The Upside Potential of Hiring Risky Workers: Evidence from the Baseball Industry

Christopher R. Bollinger, *University of Kentucky*

Julie L. Hotchkiss, *Georgia State University and the Federal Reserve Bank of Atlanta*

Making use of performance data for baseball players, this article provides empirical evidence in support of Lazear's (1998) theoretical predictions that (1) risky workers will earn a premium for their upside potential, (2) this risk premium will be higher the longer a worker's work life, and (3) firms must enjoy some comparative advantage in the labor market to be willing to pay a premium to risky workers. The validity of Lazear's predictions carries implications for wage differentials between young and old workers and between men and women.

I. Introduction

One of the best-known characteristics of financial markets is that risk carries with it greater potential for higher returns. Edward Lazear (1998)

Research support from the Georgia State University Research Initiation Grant Program is gratefully acknowledged. We thank Dan Black, John Garen, Bruce K. Johnson, Eric Johnson, Robert E. Moore, Ben Scafidi, and Bruce Seaman for comments and suggestions. The assistance of Bill Francis, with the National Baseball Hall of Fame, and of Greg Bouris, with the MLB Player's Association, is also acknowledged. Research assistance was provided by Anslie Fields and Femi Alao. The views expressed here do not necessarily reflect those of the Federal Reserve System.

[*Journal of Labor Economics*, 2003, vol. 21, no. 4]

0734-306X/2003/2104-0007$10.00

claims that this same potential exists in labor markets and that employers pay a premium to hire risky workers. "Risky workers" refers to workers with variable, or uncertain, productivity. Employers value risky workers because the worker in the upper tail of the productivity distribution can be retained and the worker in the lower tail of the productivity distribution can be terminated; risk provides option value to the worker.

Lazear develops a theoretical model in which he shows (among other things) that (1) risky workers (workers with uncertain productivity) will be paid more than safe (certain) workers, (2) the longer the work life a risky worker has the higher will be the risk premium received, and (3) in order for the firm to reap benefits of the upside potential from hiring a risky worker, the firm must have some competitive advantage (e.g., proprietary information or costs to worker mobility) over other firms. The purpose of this article is to use data from the baseball industry to test these predictions of Lazear's theoretical model. Data from the baseball industry are particularly appealing for this application because detailed information about workers' productivity (levels and variability) is available, part of workers' productivity can be argued to be firm specific, predictors for length of work life exist, and the labor market is structured in such a way that firms do enjoy competitive advantage with regard to their new hires. These are all necessary characteristics in order to formulate reasonable empirical tests of Lazear's theories.

As Kahn (2000) points out, the sports industry can provide a rich laboratory in which to explore general labor market issues, and this article presents a new application of sports data: evaluating the role of risk, or variable productivity, in hiring decisions and salary determination. The implications of Lazear's model are important to the labor market as a whole. As Lazear points out, the hiring model provides a nondiscrimination explanation of why we might expect young workers to be paid more than older workers and why we might expect males to be paid more than females, on average, ceteris paribus. His model provides a structure to the profit maximizing arguments for explaining these two outcomes but does not offer any empirical evidence for support of the model. In order to evaluate Lazear's implications for the labor market in general, it is important to find empirical support for his theoretical predictions.

II. Review of Lazear's Results

This section reviews the basic findings of Lazear (1998). We leave to the reader the exploration of the formal derivation of the results; here we only reference Lazear's main propositions and the intuition supporting those propositions.

Lazear posits two types of workers, safe and risky. The safe worker has output M_S while the risky worker has output that is random with mean M_S. Hence, on average, all workers have the same average productivity. Classical theory, with competitive labor markets and fully mobile labor, would imply that all workers in this case should be paid the same. Lazear constructs a market in which the risky worker is paid more than the safe worker in spite of the equality in their average productivity. Certain market conditions are essential for this outcome. First, employers enjoy some advantage (with respect to their workers' productivity) over other firms. This advantage may arise in a number of ways: the firm may have some proprietary information concerning the distribution of its workers' productivity, its workers may be restricted in their mobility, or there may merely be some mechanism through which workers sort into firms with which they have the highest productivity. Second, Lazear's model requires the existence of a probationary period during which the firm learns about the distribution of its workers' productivity. In addition, the firm has the ability to dismiss workers who don't achieve some desired level of productivity. And, third, some aspect of the distribution of workers' productivity is firm specific, meaning that some aspect of a worker's output at one firm is not necessarily related to the worker's output at another firm. These conditions lead to Lazear's first proposition: "Equilibrium starting wage for risky workers is higher than the equilibrium starting wage for safe workers. The premium reflects the option value of hiring a risky worker" (p. 146).

Although this is phrased in terms of "starting wages," Lazear is really working in a two-period model: probationary and nonprobationary. The result, however, extends to multiperiod models.

The intuition behind this result is found in the financial options markets. Since the firm can, at any time, remove a player with low productivity, by retaining a player the firm is exercising a call option. Call options are, of course, priced above the expected or mean value of the asset, in this case the player's marginal productivity. As has been established in financial markets, assets with higher risk will have higher option prices, ceteris paribus. The firm is paying for option value, "a firm is willing to pay a higher wage to workers for whom there is a significant upside" (Lazear, p. 148). The firm's willingness to pay this premium hinges on the firm-specific component which makes the worker more valuable to that firm than to other firms. This implies that any standard measure of the riskiness of the worker (such as variance of performance) should be positively related to the earnings of that worker, just as variance in the past performance of a stock is positively associated with the price of a call option.

Lazear's third proposition states: "The premium that the newly hired risky worker receives over the safe worker increases in the length of the worklife" (p. 149).

The reasoning behind this result is that the longer the firm has to recoup the returns to risk, the more willing it is to pay for the upside potential of the risky worker. Strictly speaking, a worker's career is the length of time a worker is tied to the employer.

