RISK PREFERENCES AND WAGE DETERMINATION IN THE MAJOR LEAGUE BASEBALL LABOR MARKET

BY

Kevin Whitaker

Submitted to Princeton University
Department of Economics
In partial fulfillment of the requirements for the A.B. degree

April 15, 2013

ACKNOWLEDGEMENTS

I would like to thank …

My adviser of two years, Prof. Andriy Norets, for aiding and overseeing two of the most rewarding projects of my life.

Sean Forman of Baseball-Reference.com, whose wonderful database enabled this research and thousands of baseball studies before it.

Vikram Rao ’11, my sounding board and muse for anything involving sports and/or economics.

And especially my family, whose endless love and support have made this, and everything else, possible.

ABSTRACT

There have been few empirical analyses of firms' risk preferences with respect to labor, largely because the necessary information is not available in many industries. In this study, a model of salary determination in Major League Baseball is developed and used to measure the risk preferences of employers. Using a sample of 1,599 contracts signed from 1993 to 2012, MLB teams are found to be risk-seeking in labor, but only for short-term contracts; in contracts spanning more than one year, teams show no significant risk preference. This pattern is consistent with a new theory of wage determination, in which firms prefer risky workers while risky workers prefer long-term income security.

TABLE OF CONTENTS

INTRODUCTION

The acquisition of new labor can be considered an investment: companies pay a defined salary to new employees, receiving a return that depends on their future productivities. During the hiring process, firms gather information about prospective workers — such as education, test scores and prior productivity —to estimate their future values; however, there will be some risk involved, as an employee's productivity is almost never fully deterministic. Furthermore, certain prospective workers may be riskier than others — such as perhaps younger employees, who have a shorter history of prior productivity from which to make inferences.

In most existing models of salary determination (such as Harris & Holmstrom, 1982; Spence, 1973), firms are assumed to be risk-neutral with respect to labor, so the variance of a worker's expected productivity does not impact his wage. There is little evidence supporting or refuting this assumption, however. In most industries, the information needed to quantify a worker's expected productivity at the time of hiring is not publicly available, so there have been few empirical tests of employers' risk preferences.

One unique sector in which productivity and risk can be estimated is the Major League Baseball labor market. In MLB, as with most professional sports leagues, players' performances and salaries over time are widely accessible, allowing for many employment theories to be tested empirically. Kahn (2000) calls professional sports a "labor market laboratory," because the availability of data has allowed researchers to study economic concepts such as monopsony, behavioral incentives and racial discrimination in empirical settings.

Professional sports are an ideal setting in which to test risk aversion in particular, because players generate extremely high revenues and demand correspondingly high salaries; the average MLB player earned $3.1 million in 2011, roughly two orders of magnitude greater than the average American's annual wage (Berry, 2011). As a result, employers' risk preferences will be much more visible in MLB than in the average labor market, because even a moderate risk affinity or aversion will have a noticeably large effect on total salaries.

For much of professional baseball's history, players were perpetually tied to an employer by the "reserve clause," which gave one team exclusive rights to each worker's labor. Players were never allowed to seek employment from another team; they could only relocate if they were traded by their original club, in which case the new team would assume their sole employment rights. Starting with the 1976 collective bargaining agreement, however, all players with at least six full years of major-league experience became eligible for free agency, earning the ability to negotiate contracts with any team and have their salaries determined competitively.

In today's MLB labor market, players with less than six years of experience are still bound to one team, just as all players were before 1976; this monopsony power allows owners to pay lower salaries than players would receive in a competitive market. For example, players signed for the 1977 season — the first year of free agency — received a higher absolute salary than similar players under the reserve clause in the previous year, and free agents' pay was more strongly correlated with their demonstrated performance (Hill & Spellman, 1983). Several studies of modern baseball's salary structure have also found that players with less than six years of experience are paid less than their free-agent counterparts, when controlling for productivity (Hakes & Turner, 2009; MacDonald & Reynolds, 1994; Marburger, 1994; Miller, 2000). The current study will focus only on free-agent contracts, because they are negotiated competitively and should be determined by players' expected performances.

This study improves on previous research in several ways. First, a theoretical model is developed to explain how firms determine appropriate salaries to offer players in professional baseball. Much work has been done on estimating a baseball player's marginal revenue product, but that is not sufficient to determine his salary, because productivity cannot be exactly known in advance. This study builds upon previous models of marginal revenue product to formalize the process by which salaries are determined, paying specific attention to unique properties of baseball players as employees. The new salary model also accounts for the possibility that employers are not risk-neutral with respect to labor.

Second, a comprehensive set of statistics and contracts is used to empirically analyze the risk preferences of baseball teams. Many previous studies of baseball salaries use limited performance metrics that measure only one or a few relevant skills. Thanks to recent innovations in baseball analysis, a more complete statistic is available that encapsulates virtually all quantifiable skills and more accurately measures players' values. Previous studies of employers' risk preferences have also relied on crude indicators to estimate performance risk; in this study, a generalized autoregressive conditional heteroskedasticity model is used to systematically predict the expected variance in performance for each player at each point in time, yielding better estimates of the actual risk assumed in each contract.

Finally, this study uses a much larger dataset than other analyses of baseball salaries. Most prior studies have used a sample of contracts signed over 1-5 years; in this analysis, all Major League Baseball players whose careers started between 1987 and 2007 are included, generating a set of 1,599 free-agent contracts signed over a 20-year span (1993-2012). Not only does the larger sample yield more powerful and precise conclusions than smaller analyses, but it also allows more specific salary patterns to be tested.

BACKGROUND LITERATURE

General properties of labor contracts and risk preferences

Many early concepts of labor contracts and wage determination are based on marginal productivity theory. According to this theory, given a constant wage interchangeable workers and decreasing returns to labor, a firm will hire additional workers until their marginal revenue product is less than the wage rate (Clark, 1891). Real workers, however, are not interchangeable; they have different skills and capacities, so they have different levels of productivity. To account for this variation, each worker can be considered a separate factor of production, for which firms will be willing to pay a wage no greater than the value of that individual's unique marginal product (Hicks, 1963). In most sectors, a worker's productivity cannot be measured directly, but studies of observable indicators support the theory's predictions; for example, Mincer (1974) shows that education and experience, which are correlated with productivity, explain a significant amount of variation in workers' income.

To fully explain how wages are determined, however, a model must account for the fact that workers' marginal revenue products are not always known at the time of hiring. Spence (1973) develops one such model in which a worker's productivity is not learned until after he is hired but prospective workers display certain signals, such as educational attainment. Based on the performance of past employees who possessed similar signals when they were hired, firms will estimate the current applicant's productivity and offer him a wage based on that prediction. If that worker is hired, his output will be observed, and that information will be incorporated into the formula for estimating future applicants' production. Spence terms this process *informational* *feedback* (p. 359), because a firm's process for predicting workers' productivities is constantly under revision.

While Spence considers a worker's salary only at the time of hiring, Harris and Holmstrom (1982) develop a dynamic model of wages in which an employee's pay is flexible over time. Harris and Holmstrom also use a feedback-based mechanism in which beliefs are updated based on observations, but in their model, the learning process is applied to each employee. As in Spence's theory, firms generate an initial estimate of productivity based on available signals, but this estimate is revised for each worker as his actual production is revealed over time. Harris and Holmstrom extend this model to show that, given symmetric information and risk-averse employees, wages are downward rigid and generally increase with experience, even if a worker's expected productivity is constant. Farber and Gibbons (1996) create another general model of learning in the labor market, in which a firm revises its previous belief of a worker's productivity as his output is observed, and his wage is updated at each point in time to equal his new expected productivity. In this model, Farber and Gibbons assert that information is acquired by "public learning" (p. 1008); that is, a worker's productivity over time is observed by not only his own employer but all other potential employers in the market.

A common feature of the Spence, Harris and Holmstrom, and Farber and Gibbons models is that firms are assumed to be risk-neutral regarding the returns to labor. However, this characteristic is not necessarily true of actual employers. Spence describes the hiring process as "investment under uncertainty" (p. 356), as firms invest a certain wage to acquire workers but receive payoffs derived from their uncertain productivities.

In standard financial theory, economic agents are not usually expected to be risk-neutral when making uncertain investments; instead, they are risk-averse and require a premium to invest in more volatile assets.

In contrast to both the risk-neutral hiring models and risk-averse investment theory, however, Lazear (1998) proposes a model in which, under certain conditions, firms are actually risk-seeking with respect to labor. According to Lazear, employers will pay a premium for risky workers — those whose expected performance distribution has more variance — when there are barriers to a worker's mobility but not to a firm's ability to terminate low-performing workers. Under such conditions, new employees provide *option value* for firms — workers whose performance is revealed to be worse than a desired level can be fired at little cost, but those who perform better than expected can be kept. Because riskier workers have greater variance in performance and are therefore more likely to exceed expectations by a great amount — and because they can simply be terminated if they under-perform — they provide more value to firms than less risky workers with the same mean expected productivity, according to Lazear.

Because it is hard to quantify risk in most sectors of the labor market, there have been few empirical tests of Lazear's theory. Burgess, Lane and Stevens (1998) find that firms in growing industries — which have a longer expected lifespan and are therefore able to benefit more from over-performing risky workers — have more turnover than those in declining industries; this evidence is consistent with a prediction of Lazear's model, but it does not prove that the risk-seeking mechanism is accurate.

Labor contracts and risk preferences in professional sports

Because a wide range of performance and salary data is meticulously recorded and publicly available, many studies of labor contracts have focused on professional sports leagues, especially Major League Baseball. The accessibility of information has allowed researchers to empirically test labor market theories and construct models that account for the unique properties of sports contracts. One of the most frequently studied topics in sports economics is the value athletes generate for their employers. Because of the distinctive nature of the sports industry, players do not have a “marginal product” in the traditional sense of making tangible goods; nevertheless, several models have been formulated to describe how these workers generate revenue.

The seminal model for estimating an athlete’s marginal revenue product is introduced by Scully (1974), who estimates a Major League Baseball player’s economic value with a two-stage process: Each team’s winning percentage in a given season is a function of its players’ productivity, and each team’s revenue is a function of its winning percentage (as well as other factors that are independent of performance). The specific inputs of the Scully model have been refined by Scully (1989), Zimbalist (1992) and many others, but his two-stage framework — in which the relationship between player performance and team revenue is moderated by winning percentage — is still generally accepted (Bradbury, 2013).

One of the most straightforward implementations of the Scully method is detailed in Scully (1989). A team’s winning percentage is expressed as a linear function of its slugging percentage (for hitters) and strikeout-to-walk ratio (for pitchers), and a team’s revenue is expressed as a linear function of its winning percentage and the population of its city:

$$WinPct = 1.09*SLG + .65*K/BB$$

$$Revenue = -1877 + 31696*WinPct + 3.31*Pop$$

A hitter's marginal revenue product, therefore, equals (31696*1.09) times his marginal contribution to the team's slugging percentage, based on coefficients estimated from data in the 1984 season.

Even before Scully's model was introduced, researchers often used the Major League Baseball labor market to study determinants of salary. Pascal and Rapping (1972) test for evidence of racial discrimination with MLB wage data from the 1968 season; they find that, controlling for performance, there was no significant impact of race on baseball salaries at that time. Hill and Spellman (1983) compare negotiated salaries in 1977, the first year of widespread free agency, to salaries of similar players under the reserve clause system the previous year; they find that players received significantly higher wages in the free-agent system, and that free agents' wages were more closely related to their performance than those of players bound by the reserve clause.

When measurements of team revenue are available, the Scully method allows a player's salary to be compared to his estimated marginal revenue product, not just to salaries of other peers. In his study of 1968 data, Scully (1974) finds that the average player's wage was merely 15-20 percent of his MRP — which is not particularly surprising, as at the time, teams still had monopsonistic power over workers due to the reserve clause. After the advent of free agency, Scully (1989) writes that free agents' salaries are generally in line with their MRPs; using a similar model, Zimbalist (1992) estimates that players in the late 1980s who had not yet reached free agency were significantly “exploited” (receiving average salaries roughly 15-50% of MRP), but that free agents actually received salaries slightly higher than their estimated MRPs.

In the last decade, new performance metrics have made it possible to estimate players’ salary functions even more precisely. One particularly influential area of study has been the search for market inefficiencies — characteristics of players that are undervalued in baseball’s labor market. This field was popularized in Lewis’ best-selling book *Moneyball* (2003), which describes how the Oakland Athletics were a successful team in the late 1990s and early 2000s despite having fewer financial resources than most of their competitors. According to Lewis, the Athletics exploited several inefficiencies — such as the fact that walks were undervalued in the baseball labor market, while runs batted in and other context-dependent statistics were overvalued — to acquire productive workers at low costs.

Several recent economic studies have found or rejected other potential inefficiencies in the MLB free-agent labor market. Hakes and Sauer (2006) confirm Lewis’ assertion that walks were undervalued around the turn of the century, though they note that the inefficiency disappeared after *Moneyball* was published and the Athletics’ strategy was made public. Healy (2007) finds evidence that baseball teams place too much weight on the most recent season of a player’s performance, and too little on seasons two or more years in the past, when estimating future productivity. Bradbury (2007) concludes that pitchers are efficiently valued in free agency, finding that their salaries are based on characteristics completely within their control (such as the ability to strike out opposing batters or avoid walks) and not determined by other statistics that also measure the performance of teammates (such as earned run average, which depends on fielders' ability as well as pitchers'). Gassko (2011) finds that teams properly value most characteristics of players but, using an advanced statistic to measure defense, concludes that fielding ability is undervalued in baseball's labor market.

Most studies of contract properties in sports have focused mainly on the mean of a player's expected performance distribution and how it relates to his salary. A few have also analyzed the variance of that distribution and the impact of performance risk on wages, however. In studying contracts in the National Basketball Association, Bodvarsson and Brastow (1998) find that basketball players who score a consistent number of points per game receive higher wages than comparable players whose performances vary more by game. Bodvarsson and Brastow explain the preference for consistent players as a desire to avoid monitoring costs, assuming that a volatile worker must be monitored more frequently than one whose output is more consistent, but their results are consistent with other theories of risk aversion.

