

A New Test of Compensating Differences: Evidence on the Importance of Unobserved Heterogeneity

Marios Michaelides*

November 2008

Abstract

A dataset on professional basketball players is assembled which contains rich measures of worker ability, measures of employer non-pecuniary characteristics, and location amenities. Using this data, we test the compensating differences theory. Empirical evidence strongly supports the theory's predictions in this context, regardless of the functional form of the wage equation. We also find that when important measures of worker heterogeneity are omitted from the specification, the quality of the results is distorted and inference on the validity of the theory is misleading. Finally, when important measures of worker skills are omitted, linear regression models outperform the Box-Cox maximum likelihood alternatives with respect to obtaining hedonic estimates that are closer to those produced in the full specification case.

JEL Classifications: J31, J32, J4

Keywords: compensating differences, job characteristics, hedonic wages

* IMPAQ International, LLC. E-mail: mmichaelides@impaqint.com. Phone: +1(202)215-1171.

I. Introduction

Wage determination is one of the most widely discussed empirical issues in labor economics and many researchers have attempted to tackle it using various theories and accompanying statistical techniques. One of the most interesting and fruitful variations is the compensating differentials theory, which suggests that workers value on-the-job consumption of job characteristics and amenities that are associated with the nature and location of their employment. This theory implies that workers are willing to forfeit part of their earnings to hold jobs with desirable attributes, while they only accept jobs with undesirable characteristics if they are offered higher wages. As a result, jobs with desirable attributes offer lower wages, whereas jobs with undesirable characteristics offer higher wages.

A basic test of the theory's validity is to estimate a wage equation that includes measures of worker skills and measures of job characteristics. According to theory, desirable job characteristics will have a negative effect on wages while undesirable job characteristics will positively affect wages. Empirical results have been mixed in terms of supporting the theory's intuition. Although many studies have found strong connections between job non-pecuniary characteristics and wages, in some studies, the signs of the estimated parameters are incompatible with the theory's predictions or lack statistical significance.

Many are the possible explanations for the inconsistent empirical evidence. The most commonly discussed empirical issue is selection due to unobserved worker skills. In the setting implied by the theory, if the unobserved portion of worker skills is correlated with the non-pecuniary job characteristics, hedonic wage estimates for these characteristics may be biased. Assuming amenities are normal goods, the direction of the unobserved heterogeneity bias would

be the opposite of the sign of the effect of amenities on wages, producing signs that are not conformable with theory.

In addition, researchers do not have complete information on working conditions, which in most contexts are quite heterogeneous across industries and occupations. This may produce biased estimates of implicit wages if the omitted job characteristics are correlated with working conditions that are included in the hedonic wage equation. Moreover, even when measures of working conditions are available, their objectivity is questionable since in most microeconomic data, they are reported by workers.

The functional form of the wage equation may also be an issue. Typically, the hedonic wage equation is estimated using the linear regression model. However, there are doubts on whether this model is appropriate to produce consistent estimates of compensating differences, especially if important portions of worker skills are unobserved. Researchers have argued that using a more flexible estimation approach, like the Box-Cox maximum likelihood model, may compensate for unobserved worker skill differences and for a potential non-linear relationship between wages and job characteristics.

In this paper, we test the compensating differences theory using information from a newly assembled dataset on professional basketball players, which does not share the same limitations as other microeconomic data. Specifically, our data contains rich measures of worker skills, which are standard in the industry, the errors in measurement of skill are limited, and employers are more homogeneous within this industry compared to those in other contexts.¹ Therefore, many of the problems considered to have an undesirable effect on the quality of the estimated hedonic parameters are less of an issue in this context.

¹ See Kahn (2000) for a discussion of the advantages of using professional sports data in economic research.

Using this data, we show that professional basketball players receive higher wages in the presence of undesirable job characteristics and lower wages in the presence of desirable job attributes, as the theory suggests. We also show that omitting important measures of player skills distorts the quality and the magnitude of the estimated compensating differentials. Moreover, when player skill heterogeneity is largely accounted for, the estimation results are not sensitive to the functional form of the wage equation, while the partially linear and the classic linear regression models outperform the Box-Cox maximum likelihood alternatives when measures of worker ability are omitted from the specification.

