and position is on vertical axis. Estimated data points only shown for age-quintile combinations where there are at least ten observations. The corresponding estimated equations are shown in Table 2.

Figure 3: Estimated indexed OPS (iOPS), by experience and talent quintile, 1985-2005.



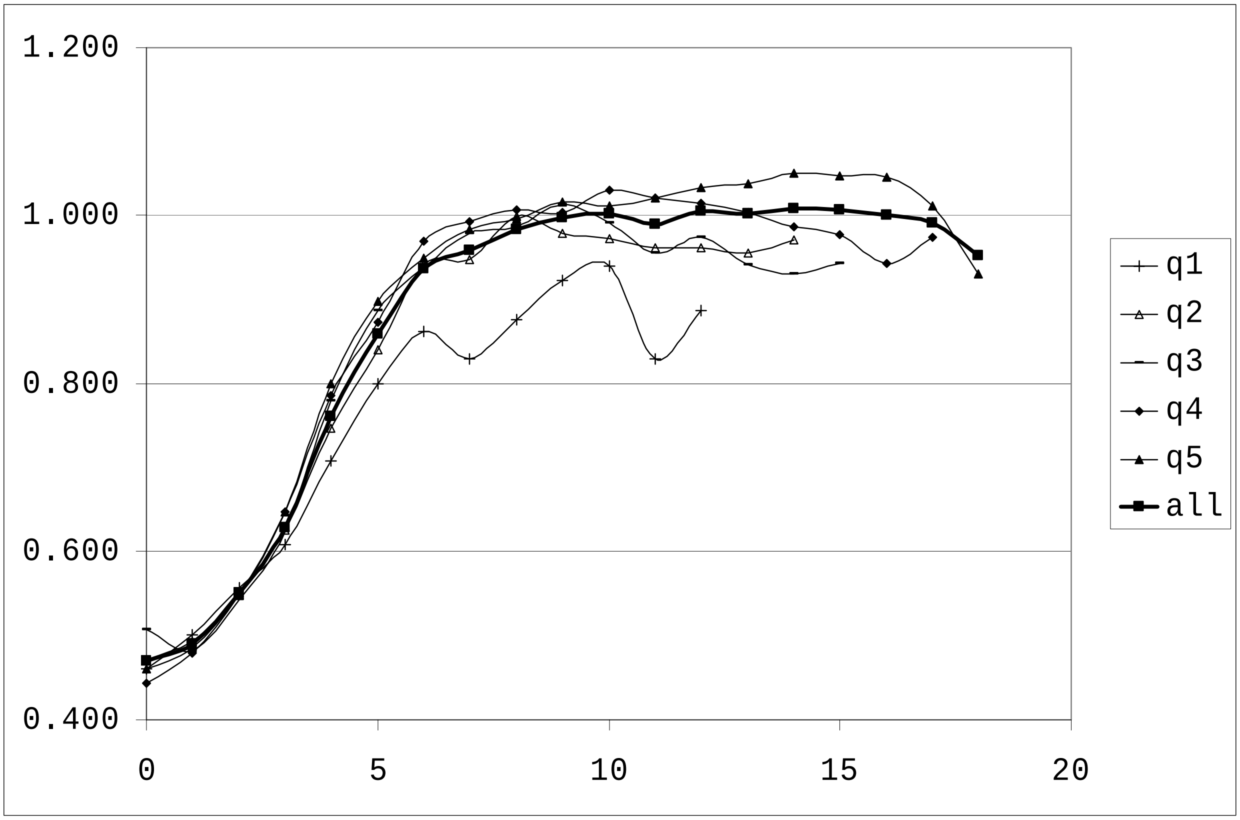
Horizontal axis represents years of experience; the vertical axis represents OPS indexed by year and defensive position. Estimated data points only shown for experience-quintile combinations with at least ten observations. The corresponding estimated equations are shown in Table 3.

Figure 4: Estimated adjusted MLB salaries (iSal), by experience and talent quintile, 1985-2005.



The horizontal axis represents years of experience; the vertical axis represents adjusted salary. Estimated data points only shown for experience-quintile combinations with at least ten observations.

Figure 5: Average ratios of adjusted (MLB) salary over indexed OPS, by experience and ability quintile, 1985-2005.



The horizontal axis measures years of experience; the vertical axis measures the ratio of adjusted salary over indexed OPS (iSal/iOPS), by quintile and experience level. Data points only shown for experience-quintile combinations with at least ten observations.