Several other studies have used contracts of young athletes to test Lazear's theory that firms are risk-seeking under certain conditions. In most American sports leagues, players enter the workforce via a draft, in which firms take turns selecting new workers. These young players can easily be released by the team that drafted them, but they cannot seek a new employer on their own until they have reached free agency (which occurs only after several years), so they fulfill Lazear's conditions for workers with option value. Hendricks, DeBrock and Koenker (2003) find some evidence suggesting that National Football League teams prefer drafting players from small colleges — who have had fewer opportunities to play against the most talented teams, and therefore may carry more uncertainty when moving to a higher level of competition — in the later rounds of the NFL Draft, which is consistent with Lazear's theory. Groothuis, Hill and Perri (2007) find slight support of Lazear's theory in the NBA Draft, claiming that increased restrictions on player mobility may cause teams to prefer younger (and presumably riskier) players.

Similarly, Bollinger and Hotchkiss (2003) find that Major League Baseball teams pay higher salaries for players with a more volatile performance history when those players' mobility is restricted by the reserve clause. When players are not bound by the reserve clause, Bollinger and Hotchkiss find an estimated effect of volatility on salary that is positive but not statistically significant. Burger and Walters (2008), however, find some evidence of risk-seeking behavior even among free agent contracts. Using a more advanced measure of performance, they show that among free agents from 1975-80, players whose production was more volatile over the previous three years received significantly higher salaries than less volatile players did; they find that teams pay a similar (though less significant) variance premium for free agents in the 1998-99 seasons. Burger and Walters describe this behavior as an example of bounded rationality in a complex decision-making setting, assuming that teams should have discounted their bids for risky free agents when they actually did not.

The current study builds upon previous literature in several ways. Scully's marginal product theory for baseball players and several general wage theories are drawn upon to construct a model of how salaries are determined in the Major League Baseball labor market. This model is adapted to incorporate employers' attitudes toward risk, and an empirical analysis improves upon existing studies of risk preferences in MLB by using more accurate estimation processes. A large dataset allows for the study of more granular patterns of salary determination, which clarify how existing risk theories are applied in an empirical setting.

THEORETICAL MODEL

How players produce value

For teams to decide how much they are willing to offer for a player's services, they must be able to estimate his expected marginal revenue product. The framework for determining a player's value was first developed by Scully (1974), who propses a two-stage model: Each player's performance affects his team's winning percentage, and a team's winning percentage affects its revenue. One implementation of this method, Scully's (1989) model, is shown in the previous section; a more general form is presented in equations 1.1 and 1.2:

$$Wins_{it} \;=\; \alpha_0 + \boldsymbol{\alpha} * \mathbf{P}_{it} \tag{1.1}$$

$$Revenue_{it} \;=\; \beta_0 + \beta_1 Wins_{it} + \beta_2 \gamma_i + \beta_3 \lambda_t \tag{1.2}$$

where $Wins_{it}$ is the number of games team *i* wins in year *t*,[1] $\mathbf{P}_{it}$ is a vector representing the team's performance in various categories (for example, Scully uses slugging percentage for hitters and strikeout-to-walk ratio for pitchers), and the regression vector $\boldsymbol{\alpha}$ represents the effect of each category in **P** on team wins. γ_i represents any team-specific determinants of revenue (such as the population of a team's city), and λ_t accounts for a time trend in MLB team revenue. None of the factors in γ_i (or λ_t) can be changed by players, so the only way players affect their team's revenue is by helping it win games.

Implicit in Scully's method is the belief that a team's performance is the sum of individual players' performances, which requires each player's value to be independent of his teammates' skills. In many sports, this assumption would be questionable — for example, a quarterback and a wide receiver with exceptional synergy might provide more

[1] Most versions of the Scully model use winning percentage instead of total wins, but since all teams play the same number of games (162), the two formulations are functionally equivalent. Total wins is used here because it is more directly estimated by the performance variable used in this study.

value to a football team than each would separately — but in baseball, there are rarely opportunities for players to affect a teammate's performance. One frequently theorized example of teammate interaction is "lineup protection," which dictates that batters will hit better when an imposing teammate is on deck; however, several studies have found that lineup protection has little or no effect (Bradbury & Drinen, 2007; Click, 2006; Tango, Lichtman, & Dolphin, 2007).

It will be assumed here that a team's aggregate performance can be expressed as the sum of each player's independent contributions. Thus, a player's marginal revenue product can be calculated as the effect of his individual performance on team wins, times the team's marginal revenue per win, as shown in equation 1.3:

$$\mathbf{P}_{it} = \sum_{j=1}^{J} \mathbf{P}_{jt} \qquad \text{(for all players } j = 1, 2, \ldots, J \text{ on team } i\text{)}$$

$$Wins_{it} = \alpha_0 + \boldsymbol{\alpha} * \sum_{j=1}^{J} \mathbf{P}_{jt}$$

$$Revenue_{it} = \beta_0 + \beta_1\left(\alpha_0 + \boldsymbol{\alpha} * \sum_{j=1}^{J} \mathbf{P}_{jt}\right) + \beta_2\gamma_i + \beta_3\lambda_t$$

$$= \beta_0 + \beta_1\alpha_0 + \left(\sum_{j=1}^{J} \beta_1\boldsymbol{\alpha} * \mathbf{P}_{jt}\right) + \beta_2\gamma_i + \beta_3\lambda_t$$

$$MRP_{jt} = \beta_1(\boldsymbol{\alpha} * \mathbf{P}_{jt}) \tag{1.3}$$

Forming beliefs of player productivity

If players' performances were known in advance, their salary determination would be straightforward — teams would be willing to pay each player a wage no greater than his marginal revenue product at each point. However, as in many labor markets, baseball teams do not have perfect knowledge of a worker's future productivity at the time of negotiation. Two important characteristics of baseball performance prevent players' MRPs from being perfectly predictable.

1. Performance is stochastic. Before 2001, San Francisco Giants shortstop Rich Aurilia had established himself as a roughly average hitter — in every season from 1997-2000, he had a batting average between .266 and .281. But in 2001, he played superbly, batting .324 for the full season and receiving the Silver Slugger award as the league's best-hitting shortstop. After that year, Aurilia immediately returned to his previous level, hitting .257 and .277 in 2002 and 2003; he retired in 2009 with a .275 career batting average and never topped .300 in another season.

It is unlikely that Aurilia's inherent talent went from average in 2000 to one of the league's best in 2001, then right back to average in 2002; he played on the same team with similar surroundings in all three seasons, so if he had truly increased his ability in 2001, he should have retained some of that skill in the following year. Instead, Aurilia's batting average in 2001 demonstrates the stochastic nature of baseball performance.

Professional baseball teams play a long season, but each player bats only a few times in each game. Over the course of a 162-game schedule, full-time players take roughly 600 at-bats, which is a large but not enormous sample of their true abilities; thus, even over a full season, performance can vary significantly due to randomness. For a player whose chance of getting a hit in every at-bat is exactly .27, the standard error over 600 at-bats is .018.[2] Even if his true talent level is known and constant, a set of possible batting averages within two standard errors of his mean ranges from .234 (well below average) to .306 (All-Star-caliber).

2. Beliefs of talent are uncertain. Of course, the exact probability of getting a hit cannot be known for any player; the same goes for all other skills. To predict a player's

---

[2] From the binomial distribution standard error formula, $SE = \sqrt{p(1-p)/n}$

future value, teams must estimate his underlying productivity, but this estimate comes with uncertainty. For a player without major-league experience, a team's prior belief of his ability comes from other factors — such as minor-league performance or scouting data — and that prior is updated as his major-league performance is revealed.

In contract negotiations, this model will assume that knowledge of a player's underlying productivity is learned symmetrically between all potential employers and the player. If knowledge was asymmetric among employers, a player's original team would be better informed about his future performance than other teams. For players currently under study — free agents who have played in MLB for at least six years — a fairly long track record of major-league performance is publicly available. Additionally, teams employ scouts to gather information about all amateur and professional players, so each team will have collected several years' worth of qualitative scouting information about any given free agent. Finally, before finalizing a contract, teams require players to pass a medical examination to discover any physical or injury-related information not reflected in their performances, which the original team could have otherwise kept private to gain a competitive advantage.

The assumption that knowledge is symmetric among teams has been contested. Lehn (1984) finds some evidence of asymmetric information between clubs in the free-agent market, though only in one specific domain — likelihood of future injury — and only among one specific positional group, pitchers; Lehn's study was also conducted at a time when performance measurement and scouting methods were less refined than they are in the present day, and when free agency was in its infancy, so teams may have still been learning how to behave rationally. A more recent study found slight but insignificant evidence of asymmetric information in baseball (Burger & Walters, 2008).

If knowledge was asymmetric between worker and firm, the player would be better informed about his future performance than teams would be. As in the first case, performance data and medical information are generally common among player and team. One remaining area in which the player could have an informational advantage is how much effort he will exert during games or training. The significance of effort in baseball has been empirically tested several times with varying results. Such studies have generally compared a player's performance in the year before free agency — which is assumed to have the most impact on his future earnings — to his performance in the first year of a long-term contract, when he will not have to negotiate for a long time and therefore has less financial motivation to give his best effort. Some analyses find that players perform worse after signing long-term contracts, which could indicate reduced effort (Perry, 2006; Martin, Eggleston, Seymour, & Lecrom, 2011), but other studies find no difference, especially when accounting for the fact that players may be more likely to sign long-term contracts after uncharacteristically good seasons (Krautmann, 1990; Maxcy, 2004). Because the evidence regarding asymmetric information in MLB (between player and team or between different teams) is conflicting and generally suggests small effects at most, this study's model will suppose that information is common among all parties; however, it should be noted that this assumption may not be completely accurate.

Consistent with previous literature on wage determination with uncertain productivity (Harris & Holmstrom, 1982; Krautmann, 1990), it is assumed here that learning occurs as a Bayesian process, where beliefs of a player's innate productivity are updated as each season's performance is revealed. To specify Krautmann's equation to baseball negotiations, teams hold a prior set of beliefs about each player's true talent in *n* categories of performance, represented by the *n*-by-1 dimensional vector $\boldsymbol{\mu}$. These beliefs are updated by the following process:

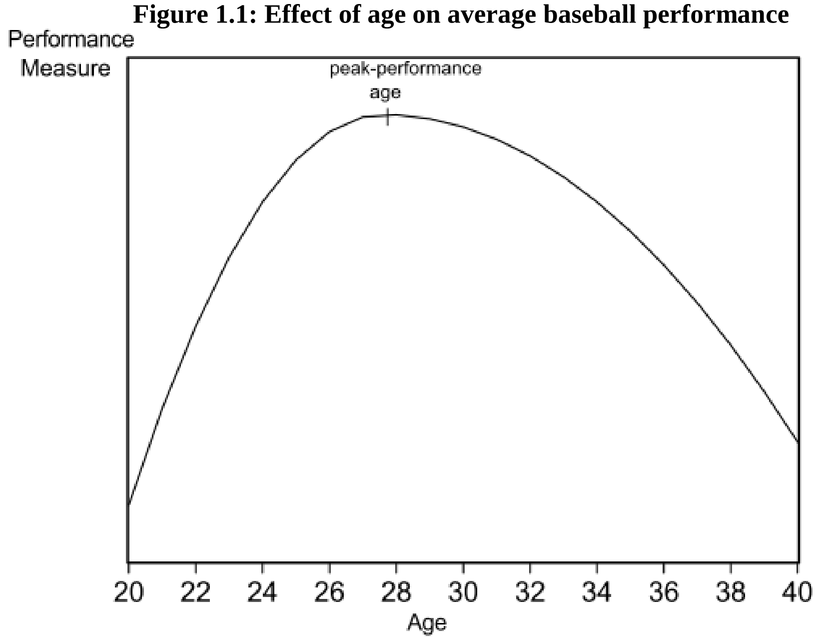
$$\boldsymbol{\mu}_{jt} \ = \ \theta \mathbf{P}_{j,t-1} + (1-\theta)\boldsymbol{\mu}_{j,t-1} \tag{1.4}$$

where $\mathbf{P}_{j,t-1}$ is player *j*'s actual performance in each category in season *t-1*, $\boldsymbol{\mu}_{jt}$ is the belief of player *j*'s underlying talent level in each category for season *t*, and θ describes the willingness to update prior beliefs based on new information.

If a player's true talent was constant over time, uncertainty would eventually diminish and beliefs of his true talent would converge to his actual productivity as more performance was revealed. But a player's true productivity is not constant; on average, it varies predictably with age, as shown in Figure 1.1 (from Fair, 2008; p. 5). The exact peak age and slopes of the curve have been contested (Bradbury, 2009; Tango, 2007), but it has been well established that, for an average player, the performance function is concave in age and peaks in his late 20s.

The year-to-year change in a player's talent is a combination of two factors: physical human capital and mental human capital. Physical human capital refers to athletic ability, including traits such as running speed and muscle strength. Most of these traits peak at a relatively young age; for instance, sprinting ability and jumping ability, two explosive physical efforts that are similar to the tasks required of a baseball player, peak on average in one's early 20s (Schulz & Curnow, 1988). Mental human capital, which encompasses traits such as the ability to anticipate the location of a pitched or

Figure 1.1: Effect of age on average baseball performance



batted ball and the knowledge of certain playing and training strategies, is likely to increase with age over the course of a player's career, though possibly at a diminishing rate. The addition of physical and mental human capital forms a productivity curve that peaks in a player's late 20s, when the rate of decline in physical skills begins to exceed the rate of increase in mental skills.

Because a player's inherent talent changes with age, and because this change happens at a relatively predictable rate, teams can factor the estimated effects of aging into their beliefs of a player's uncertain talent from equation 1.4:

$$\boldsymbol{\mu}_{jt} \;=\; \theta \boldsymbol{P}_{j,t-1} + (1-\theta)\boldsymbol{\mu}_{j,t-1} + \boldsymbol{F}(AGE_{jt}) \tag{1.5}$$

where $\boldsymbol{F}(AGE_{jt})$ is a function that describes the expected change in each category of player *j*'s talent at his age in year *t*, according to studies of past players.