II. Background

Adam Smith was the first to articulate the notion that “*the monetary and non-monetary benefits of different employments must in general be equal.*” This was perhaps the first discussion of the idea that jobs with undesirable working conditions would offer higher wages, making the sum of monetary and non-monetary benefits received equal across jobs. Formulation of this idea in a model comes down to a wage equation that is a function of worker marginal product and job attributes that affect worker utility (Rosen, 1986). The compensating differences (or hedonic wages) theory predicts that, controlling for skill differences across workers, undesirable job characteristics positively affect wages while desirable job attributes have a negative effect on wages. Intuitively, workers are willing to forfeit part of their compensation in favor of job attributes with positive utility value while they demand higher wages for non-pecuniary job characteristics that reduce worker utility.

There is substantial empirical work that investigates the validity of the compensating differentials theory using microeconomic data. Research has shown that workers receive positive wage premiums for unpleasant job characteristics, like inflexible working hours (Duncan and

Stafford, 1980; McNabb, 1989), risk of getting injured (Viscusi and Moore, 1987; Duncan and Holmlund, 1998), and income insecurity (Viscusi, 1978; Olson, 1981; Coates and Kumar, 1982). Others found that jobs of repetitive nature (Lucas, 1977), jobs with noisy environments (Duncan and Holmlund, 1998), and physically demanding jobs (Duncan and Stafford, 1980; Kim and Fishback, 1993) pay higher wages. There is also evidence that location amenities affect wages as predicted by theory (Roback, 1982; Blomquist et al, 1988).

However, there is also weak evidence of the theory's validity. Brown (1980) finds no evidence that physically demanding jobs offer higher wages, while Dorsey and Walzer (1983) and Garen (1988) find that jobs with high injury risks do not offer a significant wage premium. Moreover, in many of the cited studies, hedonic estimates for important job characteristics are often statistically insignificant, have the opposite sign than the one predicted by theory, or both. In addition, there is evidence in these studies that the magnitude and significance of estimated hedonics are sensitive to model specification.

Researchers have discussed a number of statistical nuisances that may explain the inconsistent empirical results, the most important of which, is unobserved worker heterogeneity. In most contexts, available worker characteristics only partially capture the heterogeneity in terms of worker skills. If desirable working conditions are normal goods, then they are positively correlated with unobserved worker skills. Therefore, if an important portion of worker skills is not controlled for in the wage equation, the coefficients of job characteristics will be biased against the theory's predictions (Lucas, 1977; Dorman and Hargstom, 1998). Researchers have noted the positive correlation between non-pecuniary job characteristics and unobserved worker skills (Duncan, 1976; Brown, 1980) and indicated that the produced selection bias may be quite substantial, tainting inference on the validity of the theory (Atrostic, 1982; Hwang et al, 1992).

Measurement of job characteristics may also explain some of the inconsistent empirical evidence. In available microeconomic data, respondents provide subjective responses about the quality of their workplace making the reliability of the empirical results questionable. Atrostic (1982), for example, points out that in combination with the issue of unobserved heterogeneity, the quantitative effect of measurement errors on the hedonic estimates is unknown, making inference on the validity of the theory very complicated. To avoid the measurement error issue, researchers have used data that provide objective information on job characteristics. For example, Olson (1981) uses the Dictionary of Occupational Titles, which includes job attributes by occupation, and Garen (1988) uses the Bureau of Labor Statistics measures of work hazards by industry to measure job attributes. However, these sources measure job attributes on the aggregate, thus do not capture the variation of working conditions at the employer level.

Furthermore, the sensitivity of the results to the linear specification of the wage equation that is often employed is not known to researchers, especially since the problem of unobserved heterogeneity is generally present. Previous work suggested that using the linear regression model may be inappropriate since the relationship between wages and job characteristics may be more complex (Atrostic, 1982; Anglin and Gencay, 1996; Ekeland et al, 2001). The reason is that, in addition to unobserved heterogeneity, non-linearities are reflected in the error term of the wage equation, putting in doubt the linear regression model assumption that the errors are normally distributed. Researchers have suggested that using the Box-Cox maximum likelihood model (Box-Cox MLE) instead of the linear regression model may address these issues. The Box-Cox MLE chooses the optimal transformation for the dependent and independent variables, to maximize the probability that the normality assumption holds and that consistent estimates of

the hedonic parameters are produced.² As a result, the use of the Box-Cox transformation technique is often employed when estimating hedonic price equations (Blomquist et al, 1988; Atrostic, 1982; Atkinson and Halvorsen, 1990; Cheshire and Sheppard, 1995).

In this paper, we estimate hedonic wages in a setting where rich measures of worker ability are available and show how results are affected when worker heterogeneity is not fully accounted for in the specification. We also examine how the choice of the wage equation functional form may affect estimates when good measures of worker ability are available and when such measures are not included in the specification.

