Testing the Determinants of Income Distribution: An Application of Professional Sports Data*

R. Todd Jewell
University of North Texas
Department of Economics

Michael A. McPherson
University of North Texas
Department of Economics

David J. Molina
University of North Texas
Department of Economics

August 2001

I. Introduction

In the nearly thirty years since Amartya Sen (1973) wrote in his seminal book on income distribution that "the idea of inequality is both very simple and very complex," the study of income distribution has increasingly attracted the attention of economists. The measurement used, whether the measurement is based on a specific distribution or is distribution-free, the properties of the measurement, the welfare implications of the measurement, and the relation of it to economic growth have been just a few of the issues analyzed.[1] However, data constraints have limited the ability of researchers to fully explore income distribution methodologies. For example, in general researchers must rely on sample data since the incomes of individuals in the population are unknown. Also, these studies group the data into categories, and typically the upper bound of the distribution (i.e., the income of the richest individual) is unknown. Furthermore, because no information is generally available at the individual level, researchers cannot control for individual-specific economic and demographic factors. We consider the distribution of income in populations for which income information on each individual is usually known: professional athletes. Professional sports present a myriad of opportunities for income distribution studies because the data seldom suffer from the weaknesses noted above. Analyses based on such data allow income distribution methodologies to be developed and tested that would have been impossible using more-traditional data.

In addition to methodological issues, income distribution issues have policy and business implications to the teams and fans of professional sports. For instance, in 1998 Commissioner Bud Selig formed a Blue Ribbon Panel with the purpose of describing and explaining the economic condition of Major League Baseball (MLB). Comprised of such dignitaries as Paul

[1] For an example of discussion on the different income distribution issues see *Handbook of Income Inequality Measurement*, edited by Jacques Silber, 1999.

Volker, Senator George Mitchell, and columnist George Will, this panel recently made its report (Levin, 2000). The report points out that team payrolls have become increasingly disparate; the gap between "rich" and "poor" teams is not only wide, but it is growing. Case in point is the fact that the salary of the highest-paid player in the 2000 season (Los Angeles' Kevin Brown at $15.7 million) is 95% of the entire payroll of the "poorest" team, the Minnesota Twins. The effect, according to the Panel's report, is a dramatic decline in parity and competitiveness of MLB: Since 1994, a team in the top payroll quartile has won every World Series game. In 1999, the teams with the 5 largest payrolls had an average winning percentage of 0.557, while the 5 poorest teams had a comparable figure of 0.444. The report discusses various recommendations that may narrow this gap, leading to what one might call "convergence" in team payrolls.

This paper will examine the distribution of income in MLB for the years between 1985-2000. We will *revive* (and extend) a technique first suggested by Thurow (1970) and show how it can be used to test for time trends vs. cross-sectional demographic and economic aspects influencing income distribution. Thurow (1970) discusses the impact of various measures on median income. In the present case, by calculating the Gini coefficient, we can discuss the impact that the explanatory variables have on income inequality. In the context of MLB, we attempt to shed light on the reasons why teams have greater or lesser payroll disparity. We leave the question of the effect of between-team payroll inequality on team wins and team revenues for future research.

We will compare a distribution-free Gini coefficient for each team and year with a Gini calculated from a distribution selected from the Pearson family of distributions by means of an *a priori* selection test. We then discuss the implications of both sets of Gini coefficients, including possible reasons why within-team payroll inequality has changed over time. In addition, we investigate the impact on the skewness of the chosen distribution on the variables that affect inequality. Finally, we employ a random effects model to ascertain what time-series and cross-sectional factors affect the Gini coefficients and the parameters of the pre-selected distribution. The remainder of this paper is organized as follows. In the next section, we discuss related literature. In section III, we present the methodology. In section IV, we present data and results. The final section offers concluding remarks and discusses how our results might be of broader interest to researchers concerned with income distribution topics.

II. Related Literature

The issue of whether to use an inequality measure that is based on a particular distribution or that is distribution-free has been a topic widely discussed in the income distribution literature (Silber, 1999). Ryu and Slottje (1999), in advocating the use of parametric distributions, point to several benefits of estimating the Lorenz curve in this manner. Among the benefits Ryu and Slottje (1999) point to are: the ability to summarize thousand of observation points with a few parameters; the ability to estimate the density function at any point; an increased ability to construct inequality measures; and the ability to formulate possible "laws" that would otherwise not be possible to detect.[2] However, there are problems associated with the use of a parametric Lorenz curve, such as the choice of a suitable distribution. The problem is compounded by the fact that the data are normally grouped and have an open-ended highest income category making it difficult to obtain accurate estimates. One advantage of using professional sports data is the fact that the salary is available per individual and not grouped; consequently, an actual number is available for the highest income.

[2] For instance, in this paper we can determine the different impact that time and cross-sectional information has on inequality in a manner that is not possible using the non-parametric approach.

This study analyzes payroll inequality in a professional sport by team. That is, for each year we calculate a measure of inequality for each MLB team and consider the characteristic differences in within-team payroll distributions. Our study is not, however, the first to consider income inequality within professional team sports. Depken (2000) finds that MLB teams with greater wage disparity have fewer wins, but Jewell and Molina (2001) find that this correlation is only significant after the MLB work stoppage of 1994. Furthermore, Sommers (1998) finds a negative relationship between team success and within-team payroll inequality using National Hockey League data. However, using data from the National Basketball League, Berri (2001) finds that an increase in payroll inequality within a team actually leads to an increase in wins.[3] Although the direction and magnitude of the effect of payroll inequality on team performance are still debated, it is clear that it is important to determine and analyze the factors that influence payroll inequality within professional sports teams.

In the context of the present study, the large "Kuznets curve" literature that has emerged in development economics is of particular interest. Kuznets (1955) finds evidence of an inverted U-shaped relationship between the level of development (usually measured by per capita income) and various measures of inequality. That is, Kuznets and others find that income inequality rises with per capita income at low-income levels and then falls as countries move from middle-income to high-income status. Various explanations for this relationship have been given, mostly involving the nature of structural change in the economy.[4] In a recent review of the

[3] Related studies include Fort (1992) and Jewell et al. (2001), who measure the overall Gini coefficients for the entire population of a sport's athletes in a given year ; these studies, however, do not calculate or analyze team Gini coefficients. MacDonald G. M. (1988) tangentially discusses the impact of talent on the distribution of earnings. Rottenberg (2000) discusses the distribution of income in the production allocation sense. In another use of the Gini coefficient in sports economics, Schmidt (2001) and Schmidt and Berri (2001) study the impact of competitive-balance inequality in MLB.

[4] In the development literature, the existence of a Kuznets relationship has not been settled. Paukert (1973), Papenek and Kyn (1986), Tsakloglou (1988), List and Gallet (1999), and Burger (2001) are examples of studies that have found empirical and theoretical support for the U-shaped Kuznets curve. Saith (1983) and Ram (1991) are among those who have found evidence against this relationship.

subject, Gary Fields (1987, 1998) suggests that a U-shaped pattern is more plausible than an inverted U-shape. In particular, Fields demonstrates that if average income in a population grows due to a steady growth in the relative number of wealthy individuals measures of income inequality will first decrease before eventually increasing. As this may indeed be the situation with MLB over time, it will be important to consider this. With our data, we are able to test whether the relationship proposed by Kuznets or the one proposed by Fields best describes MLB. In addition, we can evaluate the possibility that both the Kuznets and the Fields effects occur in MLB.[5] If the Kuznets effect occurs at low average salary levels and is followed by the Fields effect at higher average salary levels, then the MLB Kuznets curve will exhibit an "N-shape" or, more precisely, a "sideways, mirrored S-shape." Conversely, if the Fields effect occurs at low average salary levels and the Kuznets effect occurs at high average salary levels, the MLB Kuznets curve will be "sideways S-shaped."

III. Methodology

We use a non-parametric and a parametric approach in constructing the Gini coefficients. Once the inequality measurements are constructed, the Gini coefficient and the parameters of the underlying distribution (if applicable) are regressed on a set of team-specific data that include team performance variables as well as economic and demographic information of the metropolitan area in which the team is located. The purpose of the regression is to determine what variables have the tendency of increasing or decreasing income inequality within the team. In addition, we analyze a measure of skewness and a function of the parameters of the underlying distribution.