How salaries are determined

Using its belief of player *j*'s inherent productivity in year *t*, $\boldsymbol{\mu}_{jt}$, a team can determine the maximum salary it would be willing to pay to employ that player.[3] If teams are risk-neutral, they should be willing to pay a player up to his expected marginal revenue product, which is a function of his estimated true talent:

$$Salary_{jt} \leq \mathrm{E}\left[MRP_{jt}\right] = \beta_1(\boldsymbol{\alpha} * \boldsymbol{\mu}_{jt}) \quad (1.6)$$

where, as before, β_1 is the marginal revenue generated per win and $\boldsymbol{\alpha}$ is a coefficient vector measuring the effect on wins of each performance category represented by $\boldsymbol{\mu}$.

As an illustration, consider a simplified case in which players help their teams win only with singles, doubles, triples and home runs. Based on his prior performance and other characteristics such as age, a player and employers share a set of beliefs $\boldsymbol{\mu}_{jt} = (\mu_{\mathrm{1B,jt}};\ \mu_{\mathrm{2B,jt}};\ \mu_{\mathrm{3B,jt}};\ \mu_{\mathrm{HR,jt}})$, with $\mu_{\mathrm{1B,jt}}$ indicating the number of singles player *j* is expected to hit in year *t*, and so forth. Team *i*'s expected wins are a linear function of those four outcomes:

$$WINS_{it} = \alpha_0 + \alpha_1 \mathrm{1B}_{it} + \alpha_2 \mathrm{2B}_{it} + \alpha_3 \mathrm{3B}_{it} + \alpha_4 \mathrm{HR}_{it}$$

where $\mathrm{1B}_{it}$ is team *i*'s singles in year *t*, α_1 is the marginal contribution of each single towards team wins, and so forth. Therefore, a player's expected marginal revenue product is:

$$\mathrm{E}\left[MRP_{jt}\right] = \beta_1(\alpha_1\mu_{\mathrm{1B},jt} + \alpha_2\mu_{\mathrm{2B},jt} + \alpha_3\mu_{\mathrm{3B},jt} + \alpha_4\mu_{\mathrm{HR},jt})$$

[3] A more efficient solution would be to design a contract in which a player's pay varies according to his performance, so his salary is equivalent to his MRP in all possible states. This is not permissible in the MLB labor market, however, as the league's rules forbid teams from paying bonuses for "playing, pitching or batting skill." Teams are allowed to design contracts with variable compensation only if it is based on statistics that measure quantity of playing time, such as the number of games played, and not the quality of performance.

A more complete model would incorporate other productive activities such as walks, stolen bases and fielding.

A player's expected MRP does not represent a guaranteed return, however — only the mean value of a set of possibilities. Therefore, if teams are not risk-neutral, they may not be willing to pay a player exactly his expected MRP, adjusting their maximum offers based on the risk involved.

Consider a one-dimensional model of performance in which player *j* is believed to have talent μ_{jt} at time *t*, leading to an expected MRP of $\beta_1\alpha_1\mu_{jt}$ (where α_1 is the effect on team wins of the skill measured by μ). The player then reveals an actual performance of P_{jt} at time *t*, leaving a residual of $\varepsilon_{jt} = P_{jt} - \mu_{jt}$. That residual can be separated into two sources of error: the first is error in measuring his underlying talent, and the second is stochastic error. Let X_{jt} be player *j*'s actual talent (the true mean of his performance distribution) at time *t*. The residual $\varepsilon_{jt} = P_{jt} - \mu_{jt}$ can then be expressed as:

$$\varepsilon_{jt} = \left(X_{it} - \mu_{jt}\right) + \left(P_{jt} - X_{it}\right)$$

where $\left(X_{it} - \mu_{jt}\right)$ represents error due to uncertainty in measuring the player's talent and $\left(P_{jt} - X_{it}\right)$ represents stochastic error. Let $\varepsilon_{jt,u}$ represent the error due to uncertainty and $\varepsilon_{jt,s}$ represent stochastic error. By definition, $\mathrm{E}\left[\varepsilon_{jt,s}\right] = 0$, and if the team's prior beliefs are not systematically biased towards optimism or pessimism, $\mathrm{E}\left[\varepsilon_{jt,u}\right] = 0$. If the standard deviation of $\varepsilon_{jt,s}$ is $\sigma_{jt,s}$ and the variance of $\varepsilon_{jt,u}$ is $\sigma_{jt,u}$, the variance of the entire residual, ε_{jt}, is $({\sigma^2}_{jt,u} + {\sigma^2}_{jt,s})$.

For a team that is non-risk-neutral in hiring labor, its expected utility from a player's services will be a function of both his expected MRP and the variance of that estimation:

$$EU_{jt} \;=\; \mathrm{E}\left[MRP_{jt}\right] \;+\; \delta * \mathrm{Var}(MRP_{jt})$$

where δ indicates the team's affinity for risk. (For a risk-averse firm, $\delta < 0$; for a risk-seeking firm, $\delta > 0$.) Since MRP is a function of the player's performance, P_{jt}, the expected utility of signing player *j* for year *t* can be calculated:

$$\begin{aligned} EU_{jt} \;&=\; \beta_1\alpha_1 * \mathrm{E}\left[P_{jt}\right] + \delta * \mathrm{Var}\left(\beta_1\alpha_1 P_{jt}\right) \\ &=\; \beta_1\alpha_1 * \mathrm{E}\left[\mu_{jt} + \varepsilon_{jt}\right] + \delta{\beta_1}^2{\alpha_1}^2 * \mathrm{Var}(\mu_{jt} + \varepsilon_{jt}) \\ &=\; \beta_1\alpha_1\mu_{jt} + \delta{\beta_1}^2{\alpha_1}^2({\sigma^2}_{jt,u} + {\sigma^2}_{jt,s}) \qquad (1.7) \end{aligned}$$

The expected utility function from equation 1.7 can be substituted into equation 1.6 to determine the maximum salary teams will be willing to pay each player:

$$Salary_{jt} \;\leq\; \beta_1\alpha_1\mu_{jt} + \delta{\beta_1}^2{\alpha_1}^2({\sigma^2}_{jt,u} + {\sigma^2}_{jt,s})$$

Since free agents can negotiate with all potential employers, competitive pressures should increase their wages until they match the maximum salary teams are willing to pay. Therefore, a free agent's predicted salary is:

$$Salary_{jt} \;=\; \beta_1\alpha_1\mu_{jt} + \delta{\beta_1}^2{\alpha_1}^2({\sigma^2}_{jt,u} + {\sigma^2}_{jt,s}) \qquad (1.8)$$

DATA AND METHODOLOGY

Using the model developed in the previous section — in which a free agent's negotiated salary equals his expected MRP, with an adjustment for the variance of his expected production — the risk preferences of professional baseball teams can be studied empirically. Estimating a player's exact marginal revenue product is difficult, however, because total team revenues are not publicly known. Neither individual teams nor MLB routinely releases complete balance sheets, and the league has disputed some available estimates of team financial data (Associated Press, 2006). Additionally, owners of baseball teams often control other assets, such as stadiums, parking lots and cable television networks, whose revenues do not nominally accrue to the team but are partly determined by its performance. When such assets exist, increases in team revenue may have "spillover effects" into the revenues generated from associated properties; therefore, the team's reported revenue may significantly understate the actual team-related revenue that accrues to the owner (Fort, 2006).

Since the available estimates of team revenue may be biased, empirical measurements of a player's MRP could be biased as well. Therefore, this study will not address questions involving *absolute* MRPs, such as whether players are paid more or less than the true value of their performance. However, in the Scully model, poor information regarding team revenue only affects the relationship between team wins and team revenue — it does not affect the relationship between player performance and team wins. So it is still possible to measure each player's marginal product — his contribution to team wins — even if it cannot be precisely expressed in terms of revenue. Since each marginal win is assumed to have the same value, β_1, and since free-agent salaries are assumed to be determined in a competitive labor market, players' *relative* MRPs can be analyzed by comparing their marginal products. Therefore, even without accurate measures of team revenue, relative trends regarding how players are priced in baseball's labor market — such as whether risky workers receive higher or lower wages than consistent workers of similar average skill — can still be effectively studied.

As modeled in the previous section and shown in equation 1.8, a player's expected salary is:

$$Salary_{jt} \ = \ \beta_1\alpha_1\mu_{jt} + \delta{\beta_1}^2{\alpha_1}^2({\sigma^2}_{jt,u} + {\sigma^2}_{jt,s})$$

where μ_{jt} is a common belief of player skill that is updated over time; α_1 is a coefficient that converts the skill described by μ into marginal wins; ${\sigma^2}_{jt,u}$ and ${\sigma^2}_{jt,s}$ are the expected variance in μ_{jt} due to uncertainty and stochastic error, respectively; β_1 is the marginal revenue a team receives from each additional win; and δ denotes a team's affinity for risk. Since $\alpha_1\mu_{jt}$ represents player *j*'s expected marginal product (the wins he is expected to generate at time *t*), the salary equation can also be expressed as:

$$Salary_{jt} \ = \ \beta_1\mathrm{E}\left[MP_{jt}\right] + \delta{\beta_1}^2\mathrm{Var}(MP_{jt})$$

Therefore, teams' risk preferences, δ, can be estimated with a linear regression involving three necessary variables: Player's salaries, their expected marginal products, and the predicted variance of their marginal products.

Data and measurement

Using Baseball-Reference.com's comprehensive database of baseball statistics, a list of all players who have appeared in at least one major-league game between 1987 and 2012 has been assembled, along with each player's annual statistics in that span. Players whose careers began before 1987 were removed so that a full history of prior performance is available for each player at each point in time; since only free-agent contracts will be analyzed, the sample was further narrowed to include the careers of only players who spent at least six seasons in the major leagues (the minimum service time required for free agency). For each player, a wide range of performance data is tabulated for each season, as well as his career totals in several categories. The length, average annual salary and date of every contract signed by those players in free agency has also been recorded using Baseball-Reference.com's transactions database and salary information, which were cross-checked with media reports when necessary.

The sample of players was split into two groups, hitters and starting pitchers[4], which will be analyzed separately because they produce value in different ways. The first group includes 693 hitters, who signed a total of 1,219 free-agent contracts between 1993 and 2012; those contracts range from 1 to 10 years in length and from $109,000 to $27,500,000 in average annual salary. The second group includes 319 starting pitchers, who signed a total of 380 free-agent contracts between 1993 and 2012; those contracts range from 1 to 8 years in length and from $115,000 to $24,000,000 in average annual salary[5].

The primary performance variable used in this study is Wins Above Replacement (WAR), a statistic generated by Baseball-Reference.com that combines various

[4] Relief pitchers have been omitted from this study because there is evidence that their salaries are not determined by objective performance statistics in the same way that hitters' and starting pitchers' salaries are (Brownson, 2010; Swartz, 2012). Whether employers pay for unmeasured productive skills or are simply behaving irrationally has been contested (Moore, 2011; Swartz, 2012), but in either case, including relievers in a study that involves relating productivity and salary could be problematic.

[5] For an additional 168 hitters' contracts and 101 pitchers' contracts, salary information was not easily retrievable. In inspecting many of these cases, it was found that the vast majority were conditional contracts that paid the player a trivial wage to play in the minor leagues, with a promised salary increase if he was promoted to the major-league roster. Since these contracts are structured differently than regular free-agent contracts, with most of the salary not guaranteed, they were excluded from this analysis.

categories of player statistics and weights them based on their impact on team wins. In the previous section, a team's wins in year *t* were modeled as a function of $\mathbf{P}_t$, a vector of many different performance categories in year *t*:

$$Wins_{it} \;=\; \alpha_0 + \boldsymbol{\alpha} * \sum_{j=1}^{J} \mathbf{P}_{jt} \qquad \text{(for all players } j = 1, 2, \ldots, J \text{ on team } i\text{)}$$

WAR applies a weighting vector, $\boldsymbol{\alpha}_{\boldsymbol{w}}$, to each player's performance in order to determine the number of wins he contributes (his marginal product):

$$WAR_{jt} \;=\; \boldsymbol{\alpha}_{\boldsymbol{w}} * [\mathbf{P}_{jt}]$$

If WAR is appropriately calibrated (i.e., $\boldsymbol{\alpha}_{\boldsymbol{w}} = \boldsymbol{\alpha}$), then:

$$Wins_{it} \;=\; \alpha_0 + \sum_{j=1}^{J} [WAR_{jt}]$$

The sum of individual players' WAR, as calculated by Baseball-Reference[6], has been shown to be a very strong predictor of team wins ($R^2 = .83$), suggesting that WAR's weighting vector is appropriate and that individual WAR is a good estimate of a player's contributions to winning (DuPaul, 2012).

Previous studies have used metrics such as slugging percentage or on-base plus slugging percentage (for hitters) and strikeout-to-walk ratio or earned run average (for pitchers), to estimate players' marginal products (Scully, 1974; Zimbalist, 1992). Wins Above Replacement, a relatively recent invention, is an improvement over those measures. WAR includes the aforementioned statistics and applies more appropriate weights to them, and it also includes other areas of performance that are not used in previous models, such as baserunning and fielding ability.

---

[6] There are several versions of WAR with slightly different weighting vectors, including ones calculated by Fangraphs and Baseball Prospectus. Baseball-Reference's WAR was used in this study because it was easiest to tabulate; it has not been proven to be a worse estimate of value than any other form of WAR, and there is no indication that its usage, compared to other advanced performance metrics, will bias the results of this study in any particular way.

Another advantage of WAR is the "replacement-level" standard referenced in its name. Any system that attempts to measure a player's contribution to team wins must implicitly create a "counterfactual" estimate of how the team would have performed without him (Zimbalist, 1992). Scully's (1974) model measures players against a baseline slugging percentage of 0, suggesting that replacement players would never get a single hit in as many at-bats. Zimbalist (1992) critiques this assumption but comes up with another flawed solution, measuring a player's performance with the difference between his on-base plus slugging percentage and his team's rate, which implicitly assumes that his replacement would be an average player. The result of this methodology is that any below-average player has a negative marginal product, which is unrealistic. By definition, roughly half of major-league players are below average, so it is not reasonable to assume that an average player could easily be found as a replacement, nor that a slightly below-average player produces negative value for his team.