III. Data

This paper uses data on all professional basketball players employed in the National Basketball Association (NBA) between 1999 and 2003. This data contains player statistics that capture their on-court performance. Among other information, it includes the number of games each player played in each season, the number of points, rebounds, steals, blocked shots, assists, turnovers, and fouls the player produced per game. The data also reports each player's race, age, weight, height, position, place of birth, the year the player entered the league, and the order in which the player was selected at the annual league draft.³

Player salaries are available from a variety of sources. Our main sources of player salaries are the USA Today NBA salary database and the established pro-basketball website Hoopshype.⁴ Player salaries are protected by a Collective Bargaining Agreement (CBA) between teams and

² Cropper et al (1988) show using simulations that the Box-Cox maximum likelihood model outperforms the linear regression model in the presence of selection on unobserved characteristics, when estimating rent differentials.

³ NBA player statistics and other player information are available on the official website of the NBA (www.nba.com) and in official annual NBA guides, like the Sporting News NBA guide.

⁴ The salary reports of USA Today (www.usatoday.com) and Hoopshype (www.hoopshype.com) have been cross-checked to ensure their consistency. Salaries reported by these two sources are also consistent with salary reports available by Sports Reference, LLC (www.basketball-reference.com).

players. The CBA sets the rules under which the negotiation takes place, sets compensation thresholds, and ensures that the players' rights are reserved during their employment. Some of the conditions of the CBA have a significant effect on player compensation. In the next section, we discuss in more detail how the CBA affects player wages and how we account for these effects in our regression analysis.

Since each NBA team functions as a franchise, certain employment conditions are the same across teams. Firm heterogeneity in this context is limited to differences in location amenities, in the quality of the team's coaching staff, and in team success. We obtain measures that capture coach and team quality through official reports of the NBA, as well as from the Association of Professional Basketball Research. To account for location amenities, additional data sources are used. Weather conditions (temperature, snowfall, and precipitation) are obtained from the National Climatic Data Center and the Meteorological Service of Canada. Information on metropolitan area population, crime rates, and other location characteristics were obtained from the US Census Bureau's State and City Data Book (1997-98), the 2000 US Census, and from the Canada Uniform Reporting Crime Survey, and the 2001 Canadian Census.

IV. Estimation Results

We estimate a classic hedonic wage equation that includes all available measures of player productivity and team-specific characteristics that capture employer heterogeneity:

$$\log W_{it} = a + \sum_k P_{it}^k \beta^k + \sum_m Z_{it}^m \gamma^m + u_{it} \quad (1)$$

Note that $\log W_{it}$ is the logarithm of earnings for player i in year t . P includes player output variables and other player characteristics and Z includes team and location characteristics.⁵ Three

⁵ See Table 1 for variable description.

measures of player output are included; *Minutes* (number of minutes played per game), *Offense* (offensive output per game), and *Defense* (defensive output per game).⁶

Controlling for player output, more experienced players are likely to earn higher wages for two reasons. First, experienced players provide leadership and serve as locker room personalities, especially for teams with championship aspirations. Second, the CBA guarantees minimum compensations for veteran players, depending on the number of years they have been in the league.⁷ Similarly, players that have been with the same team for a number of years earn more, not only because of their intangible contributions, but also since the CBA mandates that players who sign multi-year contracts receive annual increases in their pay.

Another important aspect of player employment in the NBA is that prospective players enter the league by participating in an annual draft, in which teams pick the rights to sign them. The order in which a player is selected in the draft is important since the CBA ensures players who are selected high in the draft receive higher wages. On the other hand, players who are not selected in the draft and are signed as free agents have an initial contract that is not restricted by the CBA. So, we control for player draft status and the order the player was selected in the draft.

Finally, controlling for player output and other characteristics, players who enter the NBA directly from high school, that is, without playing college basketball, are likely to earn more. The reason is that such players are considered to have extremely high potential and, because of their

⁶ The standard measure of player performance is the TENDEX statistic, which was first used by the NBA. TENDEX is commonly used by coaches and basketball analysts to evaluate player output. This statistic is a weighted sum of player statistics per game played (points, rebounds, steals, blocks). TENDEX can be computed using different weights for each player statistic and can be broken down to two parts, one that measures offensive and one that measures defensive output. For the purposes of this paper, we break down the statistic into these two parts and use equal weights for all characteristics. Our empirical results are not sensitive to the use of different weights in the calculation of the offensive and defensive TENDEX.