[5] We note that it will be difficult, if not impossible, to generalize our results to (for instance) developing countries. Thus, in the context of the Kuznets curve, this study is intended only as an application of existing theory to

Non-Parametric Gini Coefficients

First, in the non-parametric approach, we use the traditional Gini coefficients based on non-grouped data. The formula used here is very straightforward and is presented in equations (1a) and (1b).

$$Gini = 1 - \frac{1}{N}\left\{1 + \frac{2}{N\overline{y}}\sum_{i=1}^{N}(N-i)y_i\right\} \tag{1a}$$

or equivalently

$$Gini = \frac{1}{2N^2\overline{y}}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|y_i - y_j\right| \tag{1b}$$

where y_i is the income of the i^{th} individual, $\overline{y}$ is the mean population income, and N is the size of the population.

Parametric Gini Coefficients

To obtain the Gini coefficient based on a parametric distribution, we select one from the Pearson family of distributions (Kendall and Stuart, 1977). One advantage of this distribution family is that most of the functional forms commonly used to analyze the size of income distribution (such as the Beta I, the Gamma, the Beta II, etc.) belong to this family. Furthermore, this particular distribution family allows discarding certain distributions *a priori* by the use of the Kappa Criterion (**K**-criterion) (Elderton and Johnson, 1969). The **K**-criterion test begins by computing the value of **K** (given in equation (3) below) using the empirical moments. The value of **K** constructed from the empirical moments is then compared to the magnitude and sign of **K** constructed from the theoretical moments of a known distribution to determine which distributions are inappropriate for the data in question. Hirschberg et al. (1988-89) list the value

professional sports markets.

of **K** for the major income distribution functions. The moments about the mean are shown in equations (2a), (2b), and (2c) (Kendall and Stuart, 1997, p. 58):

$$\mu_2 = \mu'_2 - (\mu'_1)^2 \tag{2a}$$

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)3 \tag{2b}$$

$$\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4 \tag{2c}$$

where μ'_j is the j^{th} empirical raw moment of the data under observation. The **K**-criterion is presented in equation (3) (Elderton and Johnston, 1969, p.51):

$$\kappa = \frac{\beta_1(\beta_2 + 3)^2}{4(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)} \tag{3}$$

where

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \tag{4a}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \tag{4b}$$

The results of estimating the **K**-criterion for each MLB team between the years of 1985 and 2000 show that in the majority of the cases the **K**-criterion is negative.[6] A review of the distributions presented by Hirschberg et al. (1988-89) in their Table 1 (p. 187) indicates that of the 11 distributions, only the Beta I distribution first used by Thurow (1970) is less likely to be discarded.[7] Equation (5) gives the three parameter Beta I distribution.

---

[6] In one case, the **K**-criterion could not be estimated due to insufficient observations (Texas in 1987); in another case, the **K**-criterion was positive (Baltimore in 1985).

[7] The Beta I distribution used by Thurow (1970) was only a two parameter distribution (*p*, *q*). The properties of this distribution are presented in MacDonald (1984). We use a three parameter distribution (*p*, *q*, *y*) in order to provide greater flexibility in the estimation and because it is easy to show that the **K**-criterion for the three parameter Beta I

$$F(u;p,q,y)=\frac{1}{B(p,q)}\int_{u=0}^{Y}\frac{u^{p-1}(y-u)^{q-1}}{y^{p+q}}\,du \qquad (5)$$

where B(p, q) is the complete Beta function with parameters p and q (Rainville, p. 18). MacDonald (1984) presents the Gini coefficient for the Beta I distribution in equation (5), the parameters of which were estimated using the method of moments. The raw moments of this distribution are:

$$\mu'_j=\frac{y^j B(p+q,j)}{B(p,j)} \qquad (6)$$

From equation (6) it can be shown (see Appendix One) that the first three raw moments of this distribution are:

$$\mu'_1=\frac{yp}{p+q} \qquad (7a)$$

$$\mu'_2=\frac{y^2 p(p+1)}{(p+q)(p+q+1)} \qquad (7b)$$

$$\mu'_3=\frac{y^3 p(p+1)(p+2)}{(p+q)(p+q+1)(p+q+2)} \qquad (7c)$$

The Gini coefficient presented by MacDonald (1984) can be rewritten as in equation (8) (see Appendix One):

$$Gini_{Beta\ I}=\frac{\Gamma(p+q)\Gamma(p+\frac{1}{2})\Gamma(q+\frac{1}{2})}{\Gamma(p+1)\Gamma(q)\Gamma(\frac{1}{2})\Gamma(p+q+\frac{1}{2})} \qquad (8)$$

where Γ is the gamma function.

---

is also negative and that the Gini coefficient does not include the third parameter (y). The **Κ**-criterion for the Gamma distribution (Salem and Kotz, 1974) is asymptotically $\pm\infty$.

Note that the Gini coefficient in equation (8) depends only on the parameters *p* and *q*; *y* is simply a scalar. Thurow (1970) discusses changes in the *p* and *q* parameters in terms of median income rather than the Gini coefficient. The fact that we provide the Gini coefficient for the three parameter Beta I distribution allows us to view the impact on inequality in terms of changes in *p* and *q*. This is done, as Parker (1999) does for the Beta II distribution, by plotting all relevant values of $p, q \in \Re^2$ using Mathematica (Wolfram, 1999). In Figure One, note that as *p* increases the Gini coefficient does not decrease below 0.25. On the other hand, even if the *q* increases to as high as 40, it cannot increase the value of the Gini coefficient much. In other words, as *p* increases, the Gini coefficient decreases, whereas as *q* increases, the Gini coefficient increases (i.e. $\frac{\partial G}{\partial p} < 0$, and $\frac{\partial G}{\partial q} > 0$). Figure One thus indicates the strong dominance of the *p* over the *q* parameter. That is, large changes in the q parameter lead to very small changes in the Gini coefficient whereas relatively small changes in the *p* parameter lead to relatively large changes in the Gini (i.e. $\left|\frac{\partial G}{\partial p}\right| > \left|\frac{\partial G}{\partial q}\right|$). This observation becomes crucial when examining the results of regressing these parameters on team-specific independent variables.

{INSERT FIGURE ONE}

Finally, we evaluate skewness for the distribution in equation (5). The square root of equation (4a) is a measure of skewness (Kendall and Stuart, 1977), and with the use of equations (7a-c) and (2a-b), it is trivial to show the skewness to be equations (9a-b).

$$SK = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} \qquad (9a-b)$$

$$SK = \frac{2\,(q-p)}{(2\,p+q)}\sqrt{\frac{1+\,p+q}{pq}}$$

The impact of changes in p and q on SK are:

$$\frac{\partial SK}{\partial q} = \frac{q^2 + 6p + 7pq + 4p^2}{pq\sqrt{\frac{1+p+q}{pq}}\,(q+2p)^2} > 0$$

$$(10a-b)$$

$$\frac{\partial SK}{\partial p} = -\frac{6q + 5q^2 - 5pq - 2p^2}{pq\sqrt{\frac{1+p+q}{pq}}\,(q+2p)^2} < 0$$

Furthermore, if we compare the value of these two derivates, it is clear that for equivalent changes in q and p the overall effect will be determined by the direction of the impact of q as seen in equation (11).

$$\frac{\partial SK}{\partial q} - \frac{\partial SK}{\partial p} = \frac{6(p+q)\sqrt{\frac{1+p+q}{pq}}}{(p+2q)^2} > 0 \qquad (11)$$

IV. Data and Results

A review of the MLB salary figures from 1985-2000 data provides the interesting result that the type of inequality the Blue Ribbon Panel discussed is a relatively recent phenomenon: From 1985 to 1994, the Gini coefficient for the population of teams averaged 0.148, while from 1995 to 2000, that same figure was 0.205. While this payroll inequality between teams is certainly an interesting and important issue, there are other matters of equal import not addressed in the Panel's report. Our focus is on payroll inequality *within* teams. While rich teams may be getting richer, are all players on such teams benefiting? What are the trends in within-team payroll inequality over time and across teams? How does within-team payroll inequality vary between "poor" teams and "rich" ones? Between-team payroll inequality affects MLB in that reduced parity between teams may result in a loss of interest on the part of fans of small-market teams, resulting in lower attendance and a widening of the payroll gap. Changes in within-team payroll inequality will have different, and equally important, effects. For example, a more equal distribution of payroll within a team may promote team cohesiveness, translating into team success. Alternatively, a less equal payroll distribution may signal the presence of a few high-quality players that increase the probability of a successful season.[8]