In practice, if a team did not sign a free agent to fill an open position, its alternative would be to promote a player from the minor leagues to play for minimal cost. This is the theory behind Wins Above Replacement, which constructs a baseline to represent the expected performance of a minor-league player who is cheap and easy to obtain — the practical "replacement" for an established player (Forman, 2012). The exact level of a replacement player's performance has been debated (Grosnick, 2012), but studies agree that it falls well above Scully's zero standard and well below Zimbalist's average standard, as determined from performance data (Woolner, 2006; Cameron, 2013).

A final advantage of Wins Above Replacement is that, by encapsulating all of a player's productive skills in one properly weighted metric, it allows a convenient measure of risk to be created. Before WAR and similar all-encompassing statistics were introduced, Bollinger and Hotchkiss (2003) estimate a player's productivity using nine different statistics, which may provide a reasonably accurate measure of his marginal product but which make studying risk complicated; their measure of risk includes nine variance terms and several dozen covariance terms, which are difficult to interpret. Using one comprehensive statistic, a player's risk can be expressed much more intuitively as the variance of his expected WAR at a given time.

Determining expected marginal product

Using Wins Above Replacement to measure marginal product, a player's future productivity can be estimated by regressing his WAR in year *t* on observable characteristics through year *t-1*. The player's marginal product in previous years — lagged values of WAR, as well as his average WAR over the course of his career — will be included as regressors, as they are expected to be predictive of current-season WAR. Additionally, several specific measures of performance — such as home runs and stolen bases for hitters, and strikeouts and walks for pitchers — will be included as well, because some types of skill may be more predictive of future performance than others. Finally, since athletes' performances vary as they get older, first- and second-order terms of age will be included, accounting for the fact that the age-performance relationship is nonlinear (as demonstrated by Schulz et al. (1994), Fair (2008) and others).

The full model for estimating a hitter's expected WAR is:

$$WAR_{jt} = \beta_0 + \beta_1 WAR_{j,t-1} + \beta_2 WAR_{j,t-2} + \beta_3 WAR_{j,t-3} + \beta_4 \overline{WAR}_{j,t-1} + \beta_5 HR_{j,t-1} + \beta_6 SB_{j,t-1} + \beta_7 SO_{j,t-1} + \beta_8 3B_{j,t-1} + \beta_9 GP_{j,t-1} + \beta_{10} CHR_{j,t-1} + \beta_{11} AGE_{jt} + \beta_{12} AGE^2{}_{jt} + \beta_{13} EXP_{j,t-1} \tag{2.1}$$

where WAR_{jt} is player *j*'s Wins Above Replacement in year *t*, $\overline{WAR}_{j,t-1}$ is his average WAR per season through year *t-1*, $HR_{j,t-1}$ is his home runs in year *t-1* (and SB stands for stolen bases, *SO* for strikeouts, $3B$ for triples and GP for games played), $CHR_{j,t-1}$ is his career home runs through year *t-1*, AGE_{jt} is his age in season *t* and $EXP_{j,t-1}$ is his major-league experience (in years) prior to season *t*. Since the purpose of this regression is only to generate the best possible estimate of WAR_{jt} and the precision of the coefficients is not consequential, all performance variables that improve the predictive power of the model have been included, even though some (such as age and experience) are highly correlated. Three lagged values of WAR are expected to be significant predictors of future performance based on most existing projection systems (Slowinski, 2011). Several other variables, such as doubles and fielding position, were ultimately discarded because they did not improve the model's fit.

For pitchers, WAR is estimated by the following model:

$$WAR_{jt} = \beta_0 + \beta_1 WAR_{j,t-1} + \beta_2 WAR_{j,t-2} + \beta_3 WAR_{j,t-3} + \beta_4 SO_{j,t-1} + \beta_5 BB_{j,t-1} + \beta_6 HR_{j,t-1} + \beta_7 AGE_{jt} + \beta_8 AGE^2{}_{jt} + \beta_9 AGE^3{}_{jt} \tag{2.2}$$

where WAR_{jt} is player *j*'s Wins Above Replacement in year *t*, $SO_{j,t-1}$ is the number of batters he struck out in year *t-1*, and BB and HR denote walks and home runs allowed. Unlike for hitters, a third degree of AGE proved to be a significant predictor of pitchers' performance. As before, several other metrics were ultimately discarded because they did not improve the model's fit.

Determining predicted risk

To estimate the risk preferences of employers in the MLB labor market, of course, it is necessary to develop a measure of risk for a baseball player's future performance. Previous studies have used the observed variance of a player's performance over the course of his career (Bollinger & Hotchkiss, 2003) or over the previous three seasons (Burger & Walters, 2008). However, the volatility of a baseball player's past performance is not necessarily predictive of his future risk, particularly because his underlying talent is not constant over time. For example, a young player whose performance improves rapidly in his first few seasons will have a high observed variance; but because players typically improve early in their careers, his performance may have matched expectations each year, making him relatively low-risk.

A more appropriate measure of risk is the predicted variance of a player's future performance, which may vary by player and by season. When dealing with data that involves a time-series component — such as a player's performance over several seasons — variance can be predicted by the generalized autoregressive conditional heteroskedasticity (GARCH) model developed by Bollerslev (1986). The GARCH model is based on Engle's (1982) ARCH process, which estimates the conditional variance at time t as a function of the size of the squared residuals for times prior to t. GARCH extends the ARCH model by using past predicted variances, in addition to past squared residuals, to predict conditional variance at time t. This "adaptive learning mechanism" makes the GARCH process more flexible and allows it to use more information than an ARCH model of comparable length (Bollerslev, 1986).

Cermeño and Grier (2001) develop a methodology for applying a GARCH model to panel data; this approach is useful for predicting the variance of baseball performance, as each player's productivity is observed for each year of his career and all players' performances are independent of each other. A panel data GARCH (*p, q*) model, which includes *p* lags of expected variance and *q* lagged residuals, is expressed as:

$$y_{it} = \mu + \boldsymbol{\beta}\mathbf{x}_{it} + u_{it}$$

$$u_{it} = \sigma_{it}\varepsilon_{it}, \quad \varepsilon_{it} \sim N(0,1)$$

$$\sigma^2{}_{it} = \alpha + \sum_{m=1}^{q} \gamma_m u^2{}_{i,t-m} + \sum_{n=1}^{p} \rho_n \sigma^2{}_{i,t-n}$$

It is possible that positive residuals (seasons in which a player over-performs his expectation) predict future variance differently than negative residuals. In particular, many negative residuals (seasons in which a player under-performs) happen because of injury — if a player is predicted to have positive WAR but misses a large portion of the season because he is hurt, his actual value in that season will likely be less than expected. The following season, his performance may be riskier than a healthy player's, because not only is there uncertainty about the player's underlying skill level, but there also may be uncertainty regarding how well he has recovered to full health. This is analogous to the behavior of stock markets, in which negative shocks are followed by higher volatility than positive shocks are (Rabemananjara & Zakoïan, 1993).

This behavior can be described by a threshold GARCH model, which allows negative residuals to predict future variance differently than positive residuals. One such model is the GJR-GARCH model (Glosten, Jagannathan, & Runkle, 1993), which adapts the standard GARCH equation as follows:

$$\sigma^2{}_{it} = \alpha + \sum_{m=1}^{q} \gamma_m u^2{}_{i,t-m} + \sum_{m=1}^{q} I_{i,t-m}(\tau_m u^2{}_{i,t-m}) + \sum_{n=1}^{p} \rho_n \sigma^2{}_{i,t-n}$$

where $I_{i,t-m}$ is an indicator variable that equals 1 if the residual $u_{i,t-m}$ is positive and 0 if it is negative.

Using the most common GARCH specification, a GARCH (1,1) model, as well as the GJR threshold specification, the expected mean and variance of a baseball player's performance can be modeled as:

$$WAR_{jt} = \boldsymbol{\beta}\mathbf{x}_{jt} + u_{jt} \qquad \text{(2.1 and 2.2)}$$

$$u_{jt}{}^{2} = \varepsilon_{jt}{}^{2} * Var(WAR_{jt}), \quad \varepsilon_{it} \sim N(0,1)$$

$$\mathrm{Var}(WAR_{jt}) = \alpha + \gamma u^{2}{}_{j,t-1} + I_{i,t-1}(\tau u^{2}{}_{j,t-1}) + \rho \mathrm{Var}(WAR_{j,t-1}) \qquad (2.3)$$

where WAR_{jt} represents player *j*'s actual WAR in year *t*, $\boldsymbol{\beta}\mathbf{x}_{jt}$ represents the coefficients and performance categories used to estimate player *j*'s WAR in equations 2.1 (for hitters) and 2.2 (for pitchers), $\mathrm{Var}(WAR_{jt})$ represents the predicted variance of his estimated WAR in year *t*, $u^{2}{}_{j,t-1}$ represents the squared difference between his estimated and actual WAR in year *t-1*, and $I_{i,t-1}$is a dummy variable indicating if $u_{j,t-1}$is positive.

Empirical Results

Determinants of expected performance and risk

Table 3.1 shows the results of the GJR-GARCH process for predicting WAR and variance for hitters, as shown in equations 2.1 and 2.3. The sample includes 5,946 seasons by 693 players; a GJR-GARCH (1,1) model was chosen to predict variance over specifications with additional lagged variables based on the Bayes information criterion (see Appendix A).

The estimated coefficient of lagged predicted variance is significant ($t = 44.5$, $p < 0.001$), suggesting that a GARCH specification is a better fit than an ARCH model. The estimated coefficient of the threshold indicator is also significant ($t = 9.1$, $p < .001$), suggesting that the GJR-GARCH specification is appropriate, although the direction is different than predicted; players who over-perform expectations have more risk in the following season, not those who under-perform. The mean expected WAR for hitters in

Table 3.1: Predicting WAR and risk for hitters, 1987-2012

	WAR_t		Var(WAR_t)
WAR_{t-1}	.2035**	$Error^2_{t-1}$	.0777**
	(.0203)		(.0100)
WAR_{t-2}	.1658**	$Error^2_{t-1}*I_{t-1}$	.1492**
	(.0155)		(.0165)
WAR_{t-3}	.0881**	Var(WAR_{t-1})	.7895**
	(.0138)		(.0177)
AGE	-.2597**		
	(.0681)		
AGE^2	.0032**		
	(.0011)		
Constant	5.2463**	Constant	.0433†
	(1.0340)		(.0233)

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses. Additional predictors of WAR_t were excluded from this table for brevity; the full model is reported in Appendix B.*

Table 3.2: Predicting WAR and risk for pitchers, 1987-2012

	WAR_t		$\mathrm{Var}(WAR_t)$
WAR_{t-1}	.1722**	$Error^2_{t-1}$	-.0187
	(.0268)		(.0141)
WAR_{t-2}	.0812**	$Error^2_{t-1}*I_{t-1}$	.1045**
	(.0215)		(.0254)
WAR_{t-3}	.1045**	$\mathrm{Var}(WAR_{t-1})$	.9138**
	(.0189)		(.0623)
AGE	-2.8447**		
	(1.0840)		
AGE^2	.0812*		
	(.0348)		
Constant	33.34**	Constant	.0323
	(11.15)		(.1666)

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses. Additional variables used to predict WAR_t for pitchers are also reported in Appendix B.*

this sample is 1.19, ranging from -.91[7] to 6.82; the mean expected variance in WAR is 2.33, ranging from .40 to 16.06.

Table 3.2 shows the results of the GJR-GARCH process for determining estimated WAR and variance for the sample of 319 starting pitchers over 2,296 seasons, as shown in eqations 2.2 and 2.3. As with hitters, a GJR-GARCH (1,1) model was chosen to predict pitchers' variance based on the Bayes information criterion (see Appendix A). The estimated coefficient of lagged predicted variance (t = 14.7, p < .001) and the estimated threshold coefficient (t = 4.1, p < .001) were both significant, suggesting that a GJR-GARCH model is appropriate; as with hitters, positive residuals caused the predicted variance for pitchers to be greater than negative residuals did. The mean expected WAR for pitchers in this sample is 1.51, ranging from -.59 to 7.07; the mean predicted variance in WAR is 3.18, ranging from 1.44 to 8.22.

[7] It is possible for a player to have a negative WAR; nearly 30 percent of player-seasons in the hitting sample produced negative WAR, though most were only slightly less than zero. The practical interpretation of a negative WAR is that a player's measurable performance was worse than would be expected from the average "replacement player."

Observed risk preferences for hitters

Now that expected WAR and expected variance have been estimated, it is possible to analyze MLB teams' risk preferences with respect to labor productivity. If teams are risk-neutral in the labor market, predicted variance should not affect expected salary when controlling for expected productivity. If teams are risk-averse (or risk-loving) instead, then players with high expected variance should receive lower (or higher) salaries than other players of similar expected performance. This will be tested with the following regression model:

$$SALARY_{jt} = \beta_0 + \beta_1 \mathrm{E}\left[WAR_{jt}\right] + \beta_2 \mathrm{Var}\left(WAR_{jt}\right) + \beta_3 YEAR_t + \beta_4 EXP_{j,t-1} \quad (3.1)$$

where $SALARY_{jt}$ is the average annual salary of a contract signed by player *j* at time *t*, $\mathrm{E}\left[WAR_{jt}\right]$ is player *j*'s expected Wins Above Replacement in year *t*, $\mathrm{Var}\left(WAR_{jt}\right)$ is his predicted variance, $YEAR_t$ accounts for a yearly trend in salaries and $EXP_{j,t-1}$ signifies player *j*'s major-league experience prior to year *t*. The coefficient of particular interest is β_2, which represents the amount teams are willing to pay for additional variance. This measure should be negative if teams are risk-averse, positive if they are risk-seeking and roughly zero if they are risk-neutral.

To exclude all salaries that were not determined in a competitive market, the sample has been restricted only to contracts of players who were granted free agency. Players who were officially granted free agency and then re-signed with their previous team are included, since they could negotiate with other teams, but those who signed a contract extension with their original club before reaching free agency are not.[8]

[8] For example, All-Star third baseman Chipper Jones never officially reached free agency in his 20-year career; before each of his contracts expired, he signed an extension with his original team, the Atlanta Braves.