⁷ For example, the CBA mandates that a 10-year veteran must earn at least one million dollars per year.

young age, more years to contribute positively to their teams. The intense demand for such players is likely to produce a wage premium.

There are two groups of job attributes in this context, location amenities and team quality characteristics. A typical list of location amenities is included in the wage equation. First, it is well documented that the size of an urban area has a positive correlation with wages (see for example, Gleaser and Mare, 2001). So, we expect that players in highly populated, high-crime metropolitan areas will have higher wages. Since more than 92 percent of players in the sample were born in the US, playing for a Canadian franchise may be less desirable than playing for a US franchise, thus players may receive higher wages to play in Canada. Finally, players may prefer locations with nice weather conditions (low snowfall, low rainfall, and mild temperatures) and dislike extreme temperature conditions (extremely hot summer or extremely cold winter).

A second group of job characteristics is concerned with the team's success potential. Players in more successful teams create high player popularity and the potential for higher future wages and endorsement deals. In order to capture the team's potential for future success, we construct four dummy variables: *Coach Playoff*, *Coach Ring*, *Winning Record*, and *Champs*. *Coach Playoff* equals one if the coach has previous playoff coaching experience, while *Coach Ring* equals one if the coach has led a team to an NBA championship in the past. At the same time, *Winning Record* equals one if the team had a winning record in the previous regular season, and *Champs* equals one if the team won the championship the year before.

Also, we presume a player is more likely to achieve individual fame and receive endorsement deals if he plays for an NBA franchise that is located in a metropolitan area that it is either within the state of birth of the player or the state in which the player played college basketball. Also, a white (or black) player may have more advertising appeal in a city where a

significant part of its population is white (or black). If so, players would be willing to accept lower wages to play in their birthplace or in the place they played college basketball (*Same Place*) or in metropolitan areas with higher proportions of same-race fan base (*Same Race*).

Estimation results of equation 1, using the linear regression model, are summarized in Table 2. It is noteworthy that the R-squared in all specifications is around .68, whereas typically this statistic is around .3 when we estimate a Mincerian wage equation in other contexts. This supports the idea that the control variables in this context capture to a great extent skill differences across players. Table 2 shows that player characteristics have the expected effect on wages. *Minutes*, *Offense* and *Defense* positively affect wages in all specifications. The estimated coefficients are around .008 for *Minutes*, .014 for *Offense*, and .011 for *Defense*, and are statistically significant at the 1 percent level. These results suggest that a player who plays for 5 percent more minutes than average earns 1 percent higher wages. By the same token, a player earns 1 percentage more than average if he has an offensive output that exceeds the sample mean by 10 percent, or a defensive output that exceeds the sample mean by 20 percent.

Players who enter the league through the draft (*Drafted*) earn 26 percent more than those who enter the league as free agents. In addition, players that are picked higher in the draft clearly earn higher wages. The estimated coefficient for *Draft Number* is around -.007; since the sample mean for *Draft Number* is 16, a player that is picked first in the draft earns 10.5 percent more than does the average player, whereas a player that is picked last in the draft (*Draft Number*=50), earns 20 percent less. Furthermore, players that enter the league directly from high school earn around 9 percent more.

Experienced players or long-tenure players earn a significant wage premium, not only because of their intangible contributions to the team but also because of the CBA provisions. The

coefficients of *Experience* and *Experience Square* are around .078 and -.004 respectively, which is the familiar wage-experience profile found in other contexts. Players who have been with the same team for more than 3 years earn about 8.5 more than their counterparts. Taller players earn higher wages, even after controlling for individual statistics. The estimated coefficient of *Height* is around .006; a player who is 2.10 meters tall earns 5.4 percent more, and a player that is 1.90 meters tall earns 6.2 percent less, than the average player (2.01 meters). Finally, the race or the nationality of a player, are unimportant predictors of player wages.

The estimated effects of location amenities and other team characteristics are reported in columns 2 through 7 of Table 2. Most location characteristics perform very well. Players in highly populated or high-crime metropolitan areas earn significantly higher wages. The estimated elasticity of wages with respect to the metropolitan area's population is .22. For example, Houston and San Antonio feature similar location amenities, but Houston has 71 percent higher population. Controlling for player ability and other team characteristics, players employed by the Houston Rockets will earn on average 15.8 percentage points more than players who are employed by the San Antonio Spurs.⁸ The hedonic price for *Crime* is around .05 and is statistically significant; if a player works in a city that has 1 percent higher crime rate than the sample mean (3.8 percent), he receives a wage premium of 3.9 percent.