Table One lists descriptive statistics for the entire sample. The individual player salary data used in this study are available from a number of internet sources.[9] The first two variables presented in this table correspond to two different Gini coefficients. The first, *gini1*, is the Gini coefficient estimated non-parametrically as in equations (1a) or (1b). The second Gini, *gini2,* is the Gini coefficient estimated parametrically based on the Beta I distribution and equation (8). It is important to note that the number of Gini coefficients estimated under the two scenarios (non-parametric and parametric) is not the same. There are 20 teams with 26 years of data (416 observations), 2 teams with 8 years (16), and 2 teams with 3 years (6) for a total of 438 observations. However, in 5 instances the non-parametric Gini coefficients (*gini1*) could not be obtained because the data set did not contain information on enough players. In the case of the parametric Gini (*gini2*), the **K**-criterion resulted in a negative number 436 times.[10] However, in only 407 cases did we obtain parameters (*p* and *q*) that were actually positive despite the fact that

[8] See Depken (2000) for a detailed discussion of the potential effects of within-team salary inequality. Also, see Ehrenberg and Bognanno (1990a, 1990b) for a discussion of salary inequality in the context of tournament theory and its application to an individual sport, professional golf.
[9] We rely on the collections of Rod Fort (users.pullman.com/rodfort) and Sean Lahman (baseball1.com) as well as the archives of USAToday (usatoday.com). Wherever possible, we crosschecked figures from all of these sources. In addition, we have cleaned the data so that the numbers reflect opening day salaries in most cases. For some years, we are unable to differentiate between yearly salaries and added bonuses. In the years in which we are able to separate out bonus payments, these payments do not significantly change teams' salary distributions. Thus, we are confident that the inclusion of bonuses in some years will not bias the Gini coefficients for those years. However, as is the case with some of the data stored on the internet, there may be some errors in the data.
[10] See footnote 6.

the **K**-criterion is negative.[11] Nonetheless, it is clear that for this data set that the Beta I distribution is the least likely distribution of the Pearson family to be discarded.

The other variables included in our analysis can be separated into three categories. The first category contains time-trend variables. The two remaining categories contain team-specific cross-sectional variables and variables measuring the economic and demographic conditions of the markets in which teams are located. There are four time trend variables: a trend variable (*trend*), its square ($trend^2$), and cube ($trend^3$), and a dummy variable that controls for the effects of the 1994 strike (*strike*).[12] *Strike* equals 1 if the year is 1995 or later and 0 otherwise; it is designed to ascertain whether the work stoppage of 1994 has had any impact on within-team equality.

Our team-specific measures include per capita salary (in millions of constant 1990 dollars) along with its square and cube (to test for the Kuznets and the Fields effects), the number of regular season wins (*wins*), and the number of All-Star players on the team (*all-stars*). The average age of a team's players (*age*) is included to (at least partially) control for MLB's restraints on mobility and earnings power early in a player's career. Evidence suggests that MLB players are paid less than their marginal revenue products in their first six years while they are ineligible for free agency and that after free agency, players' salaries rise considerably. All else constant, we expect that increases in average age will be associated with decreases in payroll inequality, since the older a team's players, the more of them are eligible for free agency and the tighter the salary distribution should be.[13] In addition, we include dummy variables to control for the league in which the team played (*national league*), since there may be different institutional

[11] The **K**-criterion only indicates whether a distribution can be discarded a priori. It does not indicate which distribution is the "best" fit or whether the parameters estimated will be appropriate.

[12] The three time trend variables (*trend*, $trend^2$, and $trend^3$) are included to control for any shape observed in the Kuznets Curve that is simply a product of time.

forces at work in the National and American Leagues. We include another dummy variable to control for differences between existing and expansion teams (*expansion team*). Expansion teams may operate under different cost structures or have different expectations for team success than existing teams. The market-specific variables include local MSA population and median household income. Team-specific performance measures are collected from the Total Baseball web site (total baseball.com), an online version of the official encyclopedia of MLB. MSA population and income measures are for 1990 from the US Census web site (census.gov).

{INSERT TABLE ONE}

Estimating a Simple Kuznets Curve for MLB

Table One shows that for most of the time period, computing payroll inequality based on the parameters of the Beta I distribution leads to larger Gini coefficients than for those computed using the non-parametric method, with *gini2* being on average about five percent more than *gini1*.[14] We begin by discussing the results from a pooled-data regression of the two Gini coefficients on average salary figures as reported in Table Two. This regression allows us to analyze a Kuznets curve for MLB. From the estimates presented in Table Two, we predict the Gini coefficients and graph them for different salaries in Figure Two.[15] Note that *gini2* is consistently above *gini1*, although the overall pattern with respect to salary levels is the same for both Gini coefficients.

{INSERT TABLE TWO AND FIGURE TWO}

Starting with a low average salary level, we observe the effect postulated by Kuznets followed the effect postulated by Fields. That is, the relationship between the Gini coefficient and average salary (the Kuznets curve) is the sideways, mirrored S-shaped. Based on the results

[13] There is a substantial amount of research on free agency in MLB. Readers interested in the specifics of MLB free agency are directed to Krautmann (1999).

in Table Two, payroll inequality within teams seems to increase as average salary increases to approximately $1,000,000. Payroll inequality then decreases slightly as average salary increases from $1,000,000 to approximately $2,000,000. As average salary increases past $2,000,000, payroll inequity within MLB teams rises. Thus, for teams with low payrolls, any increase in average salary will tend to increase payroll inequality, while teams with mid-range payrolls will tend to have decreased payroll inequality. However, high-payroll teams will see a fairly dramatic increase in inequality with increases in average salary. In addition, a visual inspection of Figure Two lends support to an hypothesis that within-team payroll inequality generally increases from the lowest to the highest paying teams.

The sideways, mirrored S-shaped MLB Kuznets curve can be explained by considering the potential effect on team payroll distribution from adding a superstar player. If a team signs such a player, we expect that its average salary will increase. Teams with low total payrolls (or low average salary levels) may have relatively equal within-team payroll distributions, since they may be unable to afford even one high-priced superstar. For a low-payroll team, adding the superstar player surely means paying that player much more than the current average salary. Hence, payroll inequality increases. For a mid-range-payroll team, adding the superstar player may cause a decrease in inequality, since the team may already have one or two superstars on their roster. The wealthiest MLB teams may be able to afford a number of expensive players, so these teams are expected to have the highest average salary levels. If a high-payroll team is to add another player and increase average salary, then said player would have to be an extremely high-paid mega-superstar. Signing such a player may increase payroll inequality rather dramatically.

---

[14] A complete listing of *gini1* and *gini2* by team and year is given in Appendix Two.
[15] We do not include a constant term since the Gini coefficient should be zero at a salary of $0.

More Complete Estimates of Payroll Inequality in MLB

Table Three reports the estimates from panel data regressions of Gini coefficients on the full set of determinants of payroll inequality.[16] Note that the results using *gini1* and those using *gini2* are remarkably similar. We find a significant relationship between payroll inequality and the time trend, even after controlling for *strike* as a shift parameter. The implications of our results with respect to the MLB work stoppage are especially interesting. The strike of 1994 increased within-team payroll inequality noticeably: holding other factors constant the Gini coefficient increased by between 0.11 (using *gini2*) and 0.13 (using *gini1*) from 1994 to 1995, which is an increase of approximately 20 to 25 percent. If, as some have stated (Staudohar, 1997), the players were the "winners" of the strike, then the players who have reaped the greatest rewards are those who are the highest paid on each team.

{INSERT TABLE THREE}

Table Three also indicates that both *gini1* and *gini2* continue to exhibit a sideways, mirrored S-shaped relationship with average salary, similar to the results in Table Two. As expected, average team age is negatively correlated with payroll inequality. This result suggests that labor issues associated with player mobility and market power significantly affect within-team payroll inequality; this implication becomes even more obvious when one considers the coefficient on *strike* discussed above. The number of wins decreases payroll inequality, indicating that better teams are those with less payroll inequality; however, the effect is small and insignificant when using *gini1*.[17] Expansion teams seem to have more payroll inequality than existing teams: such teams have Gini coefficients about 0.03 points higher than those that do not.

[16] We employ a random-effects, panel data estimator using the XTREG command in STATA (StataCorp, 1999). A Hausman test (available from the authors) indicates that a random-effects model is more appropriate than a fixed-effects model. We include a constant term in these regressions since some included measures are dummy variables.

[17] Due to the potential endogeneity of wins, we estimate the models presented in this paper without *wins*. These results show no significant differences from those presented in this paper and are available from the authors.

This may indicate that expansion teams are impatient to be successful and have been willing to pay for that success by signing a few high-priced stars. Perhaps most interesting of all are the coefficients on market size. Market size, measured as population or income, appears to have no significant effect on within-team payroll inequality. Although MLB may be concerned about the gap between the payrolls of large- and small-market teams, we find no evidence that the small-market/large-market distinction affects how a team distributes salaries among its players.