Table 3.3: Determinants of hitter salaries

	Avg. Salary
E[*WAR*]	2949834**
	(87570)
Var(*WAR*)	**125313**
(standardized)	**(80423)**
YEAR	140249**
(indexed)	(12267)
EXP	67565**
	(20605)
Constant	-1718482**
	(252742)
Adj. R^2	.66

****, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

The results of regression 3.1 for the sample of 1,219 free-agent contracts signed by hitters between 1993 and 2012 are reported in Table 3.3. (*YEAR* has been indexed so that 0 represents the first year of the sample in which contracts were signed, 1993, and Var(*WAR*) has been transformed so that its mean is 0 and its standard deviation is 1, for ease of interpreting coefficients.)

As expected, salary is strongly predicted by expected performance; in this sample, teams paid an average of $2.9 million per expected win in free agency. Average salaries, holding other factors constant, increased by $140,000 per year from 1993 to 2012; more veteran players also earned slightly higher salaries in free agency, even controlling for productivity, as an extra year of major-league experience was associated with a $68,000 increase in wage.

Most notably, a one-standard deviation increase in predicted variance is associated with a $125,000 increase in yearly salary, but the effect is not statistically significant (t = 1.56, p = .12). Based on this result, the assumption that MLB employers are risk-neutral in hiring labor cannot be rejected.

Table 3.4: Determinants of hitter salaries, by contract length

Contract Length:	**1-year**	**Multi-year**
	Avg. Salary	Avg. Salary
E[*WAR*]	1443147**	3376766**
	(83305)	(176605)
Var(*WAR*)	**287470****	**-105012**
(standardized)	**(76750)**	**(139967)**
YEAR	79521**	311084**
(indexed)	(8307)	(32186)
EXP	63332**	141677*
	(13267)	(64934)
Constant	-661422**	-4087880**
	(172302)	(760670)
Adj. R^2	.44	.68

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

A different pattern is found when the data is split by contract length, however. Table 3.4 shows the estimated effects of expected performance and predicted variance for the 885 contracts in the sample that guaranteed only one year of employment, as well as the sample's 334 multi-year contracts. In short-term contracts, baseball teams appear to be risk-seeking with respect to labor — a one-standard deviation increase in predicted variance corresponds to a $287,000 increase in salary (t = 3.75, p < .001). This behavior does not extend to longer-term deals, however; for multi-year contracts, riskier players are paid slightly less than safer players, though the effect is not nearly significant (t = -.75, p = .45). An appropriate statistical test (Paternoster, Brame, Mazerolle, & Piquero, 1998) confirms that the estimated variance coefficient for one-year contracts is significantly greater than the corresponding coefficient for multi-year contracts (t = 2.46, p = .007).

The model developed earlier assumes a linear relationship between marginal product and marginal revenue (each additional win increases a team's revenue by β_1), and therefore predicts a linear relationship between marginal product and salary. Some studies, however, have found empirical evidence of nonlinearity in the relationship between productivity and salary in baseball — specifically, that players are compensated as if the returns to productivity are increasing (MacDonald & Reynolds, 1994; Scully, 1989; Silver, 2005). There are a couple of possible justifications for this behavior. Most notably, when a team signs a free-agent player, it incurs not only the monetary cost of that player's salary but also the opportunity cost of his place on the roster. Each club may employ only 25 major-league players at a given time, so its total number of employees is fixed; therefore, players with extraordinary talent may be perceived as more valuable, because they provide lots of productivity at the opportunity cost of only one roster spot. (In an extreme case, one player worth 25 wins would almost certainly be preferred to 25 players worth one win apiece, because a team with the former would have many remaining positions at which it could employ other players, while a team with the latter would be filled.) Additionally, it is possible that superstars generate additional revenue beyond their contribution to team wins, such as if their individual popularity attracts fans or sells merchandise. If teams believe that the returns to productivity are increasing, then salaries can be more appropriately estimated using a quadratic model:

$$SALARY_{jt} = \beta_0 + \beta_1 \mathrm{E}\left[WAR_{jt}\right] + \beta_2 \mathrm{E}\left[WAR_{jt}\right]^2 + \beta_3 \mathrm{Var}\left(WAR_{jt}\right) + \beta_4 YEAR_t + \beta_5 EXP_{j,t-1} \tag{3.2}$$

If this model is more accurate than a linear relationship, the previously estimated coefficients of variance may have been biased. The results of regression 3.2, again separated by contract length, are shown in Table 3.5.

Table 3.5: Determinants of hitter salaries (quadratic model)

Contract Length:	**All**	**1-year**	**Multi-year**
	Avg. Salary	Avg. Salary	Avg. Salary
E[*WAR*]	1829108**	685136**	2440873**
	(136544)	(138157)	(363218)
E[*WAR*]^2	400779**	541436**	235764**
	(38513)	(79998)	(80237)
Var(*WAR*)	**58976**	**308872****	**-112124**
(standardized)	**(77351)**	**(74934)**	**(138392)**
YEAR	141233**	79244**	331315**
(indexed)	(11759)	(8104)	(31819)
EXP	62145**	53854**	156789*
	(19757)	(13107)	(64399)
Constant	-1458570**	-516067**	-3703295**
	(243548)	(169443)	(763300)
Adj. R^2	.69	.47	.69

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

The coefficient of the quadratic term is positive, and it is statistically significant in all samples; the fit of the quadratic model, as measured by adjusted R^2, is marginally better than the linear model's fit. Therefore, it is reasonable to conclude that teams act as if the marginal returns to a player's productivity are increasing.

The estimated coefficients of predicted variance, however, are largely unchanged. Within the entire sample, the estimated affinity for risk remains positive but statistically insignificant (t = .76, p = .45). But in one-year contracts, teams still appear to be risk-seeking in labor; a one-standard deviation increase in predicted variance is associated with a \$309,000 increase in expected salary (t = 4.12, p < .001). For multi-year contracts, the estimated effect of additional risk is negative but not statistically significant (t = -.81, p < .42).

Another limiting factor of the salary-determination model developed earlier is that β_1, the value of a marginal win, is assumed to be constant over time. This is not likely to be the case. Over the two decades included in this study, general inflation drove up nominal price levels substantially; if ticket prices and other sources of team income increased with inflation, the nominal marginal revenue generated by an additional win should have correspondingly increased. More dramatically, Major League Baseball's annual revenue grew at a rate far outpacing inflation; due to increased domestic and international demand for professional baseball, as well as new revenue sources such as digital content (Slater, 2012), MLB's reported revenue increased from $1.4 billion in 1995 to $7 billion in 2010 — a rise of 400 percent — while the Consumer Price Index increased by 43 percent during that time (Brown, 2011). Finally, since 1996, MLB has had some form of revenue sharing — in which each team pays a percentage of its revenue into a common pool that is then distributed equally to all 30 teams, effectively transferring money from high-revenue teams to low-revenue teams — but the terms of the agreements have changed over time (Dosh, 2007). When the degree of revenue sharing is greater, teams keep a smaller fraction of the revenue they make, reducing the private value of a marginal win; this provides another cause of change in the revenue function over time.

A more flexible model is presented in equation 3.3, where interaction terms are added between the year and both WAR and squared WAR, to account for the fact that the price teams are willing to pay for an extra win may have increased over time:

$$SALARY_{jt} = \beta_0 + \beta_1 \mathrm{E}\left[WAR_{jt}\right] + \beta_2 \mathrm{E}\left[WAR_{jt}\right]^2 + \beta_3 \mathrm{E}\left[WAR_{jt}\right] * YEAR_t + \beta_4 \mathrm{E}\left[WAR_{jt}\right]^2 * YEAR_t + \beta_5 \mathrm{Var}\left(WAR_{jt}\right) + \beta_6 YEAR_t + \beta_7 EXP_{j,t-1} \quad (3.3)$$

Table 3.6: Determinants of hitter salaries (interacted quadratic model)

Contract Length:	**All**	**1-year**	**Multi-year**
	Avg. Salary	Avg. Salary	Avg. Salary
E[*WAR*]	234734	-342093	-721748
	(310300)	(340874)	(819531)
E[*WAR*]^2	239864*	382886*	461430*
	(92774)	(200964)	(192968)
E[*WAR*]**YEAR*	130554**	82164**	283273**
	(24690)	(25863)	(68848)
E[*WAR*]^2**YEAR*	19549*	19441	-18326
	(7705)	(15837)	(16566)
Var(*WAR*)	**42140**	**271081****	**-140252**
(standardized)	**(70447)**	**(72361)**	**(125626)**
YEAR	31168*	37860**	24541
(indexed)	(14322)	(9568)	(58380)
EXP	74766**	54241**	163083**
	(18021)	(12555)	(58456)
Constant	-327678	-59364	-69938
	(242823)	(174325)	(877561)
Adj. R^2	.74	.51	.74

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

As before, *YEAR* has been transformed so that *YEAR*=0 for 1993, the first year in which contracts were signed in the sample under study. Thus, the coefficients β_1 and β_2 represent the estimated salary teams were willing to pay for marginal wins in 1993 and the coefficients β_3 and β_4 signify the yearly change in these values. The results of regression 3.3 are shown in Table 3.6.

The interaction terms of *YEAR* and both degrees of E[*WAR*] are significantly positive in the full sample, suggesting that the marginal value of wins increased between 1993 and 2012; adding these interaction terms further improved the fit of the salary-prediction model, as measured by adjusted R^2. In this interacted model, teams still appear to be risk-seeking when signing one-year contracts, paying about $270,000 more for a one-standard deviation increase in predicted variance ($t = 3.75$, $p < .001$); they are estimated to be risk-neutral or slightly risk-averse when signing multi-year contracts (t = -1.12, p = .27).

Observed risk preferences for pitchers

Further information about teams' risk preferences can be obtained from the sample of starting pitchers' contracts. Pitchers and hitters perform different functions for their teams, so their output is determined in different ways, as evidenced by the differences in regressions 2.1 and 2.2. However, their marginal products can be quantified with the same metric, Wins Above Replacement, and the average WARs and variances in this sample were roughly similar for hitters and pitchers. Therefore, there is no reason to expect that teams' risk preferences in signing pitchers are substantially different than their risk preferences for hitters.

Using the estimates of expected performance and predicted variance developed for pitchers in regressions 2.2 and 2.3, teams' risk preferences for pitchers can be analyzed with the same methods previously developed for hitters. Table 3.7 shows the results of regression 3.1, which is linear in expected WAR, applied to the sample of 380 contracts signed by starting pitchers, including 253 one-year contracts and 127 multi-year deals.

Across all sections of the sample, the estimated effect of additional risk is slightly greater for pitchers than for hitters; for one-year pitching contracts, a one-standard deviation increase in variance predicts a $482,000 increase in salary if other variables are held constant, about $200,000 greater than the corresponding effect for hitters in Table 3.4. The overall pattern, however, is the same: Teams are willing to pay a significant

Table 3.7: Determinants of pitcher salaries, by contract length (linear model)

Contract Length:	**All**	**1-year**	**Multi-year**
	Avg. Salary	Avg. Salary	Avg. Salary
E[*WAR*]	3425641**	2298199**	2791838**
	(188990)	(230396)	(316325)
Var(*WAR*)	**388267***	**482112****	**183649**
(standardized)	**(154550)**	**(155305)**	**(254204)**
YEAR	235062**	140321**	505568**
(indexed)	(29122)	(25697)	(54740)
EXP	-76669†	26549	-86503
	(44513)	(36855)	(108220)
Constant	-1565708**	-1264824**	-1843321
	(530307)	(430767)	(1313879)
Adj. R^2	.60	.46	.60

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

premium for riskier pitchers on short-term contracts (t = 3.1, p = .002), but they do not appear to have a strong risk preference in multi-year deals (t = .72, p = .47).

The linear salary model for pitchers may be incomplete for the same reasons outlined in the previous section. For pitchers as with hitters, the returns to productivity may be increasing, because highly productive players leave more available roster sports than a combination of less productive players, and because superstars may generate additional cash flows. Additionally, if marginal revenue per win has grown over time, the effect of expected performance on salary should have increased in the pitching market as well. Both of these adjustments are addressed in Table 3.8, which applies regression 3.3 (with a quadratic term of expected WAR, as well as interactions between year and both degrees of expected WAR) to starting pitchers' contracts.

Because there are many regressors and relatively few observations, particularly in the multi-year contract sample, few coefficients are estimated to be statistically significant. But the adjusted R^2 for each sample is greater than that of the corresponding linear model in Table 3.7, suggesting the quadratic and interacted model is a better fit.

Table 3.8: Determinants of pitcher salaries (interacted quadratic model)

Contract Length:	**All**	**1-year**	**Multi-year**
	Avg. Salary	Avg. Salary	Avg. Salary
E[*WAR*]	2099224*	16216	1594394
	(971771)	(1638224)	(2471944)
E[*WAR*]^2	-251343	262254	-183196
	(279466)	(789150)	(521867)
E[*WAR*]**YEAR*	86174	-7111	54612
	(78823)	(128218)	(206801)
E[*WAR*]^2**YEAR*	35347	76738	30570
	(22276)	(60069)	(42273)
Var(*WAR*)	**294703***	**409824****	**67252**
(standardized)	**(146750)**	**(145800)**	**(247289)**
YEAR	68641	81800	269405
(indexed)	(58008)	(56153)	(224161)
EXP	-50663	17538	-78125
	(42253)	(34637)	(104398)
Constant	166194	233226	1125601
	(763143)	(742364)	(2873920)
Adj. R^2	.64	.53	.63

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

Teams are still estimated to be significantly risk-seeking when negotiating one-year contracts (t = 2.81, p = .005), while their behavior is approximately risk-neutral in multi-year deals (t = .27, p = .79).

For both hitters and pitchers, the effect of teams' risk preferences is relatively small in context. According to these results, teams have been willing to pay an average of about $2-3 million for an additional expected win on the free-agent market, with the exact value rising over time and increasing for highly productive players; meanwhile, even in one-year deals, they are willing to pay only about $300,000-$400,000 for a one-standard deviation increase in predicted variance. But across independent samples of hitters and pitchers, and in several different salary models, the pattern persists: MLB teams act as if they are risk-seeking with respect to labor — but only when negotiating short-term contracts.