It is interesting to note that another prediction of Lazear's model is supported anecdotally in the baseball industry, although not tested here. He argues that even though firms may have some competitive advantage (through the reserve clause structure), they will still find additional productivity information of value. The value of such information essentially enables firms to cut off the lowest tail of the risky worker distribution, improving the probability of the upside gain; the expected value relative to the wage paid is higher. The well-organized minor league baseball system, and subsidization of that league by the major league, can be argued to illustrate how much value firms place (i.e., how much they are willing to spend) on reducing that downside potential of risky workers (see Krautmann et al. 2000).

Each of these predictions about the willingness of the firm to pay up front for a worker's riskiness hinges on the worker being tied to the firm. If the firm had to compete each year for the worker, the value of information about the worker will become public and will be incorporated into the wage.

III. The Test Environment

Testing the predictions of a theoretical model is not always straightforward. It is essential to structure the environment as closely as possible to the one assumed within the theoretical model. If the environment is not replicated and support for the model is not found, the test itself, rather than the theory, is left wide open for criticism. The environment of Lazear's model is fairly demanding. The model predictions depend on the presence of identifiable labor market imperfections (costly mobility, lack of mobility, or imperfect information), a labor market structure in which a probationary period is built in, a firm-specific component to worker productivity, and, perhaps most importantly, access to information about the riskiness of workers. We believe the labor market for baseball players comes as close as any in replicating this environment: the reserve clause structure and multiperiod contracts result in severe restrictions to mobility; teams likely gain proprietary information about their players' productivity through observing players while in the minor league; all players go through some "probationary" period, such as the minor league or college teams; team "synergy," signals and strategy unique to the team, nuances of the home field to which a player becomes accustomed, and "fan loyalty" all contribute to firm-specific productivity; teams have the ability to dismiss players during their probationary period;[1] and produc-

[1] While all contracts in major league baseball are "guaranteed," teams have markedly more flexibility in "dismissing" or reassigning reserve clause players than they do with free agents.

tivity measures of players (means and variance) are readily quantifiable and available.[2] Three predictions of Lazear's theoretical model will be tested: (1) whether there is a premium associated with the riskiness of a worker, (2) whether the risk premium diminishes over a worker's expected working life, and (3) whether imperfections in the labor market influence the determination of the risk premium.

In finding a labor market that conforms to a theoretical stylization, however, one runs the risk that the environment is not representative enough to generalize the results to other labor markets. The two most common criticisms of using baseball data to explore general labor market phenomena are that employers (team owners) may not be maximizing profits and that players are not paid the value of their marginal product (firms possess some market advantage). On the point of profit maximization, Quirk and Fort (1992) point out that if win-percent maximization (as opposed to profit maximization) were an important aspect of the owners' optimization, one should see the highest win percentage associated with the most wealthy owners. However, the 22 wealthiest team owners in 1990 had an average win percentage of .489 for their ownership tenure. Further, if maximizing some utility of win percentage were a dominant feature of the owners' decision-making process, this would imply that players would be paid more than the value of their marginal product, the evidence for which is quite to the contrary (see, e.g., Hill 1985; Bruggink and Rose 1990; and Krautmann 1999).[3] Nonetheless, we include win percent (as well as other indicators of team financial standing) as a regressor in the empirical analysis with the intention of capturing any influence this statistic might have in the determination of players' salaries.

Lazear's theory of labor market hiring requires that firms have some advantage (over other firms) in the labor market. While the baseball industry reflects this requirement nicely, some could argue that it is too restrictive of the labor market in general. However, the degree of market imperfection in Lazear's model is not very stringent. Examples of the type of advantage the employer can have over other firms include some proprietary information about the worker or limits to worker mobility. The firm need not necessarily know something about a worker's pro-

[2] It is the ready availability of productivity measures that has made the use of baseball and other sports data popular with labor economists. See, e.g., Scully (1989), Ehrenberg and Bognanno (1990); Kahn (1991); Blass (1992); Krautmann (1999); and Krautmann et al. (2000). Rosen and Sanderson (2000) provide an overview of and empirical evidence on the nuances of the labor markets in professional sports.

[3] Other studies which have assumed profit maximization by team owners include Fort and Scully (1989), Quirk and Fort (1992), Zimbalist (1992), and Quirk (1995). For an alternative view, see Vrooman (1997).

ductivity that is not known by other firms or by the worker himself. A firm may simply know what features of a worker, for example, personality style, best combine with other input factors at that firm in order to be at an advantage. In addition, other evidence in the labor market, such as firms paying for general human capital enhancements for their workers, suggests that many firms enjoy at least some local labor market advantage.[4]

A. The Data

A comprehensive database of player statistics, salary compensation, team-level statistics, and financial information has been compiled. A complete description of the variables and sources of the data are included in table 1. The data include season and career statistics for all nonpitchers who appear on a 25-man roster in the spring of any season between 1987 and 1993. Data collection ended with 1993 since both 1994 and 1995 were incomplete seasons due to the strike.[5]

The period under study is of particular interest. First and foremost, this period represents a relatively innocuous period with respect to employer-employee relationships. Prior to the late 1980s, the baseball industry underwent substantial change in employment regulation. During the early 1970s, free agency began to be implemented, but by the early 1980s the system currently in place had fully evolved. Finally, there is weak evidence (see Gius and Hylan 1996) that there was collusion in hiring practices across teams during the early and mid-1980s. However, the evidence and court decisions suggest that by 1987 collusion had been eliminated.[6]

A second reason to study this period is that it represents a period of some of the highest growth in salaries and revenues ever seen. Average salary more than doubled during this period (from approximately $450,000 to over $1 million).[7] A similar phenomenon appears for revenues, from average revenue of $30 million in 1986 to average revenue of $68

[4] See, e.g., Kaufman and Hotchkiss (2000, p. 371).

[5] Also, Florida and Colorado were expansion teams in 1993 and were eliminated from the analysis because of the special considerations given to expansion teams in their first year of hiring players.

[6] There is some disagreement as to when the collusion actually ended. However, collusion did not tend to affect reserve clause players but was rather a factor only for free agents and then, it seems, only for the top players. Our data focus primarily on reserve clause players and hence are unlikely to be affected. Further, one would expect that collusion would understate the effects of variance rather than overstate them. Hence our results in favor of this hypothesis are likely to be understatements rather than overstatements.