Players who play for a Canadian franchise earn a significant wage premium, which is estimated around 20 percent. In addition, the coefficient for *Same Race* is -0.75, implying that a player in a city that has a same race percentage that is 10 percent higher than the sample mean

⁸ Three factors may contribute to the substantial wage premium in large urban areas: (1) NBA teams in larger metropolitan areas have higher potential revenues, thus they pay more, (2) Teams in populated urban areas compensate players for the higher cost of living in that area, and (3) Players dislike playing in heavily populated cities, holding all other amenities constant.

(37 percent) earns 3 percent lower wages. The estimated coefficient for *Same Place* is $-.035$ but lacks statistical significance.

Weather conditions significantly affect wages, providing further support of the theory's validity in this context. Undesirable weather conditions, like *Snowfall* and *Rainfall* produce highly significant positive coefficients, $.075$ and $.076$ respectively. Variables that capture temperature conditions also have the expected effect on wages. Players employed in cities with mild temperature conditions, earn lower wages. Specifically, the hedonic coefficient for *Temperature* is around $-.010$, while the estimated coefficients for *Cold* and *Hot* are around $.21$ and $.25$ respectively.

Using specification 7 in Table 2, we construct the mean predicted wage for each city, evaluated at the weather characteristics of the city and at the sample means of all remaining team and player characteristics. Comparing the city-level mean predicted wage to the sample mean predicted wage, we can evaluate the importance of weather conditions in explaining player wage differences across NBA cities. Table 3 presents the mean predicted wages for each city and its percentage deviation from the predicted wages for the whole sample.

The most desirable locations with respect to weather conditions are Californian cities. It is estimated that players are willing to earn 8 percent less than average to play in Los Angeles, 6.7 percent less to play in Sacramento, and 6.5 percent less to play in Oakland. This is an expected result since Californian cities have mild weather, low precipitation, and no extreme temperature conditions. Similarly, due to desirable weather conditions, wages in Charlotte (-3.9 percent), Phoenix (-3.6 percent), Memphis (-3.5 percent), and Atlanta (-3.4 percent) are lower than average. In contrast, NBA salaries in cities with undesirable weather conditions are significantly higher than average. For example, the wage premium for undesirable weather conditions is 5.4

percent in Minneapolis, 4.6 percent in Toronto, 4.4 percent in Boston, and 4.3 percent in Chicago. These results suggest that, controlling for player skills and other team characteristics, differences in weather conditions cause substantial wage variation across cities, confirming the predictions of the compensating differences theory in this context.

Finally, going back to the results in Table 2, *Coach Ring*, *Winning Record*, and *Champs* do not have significant coefficients. The only relevant team characteristic is *Coach Playoff*; players are willing to earn 2.6 percentage points less to play for a coach with previous playoff experience. Overall, the estimation results lead to two conclusions. First, the compensating differences theory is supported in the context of professional basketball players. Undesirable firm characteristics have a positive effect on wages, while desirable non-pecuniary job attributes have a negative effect. Second, the magnitude of the estimated parameters does not vary significantly across specifications that contain different arrays of job characteristics. This suggests that when we control to a great extent for worker heterogeneity, unobserved employer heterogeneity in terms of working conditions or location amenities does not significantly affect the hedonic estimates for the employer characteristics that are included in the specification.

VI. The Effect of Unobserved Worker Heterogeneity

In this context, unobserved worker heterogeneity is less of an issue compared to microeconomic data used in previous work, thus there is more confidence in the consistency of the estimated hedonic wages. A straightforward way to test whether unobserved heterogeneity may lead researchers to falsely reject the compensating differentials theory is to omit components of worker quality in the current context and observe how it affects the results.⁹

⁹ A similar exercise was performed by Brown (1980). The author showed that by adding variables that capture unobserved worker quality in the specification does not improve support for the compensating differences theory. There are two main differences between that analysis and the analysis in this paper. First, in this context, we have

Table 4 summarizes the estimation results when *Minutes*, *Offense*, *Defense*, and other important components of player heterogeneity are omitted. Most of the estimated compensating differences in Table 4 are quite different from those in column 7 of Table 2 (full specification case) both in terms of size and of statistical significance. The estimated Canada hedonic price in Table 4 is substantially underestimated compared to the full specification case. Specifically, the *Canada* coefficient in Table 4 is between .022 and .096, compared to .199 for the full specification case, while it is only statistically significant in specifications 2 and 5. Only in specification 5 is the coefficient for *Same Race* similar to the full specification case and statistically significant, while in the other four specifications its effect is generally underestimated and lacks statistical significance.