DISCUSSION

Existing theories of risk-seeking employers

Why might employers be risk-seeking when making hiring decisions in the Major League Baseball labor market? Lazear (1998) proposes a model in which firms prefer high-variance workers to low-variance workers of comparable expected skill, because riskier employees have "option value." If they turn out to perform worse than expected, they can be fired and replaced, but if they perform better than expected, they can be retained; therefore, the total value obtained from a high-variance worker can exceed his mean expected productivity.

Lazear's model requires two necessary conditions for its risk-seeking predictions to hold. The first is that if productivity is public information, workers' mobility must be restricted; otherwise, if a risky worker performed better than expected, another employer could offer him a new salary commensurate with his revealed productivity and hire him away from his original firm. This condition is fulfilled, to some degree, in the MLB labor market; once players have signed a contract with one team, they cannot negotiate with any other employers until their original contract expires.

The second condition of Lazear's prediction is that employers can terminate workers at little cost, which is not generally true in professional baseball. MLB contracts are fully guaranteed, so teams cannot simply fire players if they are underperforming; if a player is released, he is still entitled to his entire negotiated salary. (Lazear's model would apply more directly to leagues in which many contracts are not guaranteed, such as the National Football League.) Baseball players may still have "option value" in another sense, however. Because teams can only use a certain number of players (major-league rosters are capped at 25 players, and only nine can play at any given time), using the services of one worker comes with an opportunity cost of not playing another. If a player under-performs, a team is still required to pay the nominal cost of his salary, but it can lessen the opportunity cost of his playing time by reducing his role and giving chances to other players, who might be more productive.

If this opportunity cost is a major factor in teams' hiring decisions, risky workers could be preferred to lower-variance players, because their roles can be maximized if they perform well and minimized if they perform poorly. However, this theory does not explain why risk-seeking behavior would be present only in short-term contracts. In fact, Lazear's model predicts that firms actually have a stronger preference for risky workers in long timeframes than in shorter ones, because they have more time to realize the benefits of "option value."

A second possible explanation of risk-seeking patterns comes from the variance in teams' predictions of a player's productivity. The salary model developed in this paper assumed that all parties in the labor market share the same beliefs of a player's expected performance. All teams certainly have access to most of the same information, but that does not necessarily mean they form the same beliefs — some employers may apply different weights to various factors when predicting future performance from past data. If that is the case, teams will enter free-agent bidding with different expectations of a player's productivity. Because each player will sign with the highest bidder, the final contract will be determined by only the most optimistic team's belief.

If the variance in teams' expectations is positively related to a player's predicted variance, then the highest-bidding team's belief will be further from the market average for risky free agents than for safer ones. If teams do not discount their bids for the possibility of a "winner's curse" in the free-agent market — which is possible, as Thaler (1988) finds that winner's curses exist in several other competitive industries — or if teams do not factor predicted variance into the discounting of their bids (i.e., they discount bids by the same amount for high-risk and low-risk players), then patterns consistent with risk-seeking behavior can be explained by the fact that only the highest bidder's belief is made public.

Such a mechanism would explain the appearance of risk-seeking behavior in general, but it also fails to explain why teams do not pay higher salaries for risky players in long-term contracts. If teams' beliefs of future productivity are less homogeneous for players with greater predicted variance, and if that causes higher salaries for risky players because of a type of winner's curse, then players with greater predicted variance should receive higher salaries regardless of contract length.

As another explanation of risk-seeking behavior in baseball's labor market in particular, Burger and Walters (2008) suggest that teams may "overweight good years and fall victim to over-optimism bias" (p. 242). If teams overweight good seasons, they will pay more for a player who over-performs in some years and under-performs in others than for a player who matches expectations in all seasons: the former player's "good" seasons will be better, and the fact that his "bad" seasons are worse will be given less weight. Burger and Walters' theory could explain a general association between predicted variance and salary, but it should also apply regardless of contract length, so it does not fit the pattern found in this study.

Effect of non-linear revenues on risk preferences

A new theory that could explain risk-seeking behavior in Major League Baseball is based on the relationship between wins and team revenue. Until this point, revenue has been assumed to be a linear function of wins, but research by Silver (2006) suggests that the returns to winning are nonlinear. Specifically, teams receive a constant marginal revenue for each additional win, plus a large additional lump sum if they qualify for the playoffs. Therefore, if a marginal win pushes a team above the playoff threshold, it yields a huge bonus, but if a marginal win is instead located away from the threshold, it yields a much smaller increase in revenue (which Silver estimates is roughly $1/40^{th}$ the size of the playoff bonus).

Because playoff eligibility depends on relative team performance (as of 2012, the six division winners and the teams with the two best remaining records in each league advance to the postseason), the playoff threshold varies from season to season and is not known in advance. Silver estimates the likelihood of a team reaching the playoffs at a given number of wins with a logistic regression model based on past outcomes (p. 191), which is replicated in Figure 4.1.

From this function and Silver's two-stage revenue model, a team's total revenue function can be derived. If the constant marginal revenue per win is m, and the playoff bonus is B, a team's expected revenue with w wins is $w*m + B*P(w)$, where $P(w)$ is the probability of making the playoffs with w wins, as estimated by the logistic function in Figure 4.1. A team's total revenue function is therefore nonlinear, as shown in Figure 4.2 (also from Silver, p. 192).

Figure 4.1: Expected playoff odds as a function of wins

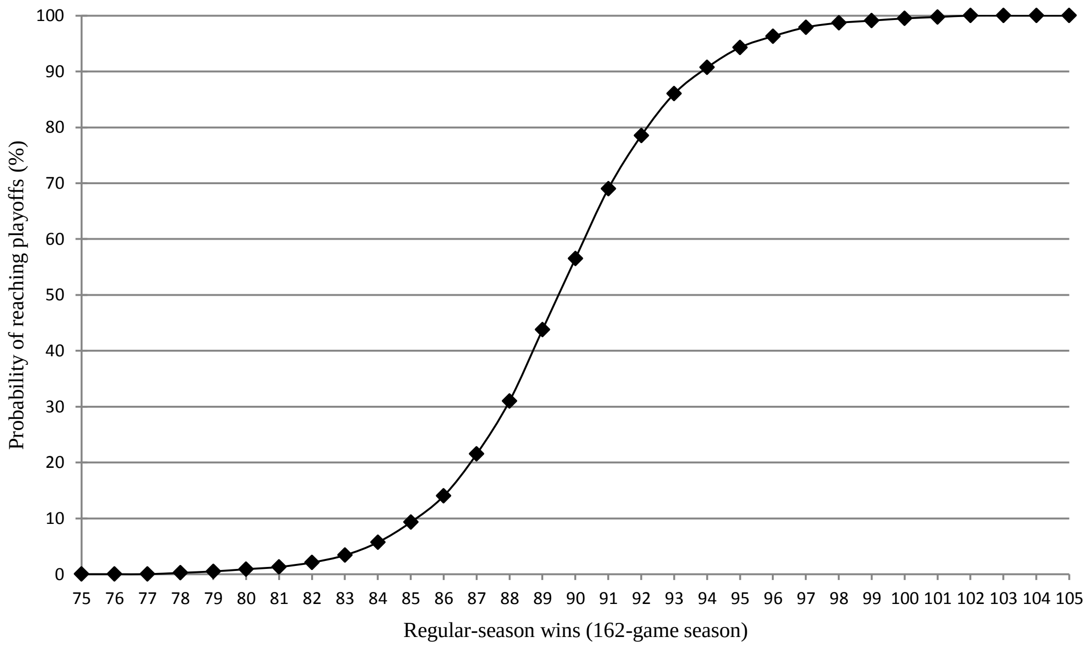
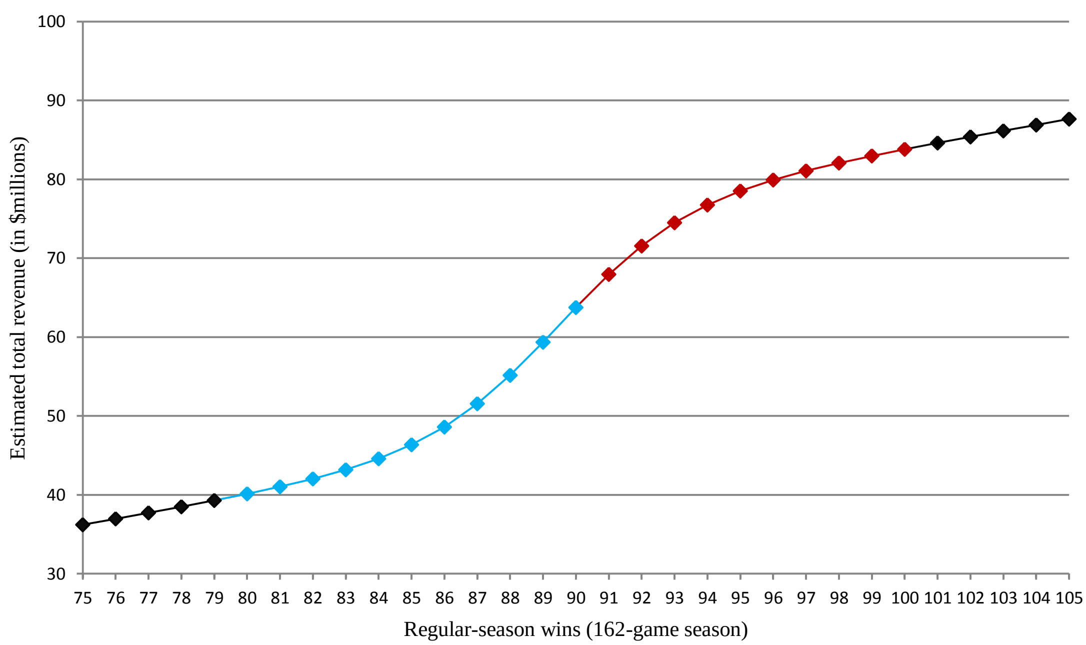


Figure 4.2: Expected total revenue as a function of wins



For any team in the convex (blue) region, the marginal value of each win is greater than the marginal value of the previous win; the reverse is true in the concave (red) region. Because the total revenue function is convex from about 81-90 wins, teams expecting to fall in that range should be risk-seeking with respect to wins if they are risk-neutral with respect to revenue.[9] Similarly, teams expecting to fall in the 91-100 win range should be risk-averse with respect to wins, if they are risk-neutral in revenue.

If an equal number of teams fell in each category, MLB employers would be risk-neutral on aggregate, as the risk-seeking and risk-averse teams would approximately cancel each other out. But empirically, teams are significantly more likely to be in the risk-loving section of the marginal revenue curve than the risk-averse section, as shown in Figure 4.3.

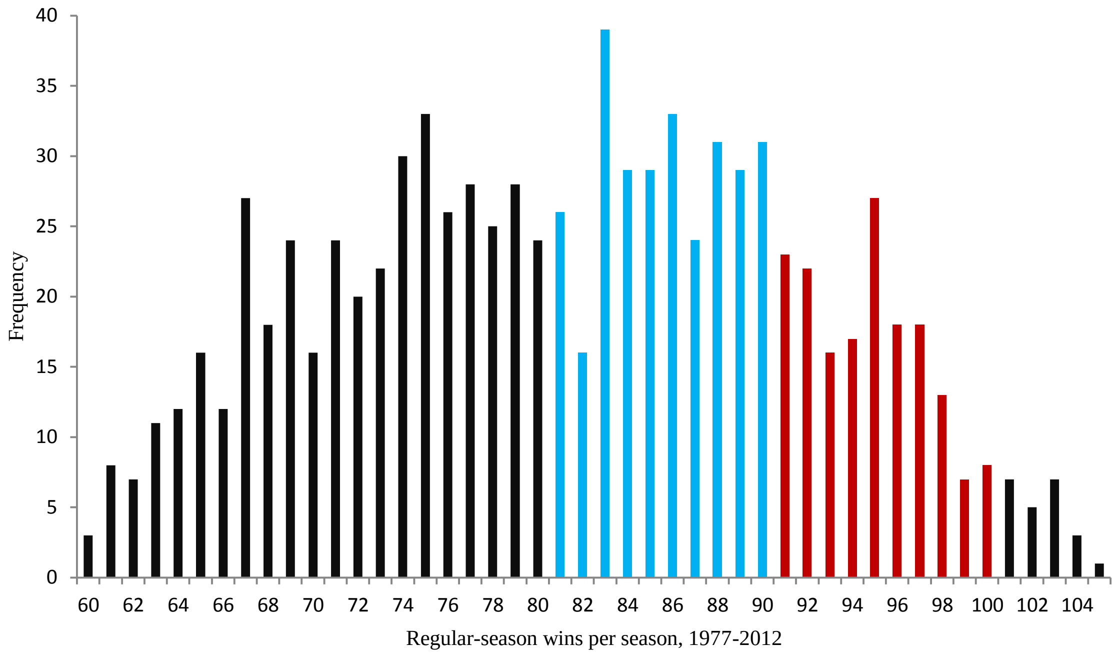
From 1977-2012[10], 287 teams finished with 81-90 wins (the convex region of the total revenue curve), while only 169 finished with 91-100 wins (the concave region). Therefore, it is reasonable to assume that more teams expect to be in the 81-90 win range than the 91-100 win range, so more teams expect to face a convex total-revenue curve than a concave one. If teams are risk-seeking in the convex region, risk-averse in the concave area and risk-neutral in the approximately linear areas, then the aggregate risk preference of MLB teams should be slightly risk-seeking.

---

[9] As an example, consider a team expected to win 85 games, which corresponds to a total revenue of $46.3 million in Figure 4.2. Assume that team can choose to employ one of two players; the first has a guaranteed marginal product of 1 win, while the second has a marginal product of 2 wins half the time and 0 wins half the time. By signing the first player, the team will expect to win 86 games, for an expected revenue of $48.6 million; by signing the second, it will expect to win either 85 and 87 games with an equal likelihood, for an expected revenue of ½(51.5) + ½(46.3) = $48.9 million. Both players have the same expected marginal product of one win, but a team in the convex region is better off signing the riskier worker. (By similar logic, a team in the concave region can be shown to prefer safe players to risky ones.)

[10] Data from 1981, 1994 and 1995 were excluded from Figure 4.3 because teams played fewer than 162 total games in those seasons due to labor stoppages. Four teams with more than 105 wins, and 26 teams with fewer than 60 wins, are not represented in Figure 4.3.