[7] The figures are based on authors' calculations. Salaries deflated by the consumer price index (CPI) are used in the empirical analysis. Clearly, the rate of inflation is not the source for the huge salary increases over the time period studied here, which is why we include demand-side controls, such as television revenues.

million in 1993. These phenomena are largely attributed to the changing role of television in the industry. Television revenues grew dramatically during this period, especially at the renegotiation period of 1989–90. Most firms experienced a jump in television revenue of over 20% (authors' calculations from *Broadcast Magazine* and *Financial World*). These observable exogenous changes in demand for the product (factors affecting marginal revenue) allow stronger identification of those factors affecting player productivity and wages.

B. The Sample

A player typically begins his journey to the major league as a player on his team's minor league team, often referred to as a "farm team." It is during this time period that the team owners learn about the distribution of players' productivity. Also during this time period (and during their first year in the major league), players are all essentially paid the same salary. This is consistent with the requirement in Lazear's model that all workers get paid the same wage during their probationary period. We expand this definition of probationary period to include the first three years of a player's reserve clause period.[8]

The analysis examines three samples. The primary sample is a pooled cross section of reserve clause players who are observed for at least 2 seasons but for no more than 5 seasons. We call this the reserve clause sample. An important subsample is the reserve clause players who have not been traded at the year of the contract. The not-traded limitation is to insure that the value of a player's riskiness derives from the firm-specific component of that riskiness, as required by Lazear's theory. For example, a player who begins his major league career in Atlanta in 1987 and is then traded for the 1991 season to Montreal will appear in the not-traded subsample only in 1988, 1989, and 1990. The player's 1987 season is not included because not enough information is available at that time to calculate performance, variance, and covariance measures. The 1991 season is excluded from the not-traded subsample because the player changed teams and not enough information is available to calculate measures with the new team. Seasons after 1991 are excluded because the player has become free agent eligible. The number of seasons a player

[8] A player is classified as a reserve clause player when he first arrives in the major league. Reserve clause players are tied to one team by an agreement that does not allow any salary negotiations during the player's first 3 years as a major league player. After 3 years, a player is entitled to salary arbitration with the team holding his contract. With 6 years of service, the player is eligible for open contract negotiations with any major league team when that player's current contract is due to expire. It is important to note that, even though free agents are considered to have a great deal more mobility than reserve clause players, they, too, are often locked into multiple-year contracts.

Table 1
Means and List of Variables included in Empirical Specification; Sample Includes Players with at Least 2 but No More Than 5 Seasons (All Seasons on the Same Team)

Variable	Description	Source	Reserve Clause	Free Agent
No. of players			934	752
Player performance:				
CATBAT	Career at bats	AC	903.8	4,312.5
CBASE	Career on base	AC	317.7	1,618.8
CPLATE	Career at the plate	AC	985.6	4,754.5
CGAMES	Career number of games	AC	288.9	1,257.8
CRUNS_B	Career runs/CBASE	AC	.38	.36
CHITS_A	Career hits/CATBAT	AC	.25	.27
CHR_A	Career home runs/CATBAT	AC	.02	.03
CRBI_A	Career runs batted in/CATBAT	AC	.11	.12
CBB_P	Career bases on balls/CPLATE	AC	.08	.09
CSO_P	Career strikeouts/CPLATE	AC	.16	.14
CSB_B	Career stolen bases/CBASE	AC	.08	.07
CCS_B	Career caught stealing/CBASE	AC	.04	.03
CFR_G	Career fielding runs/CGAMES	AC	−.001	.005
VRUNS	Variance in number of runs scored	AC	303.7	249.2
VHITS	Variance in number of hits	AC	1,120.1	758.7
VHR	Variance in number of home runs	AC	21.1	24.4
VRBI	Variance in number of runs batted in	AC	279.4	252.2
VBB	Variance in number of bases on balls	AC	173.0	158.4
VSO	Variance in number of strikeouts	AC	484.6	263.0
VSB	Variance in number of stolen bases	AC	35.5	35.2
VCS	Variance in caught stealing	AC	6.2	5.5
VFR	Variance in fielding runs variable	AC	31.3	46.5
RUNHIT	Covariance between runs and hits	AC	555.7	382.4
RUNHR	Covariance between runs and home runs	AC	62.4	54.9
RUNRBI	Covariance between runs and rbi	AC	262.6	201.0
RUNBB	Covariance between runs and bases on balls	AC	200.4	142.0
RUNSO	Covariance between runs and strikeouts	AC	329.5	163.8
RUNSB	Covariance between runs and stolen bases	AC	64.5	48.1
RUNCS	Covariance between runs and caught stealing	AC	26.4	15.2
RUNFR	Covariance between runs and fielding runs	AC	−.3	5.1
HITHR	Covariance between hits and home runs	AC	111.0	81.8
HITRBI	Covariance between hits and rbi	AC	510.1	355.0
HITBB	Covariance between hits and bases on balls	AC	370.4	224.8
HITSO	Covariance between hits and strikeouts	AC	635.0	304.7
HITSB	Covariance between hits and stolen bases	AC	119.9	75.0
HITCS	Covariance between hits and caught stealing	AC	52.3	28.4
HITFR	Covariance between hits and fielding runs	AC	2.1	8.0
HRRBI	Covariance between home runs and rbi	AC	66.2	60.6
HRBB	Covariance between home runs and bases on balls	AC	41.0	32.8
HRSO	Covariance between home runs and strikeouts	AC	77.1	43.7
HRSB	Covariance between home runs and stolen bases	AC	9.5	6.0
HRCS	Covariance between home runs and caught stealing	AC	4.2	2.8
HRFR	Covariance between home runs and fielding runs	AC	.3	1.8
RBIBB	Covariance between rbi and bases on balls	AC	179.0	119.8
RBISO	Covariance between rbi and strikeouts	AC	318.0	161.7
RBISB	Covariance between rbi and stolen bases	AC	49.3	29.8
RBICS	Covariance between rbi and caught stealing	AC	22.0	11.4
RBIFR	Covariance between rbi and fielding runs	AC	1.3	.8
BBSO	Covariance between bases on balls and strikeouts	AC	230.5	120.6