To further explore the Kuznets curve relationship, we predict *gini1* and *gini2* from Table Three, and graph these predictions at different average salary levels for an "average team."[18] The predicted Kuznets curves are presented in Figure Three. A visual inspection of the predicted Kuznets curves reveals some interesting facts. First, as expected from the numerical results, the curves are similarly shaped, both having the now familiar sideways, mirrored S-shape. That is, controlling for time effects and other variables, we still observe both the Kuznets and Fields effects in MLB. Second, the two Kuznets curves track each other, with the largest differences occurring at low average salary levels. It appears that there is a small difference in the general predicted trend: using the non-parametric Gini, there seems to be a slight upward trend in payroll inequality when moving from lowest to highest salary; using the parametric Gini, there seems to be no such trend from lowest to highest salary. Cleary, the results from the two estimations are different, but for the most part the differences are in terms of intercept shifts not in the substantive relationship between average salary and payroll inequality.

{INSERT FIGURE THREE}

[18] This average team has the following characteristics (as suggested by the sample average in Table One): the year is 1990, and the average age of the team is 28.5 years; the team has 80 wins, 2 All-Stars, and plays in the American League; the team is located in an MSA with a population of 5.6 million and a median household income of $35,390. The predictions will change somewhat with a change in any of these characteristics; however, the relative shapes of the predicted Kuznets curves will be the same, and any difference will result from an intercept shift.

Third, comparing Figure Three to Figure Two shows another intriguing result: there seems to be much less of upward trend in payroll inequality from low to high salaries in Figure Three than was the case in Figure Two. The predictions shown in Figure Three do not include the effect of time, since they show Kuznets curves for a representative team in a single year. However, time is a factor in Figure Two since the data are pooled. Therefore, we might conclude that much of the upward trend is due to structural changes over time or due to changes in time-sensitive variables, although it is clearly possible that some other mechanism is at work.

The Determinants of the Beta I Parameters

Having established that we do not lose much information when using the parametric rather than the non-parametric Gini coefficients, we now turn to an analysis of the parameters of the Beta I distribution. Table Four reports results from panel data regressions of *p* and *q* on the same explanatory variables used in Table Three. The regressions provide further insight into the reasons for the observed changes in equality. As noted in Section III, the Gini coefficient is inversely related to the parameter *p*, and positively related to the parameter *q* (i.e., $\frac{\partial G}{\partial p} < 0$, and $\frac{\partial G}{\partial q} > 0$). The first interesting observation to be made from Table Four is the fact that the only variables that significantly affect *p* are the *trend*, $trend^2$, $trend^3$, *strike*, and *age*, most of which are indicators of time.[19] Recalling from Figure One that *p* is the more dominant parameter, it appears that time (or unobserved variables correlated with time) has had the most significant impact on payroll inequality in MLB. This gives support to the tentative conclusion reached based on the time pattern of Gini coefficients in Figures Two and Three. Furthermore, the time-motivated increase in inequality was accelerated by the strike in 1994. It is worthy of note that the other variables representing cross-sectional information (team- and market-specific information) have no impact on *p*. In particular, this implies that the shape of the MLB Kuznets curve cannot be attributed to the dominant parameter *p*.

{INSERT TABLE FOUR}

Turning our attention to estimates using the parameter *q*, it is interesting to note that the *trend* measures affect *q* but *strike* does not. Note that while the coefficient signs on *trend*, $trend^2$, and $trend^3$ are the same using *p* and *q*, and recall that *p* and *q* have different impacts on the parametric Gini. Consequently, the *trend* coefficients imply that over time, there have been changes in MLB (not observed in other included variables) that have increased inequality, while at the same time there have been changes that have decreased inequality.[20] That is, the coefficients on the *trend* variables are actually measuring different impacts of time on the distribution of income. Also interesting is the fact that changes in average salary do impact *q*. Thus, it appears that the Kuznets curve relationship between the Beta I Gini coefficient and average salary comes primarily from the *q* parameter, which Figure One showed to be less dominant than *p*. In addition, it appears that *q* varies with age and enhances the effect of age on payroll inequality observed when using *p*.

A Brief Discussion of Skewness

As indicated in equation (9b), the skewness of the Beta I distribution (*SK*) can be presented in terms of *p* and *q*. Therefore, using a parametric income inequality measure allows for more detailed analysis of the entire salary distribution than if one were to use a distribution free measure. Equation (11) indicates that for equivalent changes in *p* and *q*, the change in *q* will dominate that of *p*. Based on this information, we turn out attention to Table Five, in which we

[19] Looking at the time pattern of *age*, we find no evidence to suggest that there is a time trend in the average age of players in the league as a whole or on individual teams.
[20] One could speculate that these changes are due to fluctuating market conditions not picked up by our market-size variables, but an analysis of this time trend is beyond the scope of the present paper. Given our result that time is a

present the skewness (9b) based on the annual averages of *p* and *q*. Recall that a positive (negative) skewness measure indicates the function is skewed to the right (left). There seems to be a general upward trend in skewness from 1985 to 1990, and no consistent pattern after that; note that the highest measures of skewness occur in 1995 and 1996, the years directly after the strike, which gives further evidence of effect of labor strife on the salary distribution of MLB. As suggested by equation (11), the movement of *q* has a stronger impact on the magnitude of the skewness, which implies that those variables that affect *q* as shown in Table Four are also those that will have the largest impact on *SK*.

{INSERT TABLE FIVE}

V. Conclusion

This study analyzes income distributions within teams in Major League Baseball. Our findings will be of interest to sports economists in general and to those interested in MLB in particular. We find, for example, that within-team payroll inequality has generally increased since 1985. It is also noteworthy that the strike of 1994 had the effect of dramatically increasing inequality. In addition, we find that the average age of players on a team has a significant effect on payroll inequality; specifically, teams with older players have tighter salary distributions. These findings raise a number of questions for future research. For instance, how will future labor strife and inevitable institutional changes affect income distributions in professional sports? Our findings could also be of interest to development and labor economists. We find some evidence for the effects proposed by both Kuznets and Fields with respect to income inequality

---

major determinant of within-team payroll in MLB, the reasons for this time trend should be examined more fully in future research.

and average salaries. While our results are not easily generalized to other industries or to country-level analysis, it is clear from this data that both effects need not be mutually exclusive.

Our findings may also be of interest to a broader group of researchers. There have been a number of debates in the vast literature on income distribution. Some of these have concerned whether an income inequality measure should be based on a specific distribution or be distribution free. While our results do not suggest which approach is "better," we do give evidence of the relative usefulness of parametric income inequality measures as compared to non-parametric measures. The method suggested here of regressing the impacts on the parameters, as well as on the Gini coefficient itself, provides additional insight into the determinants of inequality. In our sample of MLB teams, the dominant parameter, *p*, is mainly affected by the time trend variables; thus, we can infer that time is the strongest determinant of changes in payroll inequality within MLB teams. Furthermore, salary measures seem to impact the Gini coefficient of MLB teams only through changes in *q*; thus, the MLB Kuznets curve is a result of changes in *q*.

In the context of the Kuznets curve relationship, we find that the information employing non-parametric and parametric measures of payroll inequality in MLB have similar results. On the other hand, the use of parametric measures allow for better tracking of the various determinants of income inequality. This finding indicates that our method may be useful in analyzing data at the country level. For example, researchers studying a population's income distribution could estimate the values of the *p* and *q* parameters (and thus the Gini coefficient) even if the data were only available in quintiles or deciles. As is the case with MLB data, this decomposition would give researchers an altogether different angle in studying the determinants of changes in income inequality over time and in cross-sectional data. In addition, regressing the parameters is not hindered by the methodological problem of crossing Lorenz curves that occurs when using non-parametric Gini coefficients.

Jones et al. (1999) point out that: “(t)he existence of data for a variety of economic agents –industries (leagues), firms (teams), employees (players) – has facilitated the testing of a number of theoretical hypotheses which might otherwise be empirically barren.” We provide further evidence to support such a statement; in this case, we show that professional sports data can be used to study income distribution issues. There are other issues that could be discussed using the fruitful data base which professional sports provides, such as the welfare implications of a changing income distribution or the effects of policy intervention or institutional change on income distributions.