Figure 4.3: Distribution of wins per season, 1977-2012



Additionally, this theory might be able to explain why risk-seeking behavior is only observed in one-year contracts. Teams should be able to predict approximately how many games they will win in the upcoming season, based on the previous year's results and the other players on their roster. But forecasting wins two or more seasons in the future is much more difficult, because players' performance trajectories are harder to predict further out and more players may change teams. Because of this added uncertainty, it is possible that teams ignore the effect of their position on the total revenue curve when determining multi-year contracts — or include it at a reduced weight — even if they use that information for one-year contracts.

To test this explanation of risk-seeking behavior, the estimated effects of risk on salary can be compared for contracts signed by different groups of teams. If the nonlinear shape of the marginal revenue curve is what causes risk-seeking behavior in short-term contracts, then teams expecting to face a linear (or concave) total revenue curve should exhibit risk-neutral (or risk-averse) behavior, while teams expecting to fall in the convex region should be most willing to pay a premium for risky players.

In this model of revenue, teams make valuations in part based on their expectations of their own future performance — how many games they expect to win with and without signing a free-agent player — which are not readily available. But it is likely that a team's expected performance is strongly correlated with its performance in the previous season. Therefore, "non-contenders" will be defined as teams that won 75 or fewer games in the previous season (and should be in the linear region of the revenue curve); "contenders" will be teams that won 81-90 games in the previous season (the convex region); and "favorites" will be teams that won at least 95 games in the previous season (the concave region).[11] "Contenders" are expected to exhibit the strongest risk-seeking behavior; "non-contenders" are expected to have risk-neutral preferences and "favorites" are expected to be risk-averse.

Regression 3.3, which tested the effect of career variance on salary with an interacted and quadratic model, has been run separately for contracts signed by teams falling in each of the three categories[12]. The results (for one-year contracts only) are presented in Table 4.1 (for hitters) and Table 4.2 (for pitchers).

[11] Gaps are intentionally left between the three groups to avoid classifying teams whose expected positions on the revenue curve are ambiguous. For example, of teams that won 90-94 games from 1996-2011, only 31 percent surpassed 90 victories again the following season, suggesting that many such teams should not have expected to be above the 90-win threshold again; a majority of teams that won at least 95 games surpassed 90 in the following season, however, so it is reasonable to assume those teams expected to be in the concave region of the total revenue curve.

[12] This analysis only includes players signed in the era of eight-team playoffs, 1995-2011, because Silver's revenue analysis was based on that playoff structure; previously, only four teams advanced to the postseason in each year, so the shape of the total revenue curve was likely different. Additionally, players signed before the 1995 and 1996 seasons were excluded because labor stoppages in the preceding years (1994 and 1995) caused teams to play significantly fewer than 162 games, changing the estimated cutoff points for "contenders" and "favorites."

Table 4.1: Determinants of hitter salaries, by team expectation (1-year contracts)

	Non-contenders	***Contenders***	***Favorites***
	Avg. Salary	Avg. Salary	Avg. Salary
E[*WAR*]	1322555	-513270	-1261765
	(804808)	(939888)	(1199440)
E[*WAR*]^2	-2027344**	-25670	1297716
	(603702)	(547317)	(894141)
E[*WAR*]**YEAR*	-48837	112396	113707
	(58025)	(72071)	(91632)
E[*WAR*]^2**YEAR*	195105**	33472	-8635
	(46114)	(41511)	(62463)
Var(*WAR*)	**374880****	**529692****	**106836**
(standardized)	**(117730)**	**(159893)**	**(186378)**
YEAR	37935*	31618	15877
(indexed)	(17482)	(25150)	(186378)
EXP	72193**	79233**	33848
	(20139)	(26714)	(31567)
Constant	-172831	-73350	185408
	(322720)	(402803)	(530303)
Adj. R^2	.50	.53	.64
Observations	283	243	109

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

The results are somewhat consistent with the predictions — when signing both hitters and pitchers, "contenders" were the most risk-seeking group. Teams that had finished with 81-90 wins in the previous season, and were therefore most likely to face a convex revenue function in the following year, paid $500,000-$600,000 for a one-standard deviation increase in predicted variance. Also as expected, "favorites" were the least risk-seeking group, while "non-contenders" fell in the middle.

However, "non-contenders," who are expected to face a linear revenue function, were still risk-seeking. For hitters signed by teams that won 75 or fewer games in the previous season, risk is still a significant predictor of salary (t = 3.18, p = .002). For pitchers, the estimated effect is not statistically significant (in part because the sample size is much smaller), but its magnitude is roughly the same; "non-contenders" are

Table 4.2: Determinants of pitcher salaries, by team expectation (1-year contracts)

	Non-contenders	***Contenders***	***Favorites***
	Avg. Salary	Avg. Salary	Avg. Salary
E[*WAR*]	-2678333	-750990	-23300000
	(3730793)	(2569405)	(13700000)
E[*WAR*]^2	1118843	1250317	14200000*
	(2276272)	(1248102)	(6219512)
E[*WAR*]**YEAR*	376082	143939	1225564
	(305223)	(214454)	(936099)
E[*WAR*]^2**YEAR*	-149017	-61769	-657997
	(188979)	(104802)	(415716)
Var(*WAR*)	**293940**	**594216***	**-115584**
(standardized)	**(260327)**	**(234393)**	**(537349)**
YEAR	-17124	52516	-411555
(indexed)	(104911)	(98073)	(482115)
EXP	-3404	29228	47816
	(56309)	(59689)	(93165)
Constant	1075759	453565	8110232
	(1322020)	(1240423)	(7125027)
Adj. R^2	.26	.43	.72
Observations	76	77	33

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

estimated to pay about $300,000-400,000 for a one-standard deviation increase in predicted variance. Similarly, even though "favorites" face a concave revenue curve, they do not appear to avoid risk when signing labor contracts. The estimated effect for these teams is ambiguous — positive but small for hitters, and negative but small in a very limited sample of pitchers — but it clearly does not suggest a significant aversion to risk.

Tables 4.1 and 4.2 provide evidence that some of employers' risk preferences in the MLB labor market can be explained by nonlinearity in the total revenue function. As predicted by the theory, teams expecting to be in the convex region of the revenue curve value risk more highly than teams expecting a linear revenue curve, which in turn are more risk-seeking than those facing a concave function. But if that theory explained all of the risk-seeking behavior in MLB, teams in the linear range would have no observed preference for risk, and teams in the concave region would be strictly risk-averse. Instead, teams expecting a linear revenue function are still risk-seeking, and teams expecting a concave revenue function appear to have no preference for risk, suggesting that much of the aggregate risk-seeking behavior of MLB teams is not explained by this theory.

Incorporating risky players' contract preferences

The observed pattern of risk-seeking behavior in short-term contracts may be best explained by a theory that incorporates players' preferences as well as teams'. In the models developed to this point, players have been assumed to be passive agents in determining contracts; the hiring process has been treated as a perfectly competitive market in which teams make salary offers and players simply select the highest option. But players can have an active role in negotiations, as evidenced by the fact that nearly every player hires an agent to negotiate contracts on his behalf ("Agency Database," 2013). Therefore, a player's preferences — such as his desire for a long- or short-term contract — may ultimately affect the terms of a contract.

If baseball players are risk-averse with respect to income, as most individual economic agents are, then players whose predicted performance is relatively risky should prefer long-term contracts, because they guarantee a certain stream of future income. If those players signed short-term contracts instead, their future income would depend in part on their current-period performance, which carries substantial risk. Players whose expected performance is very certain, on the other hand, should be relatively indifferent between short- and long-term contracts; because their performance is predictable, they can estimate their future salaries rather precisely even without a long-term guarantee.

The final assumption in this theory is that teams value risky players more highly than consistent players, because they provide option value with respect to playing opportunities (as discussed earlier). When negotiating one-year contracts, teams will therefore be willing to pay risky players a higher salary than consistent players with the same expected productivity. Additionally, risky players may demand a higher salary for short-term contracts than consistent players, because the opportunity cost of not signing a long-term contract is larger for the former group. Both of these effects act in the same (positive) direction, so risky players will receive a salary premium in short-term contracts.

For long-term contracts, teams again prefer risky players to more consistent ones, so they are willing to offer them a higher salary. But because they have a strong preference for long-term contracts, risky players may be willing to take a discount to sign a long-term contract and gain income security. This effect acts negatively on risky players' salary, compared to consistent players, who are indifferent to contract length and therefore have no reason to sign long-term contracts at a discounted rate. Therefore, the bargaining outcome is ambiguous; if the effects of team preferences (positive) and player preferences (negative) approximately cancel out, risky players will ultimately negotiate similar long-term salaries as consistent players with the same expected productivity.

Thus, the results of this study — which found that risky baseball players are given greater salaries, but only in one-year contracts — are consistent with a theory of salary determination that includes players' contract preferences as well as teams' risk preferences. If teams are risk-seeking in labor and players are risk-averse with respect to future incomes, then risky players should receive higher wages in short-term contracts, but their salaries in long-term contracts may be similar to those of consistent players.

CONCLUSION

In existing models of salary determination, firms are often assumed to be risk-neutral with respect to labor productivity. But this assumption has rarely been tested, in part because the data required to measure an employee's risk are rarely available. In this study, a sample of 1,599 free-agent contracts in Major League Baseball is used to test firms' risk preferences in an empirical setting. Employers in MLB are found to be risk-seeking with respect to labor, but only in short-term contracts. When determining contracts with a length of only one season, teams are estimated to pay $300,000-$400,000 more for a player whose predicted variance in productivity is one standard deviation greater than that of another player with the same expected marginal product. In determining multi-year contracts, however, the estimated effect of risk is ambiguous and insignificant.

These results support and clarify the work of Burger and Walters (2008), who find that volatile players receive higher salaries in the MLB labor market but do not address contract length. In showing that risky workers are preferred in some conditions, this study also provides qualified evidence for Lazear's (1998) model of risk-seeking firms, although Lazear's and other existing theories do not explain why risk-seeking behavior would be limited to short-term contracts. A revised theory is proposed in which firms prefer risky workers in all situations and risky workers prefer the security of long-term contracts; under these conditions, risky workers are expected to receive higher wages than consistent workers of the same expected productivity in short-term contracts, but high-risk and low-risk players may receive similar salaries in long-term contracts, which is consistent with the observed pattern.

A few limitations of this study could have affected its results. First, the salary-determination model developed earlier assumes that all potential employers use a rational process for determining a player's expected value and risk (i.e., all relevant variables are considered and weighted appropriately). But several studies have found evidence that some characteristics of players may be overvalued or undervalued in the free-agent market (Gassko, 2011; Hakes & Sauer, 2006). This suggests that the performance-estimation models actually used by teams differed in some ways from the ones used here; therefore, this study's estimates of expected productivity and risk may not exactly match those used by employers to determine player salaries.

Similarly, some information was not included in estimating a player's expected performance and risk. A player's performance data prior to his major-league career — for instance, his minor-league and amateur statistics — were not included in this study's prediction models. Because each player had accumulated at least six years of major-league performance data at the time his free-agent contracts were signed, it is unlikely that these metrics would have added substantially to the models' accuracy, but they might have had a minor influence. More importantly, information of a qualitative nature — such as scouting reports or medical data — was not incorporated by the prediction models; teams almost certainly use qualitative information in predicting future performance, causing another potential divergence between this study's estimates of expected productivity and MLB employers' actual expectations.

Finally, some contracts include "options," which allow either the team or the player (depending on the type of option) to extend the terms of employment for an additional year at a pre-specified wage when the original contract ends. These options were not included in this study because information on options was not available in all cases, but they may affect a player's wage. Since a "team option" is favorable to the team, at the player's expense (it can be exercised if the player's value increases before the end of the contract, or declined if his value decreases), players must be given higher nominal salaries — the option's "price" — to accept this provision; similarly, to acquire a "player option," workers must accept lower nominal salaries (Clayton & Yermack, 2001). Since these options affect a player's nominal wage, if they are included more or less frequently for risky players, the failure to incorporate them in this study could have biased the results.

The unique qualities of the professional sports labor market, such as the unmatched accessibility of performance data, can often aid the analysis of economic theories, but they also make the external validity of such studies a particular concern. It seems likely that the results of this study are applicable to other professional sports, and the degree of risk-seeking is likely even greater for leagues in which players can be terminated at lower costs, such as the NFL. But many other industries operate differently than professional sports leagues, and in those sectors, the risk preferences of firms may be different. For instance, sports contracts bind a player to one team for the duration of his contract, but many labor agreements in other industries do not prevent a worker from seeking new employment at any time. In these situations, firms are less likely to be risk-seeking, because they can benefit less from over-performing workers; if an employee's productivity is revealed to be greater than expected, he can negotiate a higher salary at a different company. Similarly, if companies cannot easily terminate underperforming employees or relocate them to less important roles — such as moving a disappointing baseball player to the bench so another teammate can play more often — they will also be unlikely to seek riskier workers.

But for employers whose labor conditions are similar to those of Major League Baseball teams — companies in which workers are immobile over a certain length of time and the cost of underperforming employees can be reduced in some way — the findings of this study may be applicable. One topic particularly worthy of future research is whether risky workers actually do have a stronger preference for long-term contracts than consistent workers, as predicted, and if that preference affects their salaries. If it is found that risky workers are willing to take a salary discount in exchange for long-term security, that evidence would support the theory developed in this paper and suggest that risky workers in other industries might also only earn a salary premium in short-term contracts.

APPENDIX A: GARCH model selection

When selecting a generalized autoregressive conditional heteroskedasticity model, it is necessary to choose how many lagged squared residuals and how many lagged predictions of variance will be used for predicting future variance. Several GJR-GARCH(p, q) models were tested, where q is the number of lagged squared residuals and p is the number of lagged variance predictions. The most appropriate model was chosen with the Bayes information criterion (Schwarz, 1978):

$$BIC = -2\ln(L) + k\ln(n)$$

where $\ln(L)$ is the log-likelihood of the model being tested, k is the number of parameters used in that model and n is the number of observations. For both hitters and pitchers, a GJR-GARCH (1,1) model produced the minimum BIC, so one lagged squared residual and one lagged prediction of variance were used to predict future variance.