Table 1 (*Continued*)

Variable	Description	Source	Reserve Clause	Free Agent
BBSB	Covariance between bases on balls and stolen bases	AC	41.7	26.4
BBCS	Covariance between bases on balls and caught stealing	AC	17.8	10.1
BBFR	Covariance between bases on balls and fielding runs	AC	−.3	−1.0
SOSB	Covariance between strikeouts and stolen bases	AC	67.3	28.1
SOCS	Covariance between strikeouts and caught stealing	AC	30.2	11.7
SOFR	Covariance between strikeouts and fielding runs	AC	2.1	5.9
SBCS	Covariance between stolen bases and caught stealing	AC	10.5	7.5
SBFR	Covariance between stolen bases and fielding runs	AC	1.9	1.8
CSFR	Covariance between caught stealing and fielding runs	AC	.6	.2
SEASONS	Number of seasons played	AC	3.5	11.2
NYSALARY	Next season's salary (in U.S.$)	USASN	330,700	1,003,400
AGE	Age of player	AC	26.1	32.5
Team performance:				
LEAGUE	Indicator for league champ	TB	.09	.09
PCT	Win percentage	TB	.50	.51
Team financial performance:				
STADIUM	Seating in stadium	TB	52,698	52,115
TVRIGHTS	Local TV contract revenues (in $)	BMFW	57,591,000	62,319,000

Note.—All measures calculated over at least 2 seasons, including the current season.

Sources.—TB: *Total Baseball* (Thorn and Palmer 1995). USASN: Various issues of *USA Today* and *Sporting News* magazine, publishes roster salaries every spring. BMFW: Various issues of *Broadcast Magazine* and *Financial World* publications. AC: Authors' calculations, based on statistics obtained from Thorn and Palmer 1995.

appears in the data set is controlled for in the estimation. We might expect that the return to riskiness would be smaller when the traded players are included in the analysis. We examine both samples.

The third sample is the sample of all free-agent-eligible players. These are the players who have been free agents or have accumulated enough service to qualify as free agents. As noted above, Lazear's theory requires immobility of a player. While reserve clause players are not able to bargain with other firms, free-agent-eligible players are able to negotiate contracts with other firms. Hence, we would not expect to observe a risk premium for free-agent-eligible players.

Table 1 contains sample statistics for the reserve clause sample and the free-agent-eligible sample used in the estimation. The statistics for the subsamples are comparable. One item of note is that most of the offensive performance measures have a positive covariance, and each of these has a negative covariance with the only defensive measure of fielding runs.

IV. The Tests and Results

This section develops the specific tests for each of the three predications of Lazear's (1998) theoretical model and presents the empirical evidence for each test. Tables 2 and 3 contain the estimation results.

A. The Test for a Risk Premium

The first test conducted is to see whether variance in a player's performance contributes positively to his salary determination. Wages in Lazear's model are determined by a player's marginal product (M) and a firm-specific component (S_i). A player's marginal product has some mean M_S, while the mean of S_i is zero. Following standard labor theory, where MR is the firm's marginal revenue, consider a player's marginal revenue product: $MRP = MR \times \exp(M + S_i)$. Taking logs and letting this expression equal the wage paid to the player yields

$$\ln \text{Wage} = \ln MR + M + S_i. \tag{1}$$

Lazear's theory suggests that the variance also plays a role and uncertainty leads to the level of $(M + S_i)$ being replaced by the mean. This implies the following specification:

$$\ln \text{Wage} = \ln MR + E(M + S_i) + V(M + S_i). \tag{2}$$

Note that we keep the marginal revenue term as a part of the specification, since these variables are readily available (see below) and may be correlated with either S_i or the variance of S_i.

Following standard literature, we model the mean of $M + S_i$ (the mean of performance) as $E(M + S_i) = X\beta$, a linear model. Thus the variance of $M + S_i$ is measured as the variance of this linear term. This leads to our main specification:

$$\ln \text{Wage} = X\beta + \alpha VC(X\beta) + D\gamma. \tag{3}$$

The $VC(X\beta)$ is a quadratic form in the variance and covariance terms of the performance measures and the slope coefficients of those performance measures from the mean regression. The variable D measures the marginal revenue component, such as television revenues, stadium size, and other terms not specific to the player but affecting team marginal revenue. This

Table 2
Nonlinear Least-Squares Estimation Results of Primary Model

	Players with at Least 2 Seasons of Major League Experience		
Variables	Reserve Clause (1)	Nontraded Reserve Clause (2)	Free Agent Eligible (3)
Intercept	.444	−2.097	2.633
	(2.096)	(2.643)	(2.411)
Cruns_B	.336	.186	1.215
	(.265)	(.304)	(.957)
Chits_A	9.785***	8.636***	11.297***
	(.906)	(1.063)	(1.840)
CHR_A	5.453**	3.822	12.948**
	(2.242)	(2.538)	(5.787)
CRBI_A	2.016*	3.062**	6.301**
	(1.092)	(1.256)	(2.667)
CBB_P	2.227***	1.341*	2.006**
	(.652)	(.777)	(.999)
CSO_P	−.990**	−.973*	−2.258**
	(.441)	(.525)	(.927)
CSB_B	.594*	.809**	2.550***
	(.347)	(.398)	(.792)
CCS_B	−1.115	−.742	−.330
	(.897)	(1.042)	(2.630)
CFR_G	1.305***	1.316***	2.394***
	(.255)	(.308)	(.518)
STADIUM	2.85×10^{-6}	6.20×10^{-7}	-4.36×10^{-6}
	(2.2×10^{-6})	(2.75×10^{-6})	(3.2×10^{-6})
Ln(TVRIGHTs)	.009	.008	.016*
	(.007)	(.008)	(.008)
LEAGUE	.175**	.145	.039
	(.074)	(.092)	(.102)
PCT	.549	.742*	−.132
	(.352)	(.431)	(.480)
SEASONS	.436***	.475***	.010
	(.020)	(.025)	(.017)
Age	.206	.420**	.153
	(.158)	(.204)	(.141)
Age^2	−.005	−.009**	−.003
	(.003)	(.004)	(.002)
VCTOT	4.00×10^{-7}***	5.64×10^{-7}**	1.77×10^{-7}
	(1.55×10^{-7})	(2.35×10^{-7})	(1.32×10^{-7})
Adjusted R^2	.562	.593	.372
Number of observations	934	626	752