Table One
Descriptive Statistics
(n = 433)

Variable	Mean	Standard Deviation
gini1	0.537	0.088
*gini2**	0.565	0.100
trend	7.735	4.641
strike	0.397	0.490
real per capita salary/1 mil	0.882	0.447
age	28.556	1.181
wins	80.352	10.949
all-stars	2.185	1.363
national league	0.489	0.500
expansion team	0.050	0.219
population/1 mil	5.557	4.836
median household income/1000	35.390	4.412

* Sample size is 407 (see section IV)

Table Two
Pooled Regression Results: Kuznets Curves

Variable	Non-parametric Gini (*gini1*) n = 433 Coefficient	Std. Error	Parametric Gini (*gini2*) n = 407 Coefficient	Std. Error
real per capita salary/1mil	1.518***	0.027	1.589***	0.032
(real per capital salary/1 mil)^2	-1.207***	0.040	-1.265***	0.047
(real per capital salary/1 mil)^3	0.285***	0.014	0.298***	0.016
adjusted R-squared	0.9745		0.9697	

* significant at 10%
** significant at 5%
***significant at 1%

Table Three
Panel Data Estimates: Gini Coefficients

	Non-Parametric Gini (*gini1*) n = 433		Parametric Gini (*gini2*) n = 407	
Variable	Coefficient	Std. Error	Coefficient	Std. Error
constant	0.7267***	0.0846	0.9972***	0.0926
trend	0.0538***	0.0059	0.0491***	0.0064
$trend^2$	-0.0061***	0.0010	-0.0038***	0.0011
$trend^3$	0.0002***	0.0000	0.0001	0.0000
strike	0.1344***	0.0145	0.1126***	0.0153
real per capita salary/1 mil	0.5560***	0.0813	0.2495***	0.0974
$(real\ per\ capital\ salary/1\ mil)^2$	-0.4379***	0.0700	-0.2568***	0.0815
$(real\ per\ capital\ salary/1\ mil)^3$	0.1045***	0.0184	0.0677***	0.0209
age	-0.0178***	0.0030	-0.0220***	0.0031
wins	-0.0003	0.0003	-0.0006**	0.0003
all-stars	-0.0001	0.0024	0.0018	0.0025
national league	-0.0040	0.0058	-0.0045	0.0060
expansion team	0.0322***	0.0130	0.0350***	0.0135
population/1 mil	-0.0005	0.0006	0.0001	0.0006
median household income/1000	0.0004	0.0007	0.0002	0.0007
chi square (14)	711.84***		913.65***	

* significant at 10%
** significant at 5%
***significant at 1%

Table Four
Panel Data Estimates: Parameters of Beta I Distribution

	p n = 407		*q* n = 407	
Variable	Coefficient	Std. Error	Coefficient	Std. Error
constant	-0.1265	0.4031	35.0492***	10.7267
trend	-0.2915***	0.0278	-2.1493***	0.7994
$trend^2$	0.0324***	0.0048	0.3340**	0.1389
$trend^3$	-0.0010***	0.0002	-0.0127**	0.0058
strike	-0.3100***	0.0664	-1.4161	1.8887
real per capita salary/1 mil	-0.4995	0.4239	-32.8757***	11.9238
(real per capital salary/1 mil)^2	0.3148	0.3545	19.9295**	9.9920
(real per capital salary/1 mil)^3	-0.0618	0.0909	-3.7892	2.5667
age	0.0496***	0.0137	-0.6037*	0.3655
wins	0.0012	0.0014	0.0152	0.0382
all-stars	-0.0098	0.0109	0.0916	0.3016
national league	0.0251	0.0261	0.5998	0.5698
expansion team	-0.0569	0.0589	1.7824	1.4945
population/1 mil	-0.0009	0.0027	0.0670	0.0580
median household income/1000	0.0036	0.0032	0.0300	0.0685
chi square (14)	482.10***		61.26***	

* significant at 10%
** significant at 5%
***significant at 1

Table Five
Average *SK* for the Beta I Distribution

YEAR	***SK***
1985	-1.440
1986	0.176
1987	0.354
1988	0.375
1989	0.848
1990	1.088
1991	0.647
1992	0.546
1993	1.121
1994	0.651
1995	2.625
1996	2.025
1997	1.560
1998	0.931
1999	1.430
2000	1.167

Figure One
Sensitivity of the Parametric Gini
Coefficient (*gini2*) to *P* and *Q*