In the absence of important measures of player heterogeneity, the effect of weather conditions on wages is generally understated. In Table 4, the hedonic prices of *Snowfall* and *Rainfall* are substantially underestimated relative to the full specification case and are not statistically significant. In contrast, the three variables that capture temperature conditions remain statistically significant and their signs are according to the theory's predictions. It should be noted that although the *Temperature* parameter is similar to the full specification parameter, the *Cold* and *Hot* hedonics are generally lower than the full specification hedonics.

Same Race does not bear a significant coefficient in specifications 1 through 4 and is only statistically important in specification 5 of Table 4. Moreover, *Coach Playoff* only bears a significant parameter in specification 5. On the other hand, both *Log Pop* and *Crime Rate* have significant positive coefficients in Table 4, and in most specifications these effects are not very different to those in Table 2.

rich measures of worker quality which allow us to estimate more reliable hedonic wages. Second, we find substantial support for the theory in this context, enabling us to better observe how excluding measures of worker quality affects the estimates.

In the full specification case, ten location amenities and team characteristics were found to have a substantial effect on player wages (*Log Pop*, *Crime Rate*, *Canada*, *Same Race*, *Snowfall*, *Rainfall*, *Temperature*, *Cold*, *Hot*, and *Coach Playoff*). In specifications 1 through 4 in Table 4, where important player information is omitted from the specification, no more than three of these characteristics are statistically significant at the 5 percent level. In specification 5, where *Drafted* and *Draft Number* are included in the specification, regression results are much closer to the full specification case. Even in this case, the coefficients for *Snowfall* and *Rainfall* are statistically insignificant, while the coefficients of *Canada*, *Cold*, and *Hot* are underestimated compared to the full specification case.

This analysis shows that estimation results are sensitive to the omission of important measures of worker heterogeneity from the specification. Most hedonic estimates are underestimated, lack statistical significance, or both, in the presence of unobserved player heterogeneity. As a result, if the validity of the theory in this context was judged based solely on a given specification in Table 4, only specification 5 would provide strong support for the theory. Consequently, when important measures of worker skills are not included in the specification, inference on the validity of the compensating differences theory is tainted and may lead to ambiguous results or to the false rejection of the theory.

IV. Alternative Functional Forms of the Wage Equation

Another interesting question is whether the functional form of the wage equation affects the quality of the estimated coefficients, especially if there are unobserved components of worker ability. As discussed, when there is selection on unobserved components of worker ability, choosing the linear regression model may produce significantly biased hedonic estimates.

However, even when worker heterogeneity is accounted for, there are doubts as to whether the relationship between wages and job characteristics is linear. Ekeland et al (2001) note that even though unobserved worker ability is an important issue, restricting the relationship between amenities and wages to be linear may explain the inconsistent empirical evidence. Allowing for flexible functional forms, either by adding non-linear terms in the specification or by utilizing the Box-Cox transformation technique may produce more reliable empirical evidence.

Using the data in hand, we test the sensitivity of the estimates across different functional forms, both in the full specification case and when important measures of worker ability are omitted. To do so, we estimate the wage equation using six alternative models. Besides the classic linear regression model, we use a linear 2nd order model¹⁰, and a partially linear regression model.¹¹ We also use the three variations of the Box-Cox maximum likelihood model; Box-Cox with transformation of the dependent variable only, Box-Cox with double-side transformation using the same parameter, and Box-Cox with double-side transformation using separate parameters. This transformation may improve the plausibility of the assumption that the errors are normally distributed, especially in the case where there is selection on unobserved characteristics.¹² In its most general form, the model is as follows:

$$\frac{y^\lambda - 1}{\lambda} = a + \sum_{\kappa} \alpha_{\kappa} \frac{x_{\kappa}^{\theta} - 1}{\theta} + u \quad (6)$$

Note that y is the dependent variable and x includes the independent variables, while λ is the transformation parameter for the dependent variable and θ for the right hand-side variables.

¹⁰ This model included linear and square terms of continuous player characteristics (*Minutes*, *Offense*, *Defense*, *Draft Number*, and *Height*).