Note.—The dependent variable is next year's log annual salary deflated by the CPI; regressors are measured through the current year of observation. Standard errors are in parentheses.
* Significant at the 90% confidence level.
** Significant at the 95% confidence level.
*** Significant at the 99% confidence level; two-tailed tests.

Table 3
Nonlinear Least-Squares Estimation Results—Interactions

Variables	Players with at Least 2 Seasons of Major League Experience: Reserve Clause (1)	Nontraded Reserve Clause (2)
Intercept	1.145	−1.813
	(2.080)	(2.617)
CRUNS_B	.300	.165
	(.263)	(.301)
CHITS_A	9.734***	8.790***
	(.889)	(1.044)
CHR_A	6.061***	4.375*
	(2.221)	(2.517)
CRBI_A	1.967*	2.936**
	(1.082)	(1.240)
CBB_P	2.148***	1.332*
	(.645)	(.769)
CSO_P	−1.053**	−.963*
	(.433)	(.516)
CSB_B	.604*	.767*
	(.343)	(.393)
CCS_B	−1.079	−.683
	(.890)	(1.033)
CFR_G	1.283***	1.304***
	(.252)	(.306)
STADIUM	2.50×10^{-6}	4.23×10^{-7}
	(2.2×10^{-6})	(2.72×10^{-6})
Ln(TVRIGHTS)	.010	.010
	(.006)	(.008)
LEAGUE	.171**	.131
	(.073)	(.091)
PCT	.521	.672*
	(.349)	(.427)
SEASONS	.371***	.410***
	(.025)	(.030)
AGE	.172	.415**
	(.157)	(.202)
AGE^2	−.004	−.009**
	(.003)	(.004)
VCTOT	-1.2×10^{-6}***	-1.2×10^{-6}**
	(4.0×10^{-6})	(5.23×10^{-7})
VCTOT × SEASONS	4.48×10^{-7}***	5.17×10^{-7}***
	(1.2×10^{-7})	(1.7×10^{-7})
Adjusted R^2	.571	.603
Number of observations	934	626

Note.—The dependent variable is next year's log annual salary deflated by the CPI; regressors are measured through the current year of observation. Standard errors are in parentheses.
* Significant at the 90% confidence level.
** Significant at the 95% confidence level.
*** Significant at the 99% confidence level; two-tailed tests.

leads to the nonlinear least-squares specification we estimate. Specifically (refer to table 1 for definitions of the mnemonics):

$$\begin{aligned}\ln W = {} & \beta_0 + \beta_1 \mathrm{CRUNS\_B} + \beta_2 \mathrm{CHITS\_A} + \beta_3 \mathrm{CHR\_A} \\ & + \beta_4 \mathrm{CRBI\_A} + \beta_5 \mathrm{CBB\_P} + \beta_6 \mathrm{CSO\_P} \\ & + \beta_7 \mathrm{CSB\_B} + \beta_8 \mathrm{CCS\_B} + \beta_9 \mathrm{CFR\_G} \\ & + \alpha\left[\sum_{i=1}^{9} \beta_i^2 \operatorname{Var}(X_i) + 2\sum_{i=1}^{9}\sum_{j\neq 1} \beta_i\beta_j \operatorname{Cov}(X_i,X_j)\right] \\ & + \gamma_1 \ln \mathrm{TV} + \gamma_2 \mathrm{STADIUM} + \gamma_3 \mathrm{LEAGUE} \\ & + \gamma_4 \mathrm{PCT} + \delta_1 \mathrm{AGE} + \delta_2 \mathrm{AGE}^2 + \varepsilon,\end{aligned} \tag{4}$$

where X refers to the different career performance-level measures (*CRUNS_B, CHITS_A, . . . , CFR_G*); each of the career-level statistics is normalized by either number of times on base (*_B*), number of times at bat (*_A*), or number of times at the plate (*_P*), whichever is most appropriate for the statistic. Most of the performance variables are commonly used terms and need no lengthy explanation. The fielding runs variable is constructed by Thorn and Palmer (1995). The fielding runs variable is a position-specific measure of the runs saved by the individual player beyond what an average player in that position would save. The statistic is based on putouts, assists, errors, and double plays. More detail can be found in Thorn and Palmer (1995).[9] Career statistics are used instead of current season statistics in order to capture a more accurate picture of a player's average productivity (M_S in Lazear's model), rather than, potentially, a productivity outlier.[10] For example, for hits as a measure of performance in year t, $\mathrm{CHITS\_A}_t = \sum_{j=FrstYr}^{t} \mathrm{HITS}_j / \sum_{j=FrstYr}^{t} \mathrm{ATBATS}_j$. The above specification leads to nonlinear least-squares estimation of the parameters, where β_0–β_9 measure the return to performance levels, α measures the return to variance in performance level, γ_1–γ_4 measure the demand effects

[9] For some idea about how this variable differs across players, a particularly notable shortstop is Ozzie Guillen, with a fielding runs (FR) average of 21, whereas Jeff Blauser, considered a defensively weak infielder (also a shortstop), has an average FR of -13.

[10] Inclusion of current season statistics did not alter the primary conclusions of the analysis but was rejected in favor of the more theoretically justified specification. In addition, Lazear's model implicitly assumes, as do we, that a player's past performance and variance are good predictors of current performance and variance. Indeed, a risky player's value derives from his expected potential (upside) deviation from the mean. The use of career averages in wage equations for baseball players has been discussed by Gustafson and Hadley (1995). Our specification is similar.

on the wage, and δ_1 and δ_2 are designed to capture any additional human capital contributors that are correlated with age.