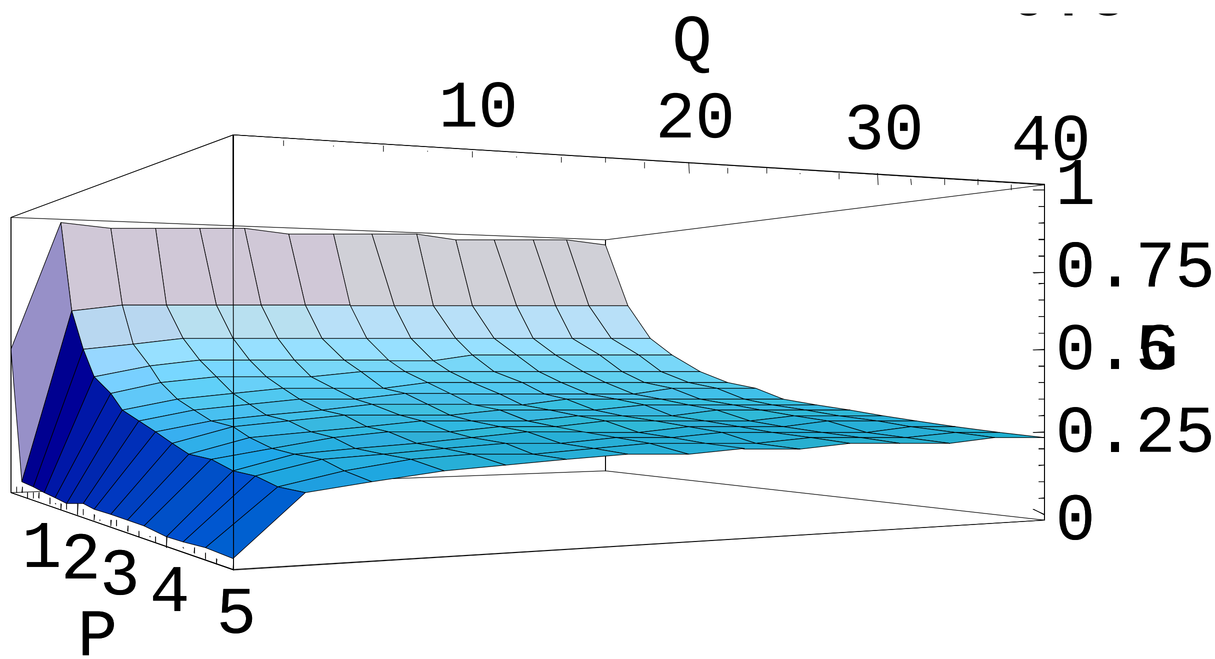


Figure Two
Simple Kuznets Curve

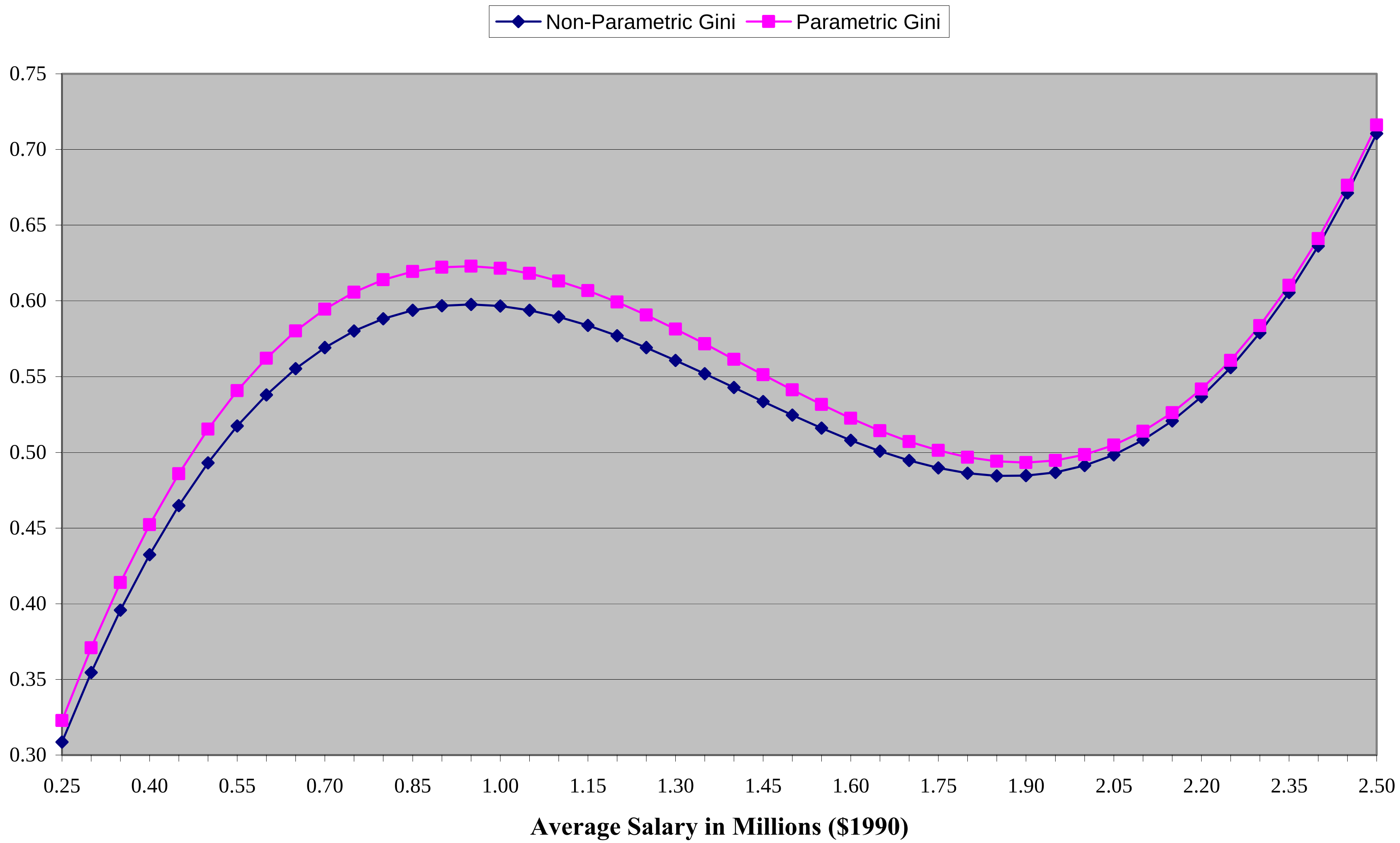
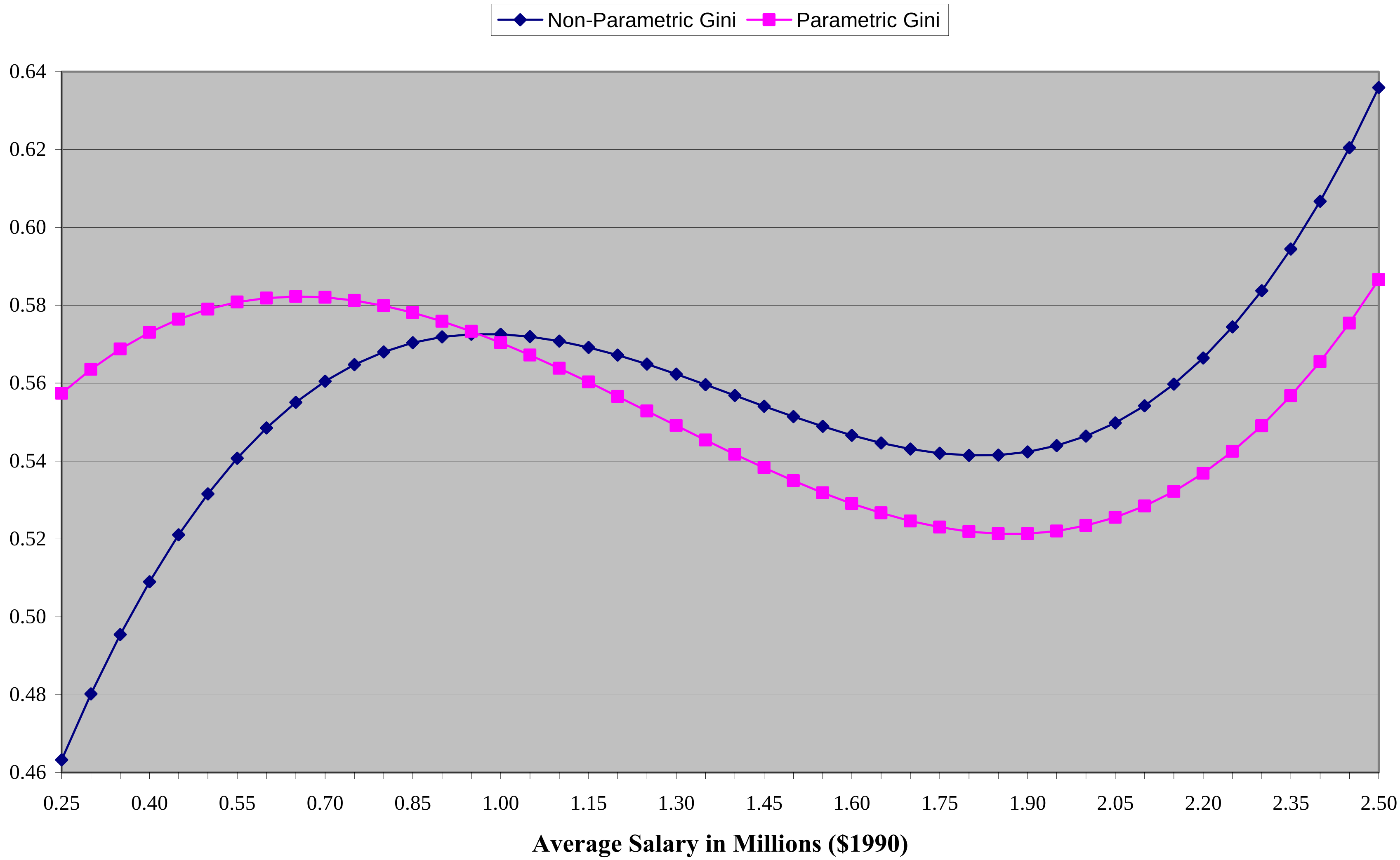


Figure Three
Kuznets Curve for Average Team



Appendix One

J. MacDonald (1984) provides the raw moment for the Beta distribution (5) in this text to be equation (6) (A.1 here):

$$\mu'_j = \frac{y^j B(p+q, j)}{B(p, j)} \tag{A.1}$$

Recall from Rainville (1960) that if the real numbers α and β are both positive then the Beta function is related to the gamma function in the following manner:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{A.2}$$

so that (A.1) can be rewritten as:

$$\mu'_j = \frac{y^j \dfrac{\Gamma(p+q)}{\Gamma(p+q+j)}}{\dfrac{\Gamma(p)}{\Gamma(p+j)}} \tag{A.3}$$

Further more, recalling the factorial function is:

$$(\alpha)_n = \prod_{k=1}^{n} (\alpha + k - 1) \tag{A.4}$$

and that the factorial function and the gamma function are related as in (A.5) (Rainville, 1960)

$$(\alpha)_n = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} \tag{A.5}$$

it is clear that (A.3) is:

$$\mu'_j = \frac{y^j (p)_j}{(p+q)_j} \tag{A.6}$$

The Gini coefficient for the Beta I distribution function in equation (5) in the text is shown to be (A.7) in J. MacDonald (1984):

$$G = \frac{B\left(p+q,\frac{1}{2}\right)B\left(p+\frac{1}{2},\frac{1}{2}\right)}{B\left(q,\frac{1}{2}\right)\pi} \tag{A.7}$$

Recalling the relation between the gamma and the beta function in (A.2) and that as Rainville (1960) shows:

$$\sqrt{\pi} = \Gamma\left(\frac{1}{2}\right) \tag{A.8}$$

it is trivial to show that (A.7) can be written as:

$$G_{Beta\ I} = \frac{\Gamma(p+q)\Gamma\left(p+\frac{1}{2}\right)\Gamma\left(q+\frac{1}{2}\right)}{\Gamma(p+1)\Gamma(q)\Gamma\left(\frac{1}{2}\right)\Gamma\left(p+q+\frac{1}{2}\right)} \tag{A.9}$$