¹¹ This model replaces each continuous measure of player skills (*Minutes*, *Offense*, *Defense*, *Draft Number*, and *Height*) with four indicator functions indicating in which quartile of the each measure's distribution the player is in.

¹² Box-Cox transformations are useful to improve the properties of the sample in cases where the disturbances of the simple linear regression model are problematic. The attractiveness of this transformation is that it allows the linear and log-linear models as special cases, depending on the estimated value of the transformation parameter.

In order to make the results comparable across specifications, the maximum likelihood estimated parameters are transformed to capture the marginal effect of each characteristic on the logarithm of earnings.¹³

Table 5 presents the estimation results for all six models when all measures of player quality are included in the specification. In columns 1 through 3, we report the hedonic estimates for the linear regression models. Results show that the estimated coefficients across the linear models are similar in size and statistical significance. Undesirable location characteristics have a positive effect on wages, whereas desirable team and location amenities are negatively related to player earnings, as predicted by theory.

Columns 4 through 6 report the linearized parameters from the Box-Cox MLE models, showing that the estimated hedonics produced are similar in size and statistical significance across the three Box-Cox models. It is also obvious that the estimated compensating differences do not vary significantly across all six models in Table 5. Therefore, based on these results, we conclude that in the presence of a rich list of measures of worker ability, the functional form of the wage equation does not significantly affect the size and significance of the estimated hedonic coefficients. Although other researchers have suggested that this may be true, to my knowledge, this is the first demonstration that the functional form does not matter when both a significant

¹³ The Box-Cox estimated coefficients represent the marginal effects of the transformed independent variables on the transformed dependent variable. To obtain the marginal effect of each control variable on the untransformed dependent variable, the estimated coefficients must be transformed to reflect the marginal effect of each characteristic on the dependent variable. If \tilde{a} is the estimated parameter, and \hat{a} the linearized coefficient, then the transformation is $\hat{a} = \tilde{a}y^{1-\lambda}$ if only the dependent variable is transformed ($\theta=1$), $\hat{a} = \tilde{a} \left(\frac{\bar{x}}{\bar{y}} \right)^{\lambda-1}$ if we restrict both sides to be transformed by the same parameter ($\lambda=\theta$), and $\hat{a} = \tilde{a} \frac{\bar{x}^{\theta-1}}{\bar{y}^{\lambda-1}}$ if we allow the two sides to be transformed by a separate parameter.

portion of worker heterogeneity is accounted for and where strong evidence exists in support of the theory.

As discussed, there is research work that discusses the merits of using the more flexible Box-Cox transformation technique as a way of reducing the effect of the unobserved skills bias when estimating hedonic wage equations. So, we use the data in hand to examine how different functional forms of the wage equation perform in terms of consistently estimating compensating differences in the absence of measures of worker heterogeneity. Specifically, we estimate the specification in Table 4, column 4 (which omits *Minutes*, *Offense*, *Defense*, *Drafted*, and *Draft Number*) using different models. By comparing these results with the full specification results in Table 5, we can evaluate which models produce consistent estimates in the presence of unobserved heterogeneity.

Table 6 presents the regression results when we omit *Minutes*, *Offense*, *Defense*, *Drafted*, and *Draft Status* under different statistical models.¹⁴ According to the results in Table 6, the partially linear model outperforms the classic linear regression model. The partially linear model produces significant estimates for *Log Pop*, *Crime Rate*, *Canada*, *Same Race*, *Temperature*, *Cold*, *Hot*, and *Coach Playoff*, while the classic linear model does not produce statistically significant parameters for *Canada*, *Same Race*, and *Coach Playoff*. In addition, the sizes of the estimated coefficients of the partially linear model are very close to those obtained in the full specification case. However, compared to the full specification results, the partially linear model underestimates the parameter of *Canada* and overestimates the effect of *Same Race*.

Results also show that all three Box-Cox MLE models perform poorly in the presence of unobserved heterogeneity. These models only produce statistically significant coefficients for

¹⁴ Since all the continuous measures of player skills are excluded from the specification, the second order linear regression model is the same as the classic linear regression model.

crime and extreme temperature conditions. So, in the presence of unobserved heterogeneity, the Box-Cox models produce results that are generally unsupportive of the theory. These results provide little support to the argument that, in the absence of important measures of worker ability, using a more flexible regression model like Box-Cox MLE may improve the consistency of the estimates.