Different authors have used different performance statistics. Some authors (e.g., Gustafson and Hadley 1995; Gius and Hylan 1996) use the same set of performance statistics as used here. Others (e.g., Krautman, Gustafson, and Hadley 2000; Poppo and Weigelt 2000) have used subsets of these variables attempting to capture "runs produced" as the measure of productivity. To summarize the performance of a player as simply runs produced appears to ignore the entertainment value of how those runs are produced. It is also difficult to determine how to split runs produced by a combined effort of two players. We argue that this is an empirical question and have attempted to include a broad spectrum of performance measures.[11]

Table 2 contains the estimation results of this model for the two reserve clause samples and the free-agent-eligible sample. For clarity, we have grouped the variance and covariance terms together and labeled this VCTOT. Hence, for example, in the second column of numbers in table 2, the coefficient on VCTOT(5.64×10^{-7}) is the estimate for α. Lazear's theory predicts this should be positive for the two reserve clause samples. Since the free-agent-eligible market does not necessarily meet the requirements for the theory, the prediction for this sample is ambiguous. Another estimation approach would be to include all variance and covariance terms as right-hand-side variables and estimate separate coefficients for each. The prediction from Lazear's model would then be that the coefficients on the variance terms must be positive (for the reserve clause sample), while the coefficients on the covariance terms should have the same sign as the product of the corresponding coefficients from the performance levels.

The estimation results in columns 1–3 of table 2 allow us to test whether a premium exists for players who exhibit greater variation in their performance (i.e., riskier players) holding performance level constant.[12] As would be expected, players who exhibit higher normalized performance outcomes, such as more hits per times at bat and more stolen bases per times on base, are paid higher salaries. Those performance measure outcomes that should negatively affect wage, such as more strikeouts, do so. The only defensive measure of performance, career fielding runs per games played, also contributes to salary as expected: higher fielding runs leads

[11] As suggested by a referee, numerous estimations were performed with varying combinations of performance measures, as well as minimizing the number of measures through all-encompassing statistics, such as total player rating. The conclusions regarding the presence of a risk premium were unchanged in all these specifications.

[12] Estimates on subsamples broken down by offensive and defensive positions are consistent with the results presented here and do not differ across subsamples.

to higher salary. The only team statistic that is of significance is whether the team was a league champion. In addition, players who have been around for more seasons earn higher salaries, in part reflecting the presence of a league minimum salary that goes to all first-season players.[13]

The coefficient of particular interest for testing the first prediction of Lazear's (1998) theory is that on VCTOT. This coefficient is positive and significantly different from zero for both of the reserve clause samples, as Lazear's theory would predict if teams associate an option value with high variance in performance. As one might expect, the coefficient for the sample including traded players is lower (although not significantly different) than the coefficient for the nontraded reserve clause sample. The coefficient for the free agents is still positive, as the theory predicted, but not significantly different than zero. Columns 1 and 2 of table 2 present evidence in favor of Lazear's theory. The third column is as expected by the theory, an ambiguous result, but neither confirms nor denies the theory.[14]

Since it is difficult to interpret the value of the coefficient α, we calculate the predicted salary given the estimation results and sample averages, then recalculate the predicted salary at variance levels that are one standard deviation above the mean variance. The result of this calculation is that reserve clause players whose variance in performance is one standard deviation above the mean variance (holding all other variables constant, including performance level) earn a salary that is 7% higher than the player with average variance levels.[15] As expected, the nontraded reserve clause players earn an 8.6% higher salary for variance in performance that is one standard deviation above the mean.

One possible concern is that the VCTOT term is not really measuring variance per se but rather other nonlinear relationships in the mean. We forward two arguments for why this is not the case. First, when the main model is estimated without the VCTOT term, both the order of magnitude and sign of the performance variables are essentially unchanged.[16] When a quadratic term is omitted from a regression, the coefficients on the remaining linear variables are substantially biased. Hence, inclusion of the

[13] Also see Krautmann et al. (2000) about the structure of salary growth during players' reserve clause period.

[14] An additional reason, suggested by a referee, why we would likely not see a risk premium among free agents is that the premium may be fully captured by signing and incentive bonuses.

[15] Increasing players' variance level by one standard deviation translates into a 151,740-unit increase in VCTOT. Players with the highest average values of predicted VCTOT include Fred McGriff (528,695), Albert Belle (633,072), and Mark McGwire (725,484); recall that the data are from pre-home-run-record years. Other average VCTOT values of interest include Don Mattingly (8,727), Andy Vanslyke (41,532), and Barry Bonds (85,295).

[16] See table A1 in the appendix.

VCTOT term would substantially alter the coefficients in the main model if this were simply a case of omitted variable bias. Since it does not, there is no evidence to support the possibility that the initial model is incorrectly specified. However, the significant coefficient on VCTOT in both the reserve clause and nontraded reserve clause samples implies that it adds explanatory power to the model and supports Lazear's theory.

Second, an alternative approach is to estimate this model as a linear model using ordinary least squares with separate coefficients on each of the 45 variance and covariance terms in equation (4). We do not present these results but will summarize them here. There is no prediction for the sign of any of these coefficients if the model is simply quadratic. Lazear's theory, however, has clear predictions for the coefficients on the variance terms: all should be positive. Our findings with this model match the Lazear predictions: all of the coefficients on the variance terms are positive, and eight of them are significant. This alone is striking. If it were simply a case of a quadratic, we would not expect this alignment. These positive coefficients, in a quadratic model, would imply that all of the responses are convex. This seems unlikely. It would imply, for example, that the negative return to strikeouts has a minimum—a point above which strikeouts begin to add to a player's income. Even if this minimum is outside the range of the data, it would be surprising if the negative impact of strikeouts were mitigated as the number of strikeouts increased over any range. The results for the covariance terms are less clear. While Lazear's theory is silent regarding these terms, our operationalization of the theory predicts the coefficients on the covariance terms should have the same sign as the product of the coefficients on the corresponding level terms. We find that the signs on the covariance terms are correct for 16 of the 36 terms. Of the terms that have signs opposite what the variance model would predict, only four are statistically significant. Of the cases where the signs do match, six are statistically significant. Hence, the variance terms all exactly match the prediction of the model, and these are clearly the most important terms in the model. And a preponderance of the covariance terms also match the prediction from our operationalization of the model. It should be noted that failure of the covariance terms would not refute Lazear's theory, only our operationalization of it.