Appendix Two
Gini Coefficients by Team and Year

		1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	*average*	*difference*
Anaheim	***gini1***	0.340	0.440	0.441	0.475	0.435	0.493	0.505	0.552	0.647	0.570	0.684	0.630	0.561	0.572	0.588	0.642	*0.536*	*0.019*
	gini2	0.343	0.424	0.448	0.487	0.441	0.502	0.514	0.561	0.686	0.595	0.736	0.682	0.584	0.596	0.621	0.664	*0.555*	
Arizona**	***gini1***														0.649	0.575	0.562	*0.595*	*0.031*
	gini2														0.715	0.595	0.571	*0.627*	
Atlanta	***gini1***	0.298	0.437	0.454	0.526	0.550	0.560	0.461	0.490	0.501	0.521	0.620	0.650	0.623	0.645	0.564	0.574	*0.530*	*0.016*
	gini2	0.308	0.445	0.461	0.542	0.588	0.601	0.469	0.502	0.516	0.527	0.637	0.672	0.640	0.664	0.572	0.584	*0.546*	
Baltimore	***gini1***	0.297	0.388	0.524	0.555	0.624	0.427	0.554	0.522	0.586	0.555	0.884	0.605	0.545	0.460	0.494	0.547	*0.535*	*0.025*
	gini2	****	0.389	0.531	0.567	0.683	****	0.607	0.536	0.609	0.570	0.666	0.619	0.554	0.468	0.497	0.550	*0.560*	
Boston	***gini1***	0.368	0.462	***	0.550	0.514	0.537	0.471	0.491	0.554	0.534	0.646	0.616	0.576	0.567	0.539	0.533	*0.531*	*0.011*
	gini2	0.373	0.452	***	0.560	0.526	0.552	0.479	0.500	0.571	0.544	0.694	0.636	0.588	0.576	0.546	0.527	*0.542*	
Chicago (NL)	***gini1***	0.273	0.375	0.456	0.563	0.578	0.559	0.555	0.574	0.577	0.564	0.676	0.623	0.602	0.570	0.401	0.591	*0.533*	*0.032*
	gini2	0.275	0.388	0.469	0.577	0.623	0.614	0.574	0.603	0.598	0.579	0.743	0.653	0.622	0.579	0.539	0.602	*0.565*	
Chicago (AL)	***gini1***	0.396	0.501	***	0.487	0.483	0.512	0.575	0.590	0.505	0.491	0.608	0.626	0.666	0.708	0.516	0.646	*0.554*	*0.041*
	gini2	0.401	0.514	***	0.509	0.497	****	0.606	0.617	0.515	0.496	0.618	0.639	0.692	0.761	0.748	0.723	*0.595*	
Cincinnati	***gini1***	0.444	0.492	0.490	0.455	0.524	0.533	0.520	0.514	0.541	0.505	0.608	0.631	0.634	0.603	0.593	0.570	*0.541*	*0.026*
	gini2	0.456	0.492	0.509	0.466	0.545	0.573	0.531	0.520	0.553	0.554	0.619	0.667	0.661	0.687	0.646	0.602	*0.567*	
Cleveland	***gini1***	0.318	0.473	0.473	0.467	0.485	0.484	0.564	0.341	0.429	0.493	0.579	0.567	0.542	0.482	0.558	0.539	*0.487*	*0.036*
	gini2	****	0.489	0.483	0.484	****	0.501	0.586	****	0.437	0.501	0.594	0.577	0.547	0.486	0.570	0.549	*0.523*	
Colorado*	***gini1***									0.409	0.575	0.642	0.628	0.622	0.567	0.535	0.542	*0.565*	*0.028*
	gini2									0.433	0.607	0.674	0.654	0.643	0.586	0.554	****	*0.593*	
Detroit	***gini1***	0.308	0.400	0.425	0.443	0.504	0.538	0.401	0.597	0.549	0.432	0.699	0.717	0.548	0.491	0.542	0.536	*0.508*	*0.019*
	gini2	0.313	0.407	0.432	0.449	0.513	0.560	0.409	0.617	0.565	0.437	0.736	0.840	****	0.517	0.561	0.546	*0.527*	
Florida*	***gini1***									0.622	0.634	0.684	0.639	0.613	0.728	0.609	0.605	*0.642*	*0.081*
	gini2									0.676	0.693	0.773	0.670	0.632	0.796	0.755	0.789	*0.723*	
Houston	***gini1***	0.330	0.398	0.407	0.426	0.431	0.477	0.590	0.453	0.575	0.565	0.695	0.649	0.631	0.611	0.522	0.586	*0.522*	*0.024*
	gini2	0.335	0.402	0.418	0.433	0.439	0.488	0.641	0.469	0.597	0.587	0.747	0.700	0.684	0.649	0.532	0.614	*0.546*	
Kansas City	***gini1***	0.374	0.445	0.543	0.494	0.508	0.522	0.520	0.544	0.532	0.508	0.691	0.629	0.613	0.602	0.593	0.532	*0.541*	*0.025*
	gini2	0.379	0.449	0.555	0.500	0.524	0.534	0.527	0.563	0.542	0.513	0.773	0.690	0.643	0.630	0.655	0.573	*0.566*	
Los Angeles	***gini1***	0.361	0.457	0.493	0.490	0.504	0.525	0.496	0.468	0.557	0.491	0.660	0.588	0.573	0.601	0.535	0.581	*0.524*	*0.012*
	gini2	0.367	0.460	0.503	0.503	0.509	0.541	0.504	0.472	0.569	0.504	0.691	0.609	0.590	0.622	0.543	0.590	*0.536*	
Milwaukee	***gini1***	0.301	0.508	0.544	0.523	0.560	0.586	0.423	0.543	0.587	0.559	0.635	0.631	0.599	0.549	0.566	0.587	*0.544*	*0.031*
	gini2	0.307	0.526	0.580	0.547	0.588	0.630	0.512	0.556	0.609	0.584	0.723	****	0.666	0.572	0.582	0.635	*0.574*	
Minnesota	***gini1***	0.362	0.478	***	0.496	0.590	0.549	0.471	0.556	0.599	0.582	0.706	0.706	0.657	0.518	0.537	0.497	*0.554*	*0.034*

	gini2	0.365	0.487	***	0.506	0.620	****	0.483	0.575	0.626	0.596	0.767	0.792	0.705	0.530	0.584	****	*0.587*	
Montreal	***gini1***	0.351	0.548	0.531	0.485	0.493	0.527	0.613	0.599	0.630	0.637	0.541	0.607	0.573	0.407	0.502	0.594	*0.540*	*0.054*
	gini2	0.354	0.571	****	0.497	0.500	0.547	0.662	0.652	0.714	0.687	****	0.680	0.631	****	0.574	0.646	*0.594*	
New York (NL)	***gini1***	0.477	0.536	0.527	0.493	0.547	0.522	0.461	0.536	0.574	0.637	0.697	0.621	0.646	0.533	0.484	0.506	*0.550*	*0.019*
	gini2	0.512	0.523	0.538	0.499	0.564	0.543	0.468	0.544	0.585	0.665	0.767	0.663	0.681	0.545	0.491	0.509	*0.569*	
New York (AL)	***gini1***	0.325	0.407	0.527	0.438	0.557	0.488	0.471	0.551	0.559	0.497	0.600	0.574	0.529	0.487	0.515	0.594	*0.508*	*0.010*
	gini2	0.333	0.415	0.546	0.448	0.571	0.496	0.475	0.560	0.570	0.502	0.613	0.583	0.542	0.496	0.520	0.605	*0.517*	
Oakland	***gini1***	0.366	0.537	0.516	0.427	0.408	0.500	0.424	0.558	0.574	0.552	0.698	0.701	0.671	0.578	0.550	0.580	*0.540*	*0.031*
	gini2	0.369	0.560	0.527	0.431	0.412	0.513	0.435	0.568	0.594	0.568	0.738	0.817	0.817	****	0.597	0.616	*0.571*	
Philadelphia	***gini1***	0.432	0.550	0.531	0.472	0.512	0.506	0.524	0.570	0.522	0.454	0.687	0.699	0.669	0.609	0.588	0.559	*0.555*	*0.039*
	gini2	****	****	0.543	0.481	0.534	0.532	0.541	0.599	0.537	0.461	0.733	0.762	0.725	0.644	0.651	0.581	*0.595*	
Pittsburgh	***gini1***	0.360	0.568	0.489	0.401	0.425	0.432	0.498	0.593	0.595	0.617	0.626	0.636	0.412	0.451	0.440	0.524	*0.504*	*0.053*
	gini2	0.363	0.592	0.519	****	****	0.440	0.517	0.609	0.617	0.662	0.691	0.681	****	****	0.455	0.544	*0.558*	
San Diego	***gini1***	0.433	0.458	0.499	0.506	0.508	0.512	0.566	0.575	0.655	0.620	0.677	0.583	0.528	0.509	0.546	0.606	*0.549*	*0.025*
	gini2	0.447	0.465	0.508	0.527	0.517	0.525	0.587	0.592	0.696	0.691	0.741	0.611	0.551	0.525	0.557	0.637	*0.574*	
San Francisco	***gini1***	0.345	0.386	0.424	0.312	0.399	0.530	0.529	0.517	0.586	0.498	0.692	0.705	0.590	0.522	0.562	0.602	*0.513*	*0.010*
	gini2	0.347	0.398	0.431	0.313	0.404	0.548	0.538	0.523	0.599	0.509	0.739	0.750	****	****	0.589	0.619	*0.522*	
Seattle	***gini1***	0.319	0.365	***	0.520	0.500	0.469	0.441	0.556	0.546	0.615	0.680	0.677	0.640	0.589	0.525	0.522	*0.531*	*0.037*
	gini2	****	0.372	***	0.562	0.514	0.492	0.457	0.572	0.556	0.651	0.725	0.713	0.670	0.603	0.529	0.529	*0.568*	
St. Louis	***gini1***	0.453	0.545	0.579	0.536	0.516	0.556	0.544	0.510	0.566	0.552	0.605	0.578	0.607	0.577	0.586	0.550	*0.554*	*0.019*
	gini2	0.464	0.570	0.603	0.548	0.523	0.578	0.564	0.524	0.597	0.565	0.627	0.594	0.631	0.588	0.612	0.569	*0.572*	
Tampa Bay**	***gini1***														0.644	0.617	0.618	*0.626*	*0.045*
	gini2														0.713	0.659	0.642	*0.671*	
Texas	***gini1***	0.303	0.634	***	0.412	0.535	0.536	0.548	0.640	0.639	0.574	0.659	0.606	0.661	0.548	0.497	0.552	*0.556*	*0.028*
	gini2	0.306	****	***	****	0.559	0.576	0.569	0.672	0.668	0.602	0.691	0.633	0.694	0.558	0.503	0.560	*0.584*	
Toronto	***gini1***	0.335	0.426	0.481	0.505	0.545	0.469	0.470	0.519	0.553	0.543	0.692	0.689	0.618	0.539	0.553	0.601	*0.533*	*0.018*
	gini2	0.339	0.432	0.488	0.514	0.563	0.483	0.486	0.525	0.562	0.562	0.720	0.746	0.639	0.546	0.570	0.643	*0.551*	
Average	***gini1***	*0.357*	*0.470*	*0.493*	*0.479*	*0.509*	*0.513*	*0.508*	*0.537*	*0.563*	*0.549*	*0.663*	*0.636*	*0.598*	*0.564*	*0.541*	*0.569*	0.534	*0.026*
	gini2	*0.366*	*0.468*	*0.505*	*0.498*	*0.532*	*0.538*	*0.528*	*0.561*	*0.586*	*0.572*	*0.703*	*0.679*	*0.641*	*0.602*	*0.580*	*0.601*	0.560	