Table A.1: GARCH model selection for hitters

Model	Observations	Log Likelihood	BIC
GJR-GARCH (1,1)	5946	-10514	**21184.49**
GJR-GARCH (1,2)	5946	-10512	21197.83
GJR-GARCH (1,3)	5946	-10509	21209.94
GJR-GARCH (2,1)	5946	-10514	21192.64

Table A.2: GARCH model selection for pitchers

Model	Observations	Log Likelihood	BIC
GJR-GARCH (1,1)	2296	-4565	**9238.57**
GJR-GARCH (1,2)	2296	-4564	9252.49
GJR-GARCH (1,3)	2296	-4562	9262.87
GJR-GARCH (2,1)	2296	-4564	9243.45

APPENDIX B: Full models of expected performance

Table B shows the full regression models used to predict hitters' and pitchers' expected productivity, which are partly displayed in Tables 3.1 and 3.2.

Table B: Predicting performance for hitters and pitchers

Hitters:	WAR_t	**Pitchers:**	WAR_t
WAR_{t-1}	.2035**	WAR_{t-1}	.1722**
	(.0203)		(.0268)
WAR_{t-2}	.1658**	WAR_{t-2}	.0812**
	(.0155)		(.0215)
WAR_{t-3}	.0881**	WAR_{t-3}	.1045**
	(.0138)		(.0189)
AvgWAR	.2090**	SO_{t-1}	.0129**
	(.0285)		(.0013)
HR_{t-1}	.0277**	BB_{t-1}	-.0125**
	(.0033)		(.0022)
SB_{t-1}	.0061**	HR_{t-1}	-.0165**
	(.0020)		(.0056)
SO_{t-1}	-.0022**	*AGE*	-2.8447**
	(.0009)		(1.0840)
$3B_{t-1}$	.0180*	AGE^2	.0812*
	(.0081)		(.0348)
$GAMES_{t-1}$	-.0012†	AGE^3	-.0008*
	(.0007)		(.0004)
CareerHR	-.0015**		
	(.0004)		
AGE	-.2597**		
	(.0681)		
AGE^2	.0032**		
	(.0011)		
EXP	-.0084		
	(.0134)		
Constant	5.2463**	Constant	33.34**
	(1.0340)		(11.15)

***, *, † indicate significance at the 1, 5 and 10% levels, respectively; standard errors in parentheses.*

REFERENCES

Agency Database. (2013). *MLB Trade Rumors*. Retrieved April 10, 2013, from http://www.mlbtraderumors.com/agencydatabase

Associated Press. (2006). MLB: Forbes analysis "misstates" team finances. *ESPN.com*. Retrieved February 8, 2013, from http://sports.espn.go.com/mlb/news/story?id=2417138

Baseball-Reference.com. (2012). *Sports Reference*. Retrieved December 7, 2012, from http://www.baseball-reference.com/

Berry, A. (2011). Average salary for Major Leaguers on the rise. *MLB.com*. Retrieved April 11, 2013, from http://mlb.mlb.com/news/article.jsp?ymd=20111205&content_id=26096930&vkey=news_mlb&c_id=mlb

Bodvarsson, O. B., & Brastow, R. T. (1998). Do Employers Pay for Consistent Performance? Evidence from the NBA. *Economic Inquiry*, *36*(1), 145–160.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollinger, C. R., & Hotchkiss, J. L. (2003). The upside potential of hiring risky workers: Evidence from the baseball industry. *Journal of Labor Economics*, *21*(4), 923–944.

Bradbury, J. C. (2007). Does the Baseball Labor Market Properly Value Pitchers? *Journal of Sports Economics*, *8*(6), 616–632.

Bradbury, J. C. (2009). Peak athletic performance and ageing: evidence from baseball. *Journal of Sports Sciences*, *27*(6), 599–610.

Bradbury, J. C. (2013). What is Right with Scully-Estimates of a Player's Marginal Revenue Product. *Journal of Sports Economics*, *14*(1), 97–105.

Bradbury, J. C., & Drinen, D. J. (2007). Pigou at the Plate: Externalities in Major League Baseball. *Journal of Sports Economics*, *9*(2), 211–224.

Brown, M. (2011). MLB Revenues Grown From $1.4 Billion in 1995 to $7 Billion in 2010. *The Biz of Baseball*. Retrieved March 5, 2013, from http://www.bizofbaseball.com/index.php?option=com_content&view=article&id=5167:mlb-revenues-grown-from-14-billion-in-1995-to-7-billion-in-2010

Brownson, C. (2010). Does anyone see a pattern here? *The Hardball Times*. Retrieved April 6, 2013, from http://www.hardballtimes.com/main/blog_article/does-anyone-see-a-pattern-here/

Burger, J. D., & Walters, S. J. K. (2008). The existence and persistence of a winner's curse: new evidence from the (baseball) field. *Southern Economic Journal*, *75*(1), 232–245.

Burgess, S., Lane, J., & Stevens, D. (1998). Hiring risky workers: some evidence. *Journal of Economics & Management Strategy*, *7*(4), 669–676.

Cameron, D. (2013). The Recent Examples of a Replacement Level Player. *FanGraphs Baseball*. Retrieved February 20, 2013, from http://www.fangraphs.com/blogs/index.php/the-recent-examples-of-a-replacement-level-player/

Cermeño, R., & Grier, K. B. (2001). Modeling GARCH processes in panel data: Theory, simulations and examples. Working paper. Retrieved from http://faculty-staff.ou.edu/G/Kevin.B.Grier-1/pg01.pdf

Clark, J. B. (1891). Distribution as Determined by a Law of Rent. *The Quarterly Journal of Economics*, *5*(3), 289–318.

Clayton, M., & Yermack, D. (2001). Major League Baseball player contracts: an investigation of the empirical properties of real options. Working Paper. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1297052

Click, J. (2006). Was Billy Martin Crazy? In J. Keri (Ed.), *Baseball Between the Numbers* (pp. 35–47). New York: Basic Books.

Consumer Price Index. (2013).*Bureau of Labor Statistics*. Retrieved March 4, 2013, from ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

Dosh, K. (2007). Can Money Still Buy the Postseason in Major League Baseball? *University of Denver Sports and Entertainment Law Journal*, *3*, 1–44.

DuPaul, G. (2012). What is WAR good for? *The Hardball Times*. Retrieved February 16, 2013, from http://www.hardballtimes.com/main/article/what-is-war-good-for/

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50*(4), 987–1007.

Fair, R. C. (2008). Estimated age effects in baseball. *Journal of Quantitative Analysis in Sports*, *4*(1), 1–39.

Farber, H. S., & Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, *111*(4), 1007–1047.

Forman, S. (2012). Baseball-Reference.com WAR Explained. *Baseball-Reference.com*. Retrieved February 16, 2013, from http://www.baseball-reference.com/about/war_explained.shtml

Fort, R. D. (2006). The value of major league baseball ownership. *International Journal of Sport Finance*, *1*(1), 9–20.

Gassko, D. (2011). *Does the Baseball Free Agent Market Properly Value Performance?* Thesis, Rice University.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801.

Groothuis, P. A., Hill, J. R., & Perri, T. J. (2007). Early Entry in the NBA Draft: The Influence of Unraveling, Human Capital, and Option Value. *Journal of Sports Economics*, *8*(3), 223–243.

Grosnick, B. (2012). Introducing the WAR Index. *Beyond the Box Score*. Retrieved February 20, 2013, from http://www.beyondtheboxscore.com/2012/10/10/3475808/MLB-WAR-index-fWAR-bWAR-rWAR-WARP

Hakes, J. K., & Sauer, R. D. (2006). An Economic Evaluation of the Moneyball Hypothesis. *Journal of Economic Perspectives*, *20*(3), 173–185.

Hakes, J. K., & Turner, C. (2009). Pay, productivity and aging in Major League Baseball. *Journal of Productivity Analysis*, *35*(1), 61–74.

Harris, M., & Holmstrom, B. (1982). A Theory of Wage Dynamics. *The Review of Economic Studies*, *49*(3), 315–333.

Healy, A. (2007). Do Firms Have Short Memories?: Evidence From Major League Baseball. *Journal of Sports Economics, 9*(4), 407–424.

Hendricks, W., DeBrock, L., & Koenker, R. (2003). Uncertainty, hiring and subsequent performance: The NFL draft. *Journal of Labor Economics*, *21*(4), 857–886.

Hicks, J. R. (1963). *Theory of Wages* (2nd ed., p. 388). New York: St. Martin's Press, Inc.

Hill, J. R., & Spellman, W. (1983). Professional Baseball: The Reserve Clause and Salary Structure. *Industrial Relations: A Journal of Economy and Society*, *22*(1), 1–19.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *The Journal of Economic Perspectives*, *14*(3), 75–94.

Krautmann, A. C. (1990). Shirking or Stochastic Productivity in Major League Baseball? *Southern Economic Journal*, *56*(4), 961–968.

Lazear, E. P. (1998). Hiring risky workers. In I. Ohashi & T. Tachibanaki (Eds.), *Internal Labour Markets, Incentives and Employment* (pp. 143–158). New York: St. Martin's Press, Inc.

Lehn, K. (1984). Information asymmetries in baseball's free agent market. *Economic Inquiry*, *22*(1), 37–44.

Lewis, M. (2003). *Moneyball: The art of winning an unfair game* (p. 320). New York: W.W. Norton & Company.

MacDonald, D. N., & Reynolds, M. O. (1994). Are baseball players paid their marginal products? *Managerial and Decision Economics*, *15*(5), 443–457.

Major League Rules. (2008). Accessed via BizofBaseball.com.

Marburger, D. R. (1994). Bargaining power and the structure of salaries in major league baseball. *Managerial and Decision Economics*, *15*(5), 433–441.

Martin, J. A., Eggleston, T. M., Seymour, V. A., & Lecrom, C. W. (2011). One-Hit Wonders: A Study of Contract-Year Performance among Impending Free Agents in Major League Baseball. *NINE: A Journal of Baseball History and Culture*, *20*(1), 11–26.

Maxcy, J. G. (2004). Motivating long-term employment contracts: Risk management in major league baseball. *Managerial and Decision Economics*, *25*(2), 109–120.

Miller, P. A. (2000). A theoretical and empirical comparison of free agent and arbitration-eligible salaries negotiated in major league baseball. *Southern Economic Journal*, *67*(1), 87–104.

Mincer, J. (1974). *Schooling, experience, and earnings* (p. 152). New York: Columbia University Press.

Moore, J. (2011). Can WPA Explain How Teams Buy Relievers? *FanGraphs Baseball*. Retrieved April 6, 2013, from http://www.fangraphs.com/blogs/index.php/can-wpa-explain-how-teams-buy-relievers/

Pascal, A. H., & Rapping, L. A. (1972). The Economics of Racial discrimination in Organized Baseball. In A. H. Pascal (Ed.), *Racial Discrimination in Economic Life* (pp. 119–156). Lexington, Mass: Lexington Books.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, *36*(4), 859–866.

Perry, D. (2006). Do Players Perform Better in Contract Years? In J. Keri (Ed.), *Baseball Between the Numbers* (pp. 199–206). New York: Basic Books.

Rabemananjara, R., & Zakoïan, J.-M. (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, *8*(1), 31–49.

Schulz, R., & Curnow, C. (1988). Peak Performance and Age Among Superathletes: Track and Field, Swimming, Baseball, Tennis, and Golf. *Journal of Gerontology: Psychological Sciences*, *43*(5), 113–120.

Schulz, R., Musa, D., Staszewski, J., & Siegler, R. S. (1994). The relationship between age and major league baseball performance: implications for development. *Psychology and Aging*, *9*(2), 274–86.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461–464.

Scully, G. W. (1974). Pay and performance in major league baseball. *The American Economic Review*, *64*(6), 915–930.

Scully, G. W. (1989). *The Business of Major League Baseball* (p. 212). Chicago: University of Chicago Press.

Silver, N. (2005). Lies, Damned Lies: A Mulligan on Guzman. *Baseball Prospectus*. Retrieved April 8, 2013, from http://www.baseballprospectus.com/article.php?articleid=4535

Silver, N. (2006). Is Alex Rodriguez Overpaid? In J. Keri (Ed.), *Baseball Between the Numbers* (pp. 174–199). New York: Basic Books.

Slater, C. (2012). MLB Advanced Media's Bob Bowman Is Playing Digital Hardball. And He's Winning. *Fast Company*. Retrieved March 5, 2013, from http://www.fastcompany.com/1822802/mlb-advanced-medias-bob-bowman-playing-digital-hardball-and-hes-winning

Slowinski, S. (2011). The Projection Rundown: The Basics on Marcels, ZiPS, CAIRO, Oliver, and the Rest. *FanGraphs Baseball*. Retrieved February 22, 2013, from http://www.fangraphs.com/library/index.php/the-projection-rundown-the-basics-on-marcels-zips-cairo-oliver-and-the-rest/

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Swartz, M. (2012). Positional Differences in the Price of WAR. *FanGraphs Baseball*. Retrieved April 6, 2013, from http://www.fangraphs.com/blogs/index.php/positional-differences-in-the-price-of-war-2/

Tango, T. (2007). Basic Aging Curve for Hitters, 1957-2006. *The Book - Playing the Percentages in Baseball*. Retrieved December 7, 2012, from http://www.insidethebook.com/ee/index.php/site/article/basic_aging_curve_for_hitters_1957_2006/

Tango, T., Lichtman, M., & Dolphin, A. (2007). *The Book: Playing the Percentages in Baseball* (p. 386). Dulles, Virginia: Potomac Books Inc.

Thaler, R. H. (1988). Anomalies: The winner's curse. *The Journal of Economic Perspectives*, *2*(1), 191–202.

Woolner, K. (2006). Why Is Mario Mendoza So Important? In J. Keri (Ed.), *Baseball Between the Numbers* (pp. 157–173). New York: Basic Books.

Zimbalist, A. (1992). Salaries and Performance: Beyond the Scully model. In P. M. Sommers (Ed.), *Diamonds are Forever* (pp. 109–133). Washington, D.C.: The Brookings Institution.

PLEDGE OF HONOR

This paper represents my own work in accordance with University regulations.

Kevin Whitaker

Class of 2013