An additional concern might be that since teams can get rid of lower-performing players, a positive correlation between variance and improvement in performance develops in the data. If this is the case, then VCTOT may merely be proxying for player improvement rather than actual uncertainty in performance. To eliminate this concern, we estimated three specifications in which measures of player improvement over the time period were added to the model as regressors. The three improvement measures explored were (1) the difference between the first year performance statistics and the current year, (2) the difference between the current year performance and the career performance, and (3) the current year performance (also a measure of deviation from career statistics). In all of these specifications, the coefficient corresponding to VCTOT remained virtually unchanged from that reported in column 2 of table 2, indicating that VCTOT is not merely proxying for improvement in player performance.

B. The Test for Increasing Risk Premium for Workers with Longer Expected Careers

In order to see whether the premium on risk is higher for players with longer expected careers, we estimate the model detailed above with an additional interaction term.[17] Since the term of interest is only significant in the reserve clause player samples, we limit further analysis to these samples. We interact the riskiness measure with the number of seasons a player has been in the major league (SEASONS). As is well known, many players fail to make it through the initial reserve clause period. Those who do, however, are typically in the game for many years to come (particularly if the player has remained with the same team over this time period). Hence, we argue that the best indicator of expected length of service is current length of service. In other words,

$$\frac{\partial E[\text{careerlength}|\text{SEASONS} = t]}{\partial \text{SEASONS}} > 0. \tag{5}$$

Therefore, Lazear's model would predict that the coefficient on the interaction term will be positive.

Columns 1 and 2 of table 3 present the results appropriate for testing the prediction that the longer a worker's expected career, the higher the premium for risk. The argument, again, is that firms will be willing to pay more for a risky worker's option value if they have longer to reap the rewards of the investment. Since the expected career of players increases the longer one observes them on a major league roster, the premium for risk should increase with the number of seasons a player has been in the major league (remembering that the samples contain players with at most 5 seasons of play). We would expect the result to be most pronounced in the sample of nontraded reserve clause players, since this constitutes the group with the largest mobility restrictions. The coefficient on the interaction term (VCTOT × SEASONS) is positive and significantly different from zero in both columns, meaning that the return to overall variance is higher the larger SEASONS is and the longer the

[17] Recall that "career" in Lazear's model refers to the length of time the player is tied to his team. Further analysis of this prediction with information on contract length would be of interest.

player's expected career is. Also note that the partial derivative for VCTOT (at the mean value for SEASONS) remains relatively unchanged at 6.10×10^{-7} for the nontraded reserve clause players. Similarly, for the whole reserve clause players, the derivative for VCTOT evaluated at the mean is 3.68×10^{-7}. It is natural to ask at what value of SEASONS the return to variance becomes negative. For the nontraded sample, SEASONS would have to be less than 2.3 years. Similarly for all reserve clause players, SEASONS would need to be less than 2.7. Hence, only the players in the first 2 years of playing the major leagues have a zero or negative risk premium. This is consistent with the theory since insufficient data may prevent the firm from determining the variance of performance for at least the first season. The fact that the risk premium is more pronounced as more data become available is certainly consistent with theory.

V. Conclusions

This article presents empirical support for a theory developed by Lazear (1998) that predicts that workers who are risky (i.e., performance is uncertain) will be paid more than workers who have thse same average performance yet are more predictable in their output. The reason for this is that firms are willing to pay for the potential upside of the variable performance of the risky worker under conditions that allow them to recoup the benefits of the better performance. The empirical tests make use of baseball industry data where measures of performance (level and variance) are readily available.

We find that an increase in the variance of a player's performance measures by one standard deviation, ceteris paribus, will raise his salary by 7%. Players whose expected careers are longer are also found to earn a higher premium for their riskiness. We also confirm empirically that firms must enjoy some degree of market power in order to be willing to pay a premium for risk.

Establishing validity of Lazear's theory means that the implications of his model should be taken seriously in offering potential explanations for observed labor market phenomena. Lazear discusses two phenomena that his theory could help to explain: that younger workers are paid more than older workers and that men are paid more, on average, than women. The higher payment to young workers can be explained in the context of Lazear's model in two ways. First, the performance of younger workers will typically be more uncertain than that of older workers—leading to a risk premium. Second, young workers have a longer work life—leading to a higher risk premium for young workers than for similarly risky older workers.

In addition, Lazear cites evidence that female workers are more certain (less risky) than male workers. For example, males have higher variance on normalized (same mean across males and females) IQ tests; and while males and females typically perform at similar achievement levels, the performance of males is more variable.[18] According to Lazear's model, at least part of the wage differential between males and females is a result of the risk premium being paid to the riskier male hires.

Clearly, finding support for the predictions of Lazear's theoretical model does not prove the implications suggested by the results. Future work might expand this analysis to other industries in which there might be some quantitative measure of output, examples of which are hard to find.

[18] References on these observations made by Lazear include Doolittle and Welch (1989) and Han and Hoover (1994). Also see Graham et al. (2000) for evidence that college grade point average (GPA) has a higher variance among a sample of men than women, even when both samples have the same mean level GPA.