* Expansion team in 1993

** Expansion team in 1998

*** Observation missing due to limited sample size (less than ½ opening-day roster)

**** Observation missing due to misfit of Beta I distribution

References

Berri, D. J. 2001. "Mixing the Princes and the Paupers and Other Determinants of Worker Productivity in the National Basketball Association." Unpublished manuscript.

Burger, J. D. 2001. "U or Inverted-U? That is the Question." *Review of Development Economics* 5:151-156.

Depken, C. A. 2000. "Wage Disparity and Team Productivity: Evidence from Major League Baseball." *Economics Letters* 67:87-92.

Elderton, W. P., and N. L. Johnson. 1969. *Systems of Frequency Curves*. Cambridge Univeristy Press.

Ehrenberg, R. G., and M. L. Bognanno. 1990a. "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour." *Industrial and Labor Relations Review* 43:74S-88S.

Ehrenberg, R. G., and M. L. Bognanno. 1990b. "Do Tournaments Have Incentive Effects." *Journal of Political Economy* 98:1307-1324.

Fields, G. 1987. "Measuring Inequality Change In An Economy With Income Growth," *Journal of Development Economics* 26: 357-374.

Fields, G. 1998. "Do Inequality Measures Measure Inequality," in *The Distribution of Welfare and Household Production: International Perspectives*, edited by S. P. Jenkins, A. Kapteyn, and B. M. S. Var Praag. Cambridge University Press.

Fort, R. 1992. "Pay and Performance: Is the Field of Dreams Barren?" in *Diamonds are Forever: The Business of Baseball* edited by P. M. Sommers. Brookings Institution Press.

Hirschberg, J. G., D. J. Molina, and D. J. Slottje. 1988-89. "A Selection Criterion for Choosing Between Functional Forms of Income Distribution." *Econometric Reviews* 7:183-197.

Jewell, R. T., M.A. McPherson, and J. W. Wenger. 2001. "Teaching Income Inequality More Effectively: An Example Using Major League Baseball." Unpublished manuscript.

Jewell, R. T., and D. J. Molina. 2001. "The Effect on Salary Distribution on Production: An Analysis of Major League Baseball." Unpublished manuscript.

Jones, J. C. H., S. Nadeau, and W. D. Walsh. 1999. "Ethnicity, Productivity, and Salary: Player Compensation and Discrimination in the National Hockey League." *Applied Economics* 31:593-608.

Kendall, Sir M., and A. Stuart. 1977. *The Advanced Theory of Statistics, Vol. 1*. MacMillan Publishing.

Krautmann, A. C. 1999. “What’s Wrong with Scully-Estimates of a Player’s Marginal Revenue Product.” *Economic Inquiry* 37:369-381.

Kuznets, S. 1955. “Economic Growth and Income Inequality.” *American Economic Review* 49:1-28.

Levin, R. C., G. J. Mitchell, P. A. Volker, and G. F. Will. 2000. “Report of the Independent Members of the Commissioner's Blue Ribbon Panel on Baseball Economics.” Major League Baseball.

List, J.A. and C.A. Gallet. 1999. “The Kuznets Curve: What Happens After the Inverted-U?” *Review of Development Economics* 3:200-206.

MacDonald, G. M. 1988. “The Economics of Rising Stars.” *American Economic Review* 78:155-166.

MacDonald, J. B. 1984. “Some Generalized Functions for the Size Distribution of Income.” *Econometrica* 52:647-664.

Papanek, G. F. and O. Kyn. 1986. “The Effect on Income Distribution of Development, the Growth Rate, and Economic Strategy.” *Journal of Development Economics* 23:55-65.

Parker, S. C. 1999. “The Beta as a Model for the Distribution of Earning.” *Bulletin of Economic Research* 5:243-251.

Paukert, F. 1973. “Income Distribution at Different Levels of Development: A Survey of the Evidence.” *International Labor Review* 108:97-125.

Rainville, E. D. 1960. *Special Functions*. Macmillan Publishing.

Ram, R. 1991. “Kuznets’s Inverted-U Hypothesis: Evidence from a Highly Developed Country.” *Southern Economic Journal* 57:1112-1123.

Rottenburg, S. 2000. “Resource Allocation and Income Distribution in Professional Sports.” *Journal of Sports Economics* 1:11-20.

Ryu, H. K., and D. J. Slottje. 1999. “Parametric Approximations of the Lorenz Curve,” in *Handbook of Income Inequality Measurement* edited by J. Silver. Kluwer Academic Publishers.

Salem, A. B. Z. and T. D. Mount. 1974. “A Convenient Descriptive Model of Income Distribution: The Gamma Density.” *Econometrica* 42:1115-1127.

Saith, A. 1983. “Development and Distribution: A Critique of the Cross-Country U-Hypothesis.” *Journal of Development Economics* 13:367-382.

Schmidt, M. B. 2001. “Competition in Major League Baseball: The Impact of Expansion.” *Applied Economic Letters* 8:21-26.

Schmidt, M. B., and D. J. Berri. 2001. “Competitive Balance and Attendance: The Case of Major League Baseball.” *Journal of Sports Economics* 2:145-167.

Sen, A. 1973. *On Economic Inequality*. Oxford University Press.

Silber, J. 1999. *Handbook of Income Inequality Measurement*. Kluwer Academic Publishers.

Sommers, P. M. 1998. “Work Incentives and Salary Distributions in the National Hockey League.” *Atlantic Economic Journal* 26:119.

StataCorp. 1999. *Stata Statistical Software: Release 6.0*. Stata Corporation.

Staudohar, P. D. 1997. “The Baseball Strike of 1994-95.” *Monthly Labor Review* 120:21-27.

Thurow, L. 1970. “Analyzing the American Income Distribution.” *American Economic Review* 60:261-269.

Tsakloglou, P. 1988. “Development and Inequality Revisited.” *Applied Economics* 20:509-531.

Wolfram. 1999. *Mathematica Version 4.0.2.0*.