Appendix

Table A1
Model Estimates without VCTOT Term

Variables	Players with at Least 2 Seasons of Major League Experience: Reserve Clause	Nontraded Reserve Clause	Free Agent Eligible
Intercept	.791	−1.766	2.292
	(2.103)	(2.662)	(2.405)
CRUNS_B	.392	.229	1.225
	(.268)	(.309)	(.963)
CHITS_A	10.524***	9.367***	11.771***
	(.896)	(1.073)	(1.850)
CHR_A	5.562	3.655	12.708**
	(2.336)	(2.684)	(5.955)
CRBI_A	2.324**	3.622***	6.601**
	(1.141)	(1.330)	(2.713)
CBB_P	2.516***	1.655**	2.040**
	(.681)	(.825)	(1.021)
CSO_P	−1.028**	−.973*	−2.045**
	(.456)	(.553)	(.947)
CSB_B	.557	.782*	2.562***
	(.352)	(.408)	(.799)
CCS_B	−.950	−.467	−.116
	(.908)	(1.058)	(2.646)
CFR_G	1.383***	1.382***	2.351***
	(.258)	(.314)	(.520)
STADIUM	2.69×10^{-6}	3.27×10^{-6}	-4.06×10^{-6}
	(2.2×10^{-6})	(2.77×10^{-6})	(3.2×10^{-6})
Ln(TVRIGHTS)	.009	.007	.017**
	(.0065)	(.008)	(.009)
LEAGUE	.171**	.151	.054
	(.074)	(.093)	(.102)
PCT	.489	.656	−.206
	(.354)	(.433)	(.478)
SEASONS	.445***	.489***	.010
	(.020)	(.025)	(.017)
AGE	.172	.390**	.165
	(.159)	(.205)	(.141)
AGE^2	−.004	−.008**	−.003
	(.003)	(.004)	(.002)
Adjusted R^2	.557	.593	.371
Number of observations	934	626	752

Note.—The dependent variable is next year's log annual salary deflated by the CPI, regressors are measured through the current year of observation. Standard errors are in parentheses.
* Significant at the 90% confidence level.
** Significant at the 95% confidence level.
*** Significant at the 99% confidence level; two-tailed tests.

References

Blass, Asher A. "Does the Baseball Labor Market Contradict the Human Capital Model of Investment?" *Review of Economics and Statistics* 74 (May 1992): 261–68.

Bruggink, Thomas, and Rose, David Jr. "Financial Restraint in the Free

Agent Labor Market for Major League Baseball: Players Look at Strike Three." *Southern Economic Journal* 65 (April 1990): 1029–43.
Doolittle, A., and Welch, C. "Gender Differences in Performance on a College-Level Achievement Test." ACT research report series 1989-03-00. Working paper presented at the annual meeting of the American Educational Research Association, San Francisco, March 27-31, 1989.
Ehrenberg, Ronald G., and Bognanno, Michael L. "Do Tournaments Have Incentive Effects?" *Journal of Political Economy* 98 (December 1990): 1307–24.
Fort, Rodney, and Quirk, James. "Cross-Subsidization, Incentives and outcomes in Professional Team Sports Leagues." *Journal of Economic Literature* 33 (September 1995): 1265–99.
Gius, Mark Paul, and Hylan, Timothy R. "An Interperiod Analysis of the Salary Impact of Structural Changes in Major League Baseball: Evidence from Panel Data." In *Baseball Economics: Current Research*, edited by John Fizel, Lawrence Hadley, and Elizabeth Gustafson. Westport, CT: Greenwood Praeger, 1996.
Graham, Mary; Hotchkiss, Julie L.; and Gerhart, Barry. "Discrimination by Parts: A Fixed-Effects Analysis of Starting Pay Differentials across Gender." *Eastern Economic Journal* (Winter 2000): 9–28.
Gustafson, Elizabeth, and Hadley, Lawrence. "Arbitration and Salary Gaps in Major League Baseball." *Quarterly Journal of Business and Economics* 34 (Summer 1995): 32–46.
Han, Lei, and Hoover, H. D. "Gender Differences in Achievement Test Scores." Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, April 5–7, 1994.
Hill, James R. "The Threat of Free Agency and Exploitation in Professional Baseball: 1976–1979." *Quarterly Review of Economics and Business* 25 (Winter 1985): 68–82.
Kahn, Lawrence M. "Discrimination in Profession Sports: A Survey of the Literature." *Industrial and Labor Relations Review* 44 (April 1991): 395–418.
———. "The Sports Industry as a Labor Market Laboratory." *Journal of Economic Perspectives* 14 (Summer 2000): 75–94.
Kaufman, Bruce E., and Hotchkiss, Julie L. *The Economics of Labor Markets*. New York: Dryden, 2000.
Krautmann, Anthony. "What's Wrong with Scully—Estimates of a Player's Marginal Revenue Product." *Economic Inquiry* 37 (April 1999): 369–81.
Krautmann, Anthony; Gustafson, Elizabeth; and Hadley, Lawrence. "Who Pays for Minor League Training Costs?" *Contemporary Economic Issues* 18 (January 2000): 37–47.
Lazear, Edward P. "Hiring Risky Workers." In *Internal Labour Market, Incentives, and Employment*, edited by Isao Ohashi and Toshiaki Tachibanaki. New York: St. Martin's, 1998.
Poppo, Laura, and Weigelt, Keith. "A Test of the Resource-Based Model Using Baseball Free Agents." *Journal of Economics and Management Strategy* 9 (Winter 2000): 585–614.

Quirk, James, and Fort, Rodney. *Pay Dirt: The Business of Professional Sports Teams*. Princeton, NJ: Princeton University Press, 1992.

Rosen, Sherwin, and Sanderson, Allen. "Labor Markets in Professional Sports." Working Paper no. 7573. Cambridge, MA: National Bureau of Economic Research, 2000.

Scully, Gerald. *The Business of Baseball*. Chicago: University of Chicago Press, 1989.

Thorn, John, and Palmer, Peter, eds., with Michael Gershman. *Total Baseball*. 4th ed. New York: Viking, 1995.

Vrooman, John. "A Unified Theory of Capital and Labor Markets in Major League Baseball." *Southern Economic Journal* 63 (January 1997): 594–619.

Zimbalist, Andrew. *Baseball and Billions.* New York: Basic, 1992.