VI. Conclusions

Evidence on the validity of the compensating differentials theory has been inconclusive, many argue because of the unobserved differences in worker skills, the poor observation and measurement of working conditions, and the choice of the functional form of the wage equation. We utilize data on professional basketball players, where rich measures of worker heterogeneity are available, measurement issues are limited, and important working conditions are observable. Using this data, we produce significant evidence supporting the compensating differences theory in this context. We also conclude that, when rich measures of worker heterogeneity are available, different functional forms of the wage equation do not produce dissimilar results.

Using this data, we also find that omitting important measures of player heterogeneity may distort the quality of the estimates. Parameters for important team and location characteristics lose statistical significance or their magnitudes are quite smaller compared to those obtained in the full specification models. Moreover, we find that linear regression models are more effective in estimating hedonics in the presence of unobserved worker heterogeneity compared to the Box-Cox MLE alternatives.

The findings in this paper do not imply that the hedonic wages theory would be validated in other contexts had better measures of worker heterogeneity been available. It rather suggests that, when important components of worker heterogeneity are absent, empirical results are

distorted and may provide wrong inference on the validity of the compensating differentials theory. When the problem of unobserved worker heterogeneity exists, it is difficult both to derive credible conclusions on whether the theory is connected to the data and on the actual magnitude of the effect of job characteristics on wages.

References

- Anglin, M.P. and Gencay, R. (1996). Semi-parametric Estimation of a Hedonic Price Function, *Journal of Applied Econometrics*, 11, 633-648.
- Atrostic, B.K. (1982). The Demand for Leisure and Non-Pecuniary Job Characteristics. *American Economic Review*, 72, 428-440.
- Blomquist, C.G., Berger, M.C., and Hoehn, J.P. (1988). New Estimates of Quality of Life in Urban Areas, *American Economic Review*, 78, 89-107.
- Brown, C. (1989). Equalizing Differences in the Labor Market. *Quarterly Journal of Economics*, 94, 113-134.
- Box, G.E.P. and Cox, D.R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society*, 26, 211-252.
- Camerer, C.F. (1989). Does the Basketball Market Believe in the Hot Hand? *American Economic Review*, 79, 1257-1261.
- Coates, M.L. and Kumar, P. (1982). Occupational Earnings, Compensating Differentials and Human Capital: An Empirical Study. *Canadian Journal of Economics*, 15, 442-457.
- Cropper, L.M, Deck, B.L, McConnell, E.K. (1988). On the Choice of Functional Form for Hedonic Price Functions. *Review of Economics and Statistics*, 70, 255-260.
- Dorman, P., Hagstrom, P. (1998). Wage Compensation for Dangerous Work Revisited. *Industrial and Labor Relations Review*, 52, 116-135.
- Dorsey, S. and Walzer, N. (1983). Workers' Compensation, Job Hazards and Wages. *Industrial and Labor Relations Review*, 36, 642-654.
- Duncan, J.G. (1976). Earnings Functions and Non-pecuniary Benefits. *Journal of Human Resources*, 11, 462-483.
- Duncan, J.G. and Holmlund, B. (1983). Was Adam Smith Right After All? Another Test of the Theory of Compensating Wage Differentials. *Journal of Labor Economics*, 1, 366-379.
- Duncan, J.G. and Stafford, P.F. (1980). Do Union Members Receive Wage Differentials? *American Economic Review*, 70, 355-371.
- Ekeland, I., Heckman, J.J., and Nesheim, L. (2001). Identification and Estimation of Hedonic Models. *CEMMAP Working Paper*, CWP07/02.
- Fishback, P.V. and Kim, S.W. (1993). Institutional Change, Compensating Differentials and Accident Risk in American Railroading 1892-1945. *Journal of Economic History*, 53, 796-823.

- Garen, J.E. (1988). Compensating Wage Differentials and the Endogeneity of Job Riskiness. *Journal of Labor Economics*, 70, 9-16.
- Hausman, J.A. and Leonard, G.K. (1997). Superstars in the National Basketball Association, Economic Value and Policy. *Journal of Labor Economics*, 15, 586-624.
- Hwang, H., Reed, W.R., and Hubbard, C. (1988). Compensating Wage Differentials and Unobserved Productivity. *Journal of Political Economy*, 100, 904-921.
- Kahn, L.M. (2000). The Sports Business as a Labor Market Laboratory. *Journal of Economic Perspectives*, 14, 75-94.
- Kahn, L.M. and Sherer P.D. (1988). Racial Differences in Professional Basketball Players' Compensation. *Journal of Labor Economics*, 6, 40-61.
- Linneman A. (1980). Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market. *Journal of Urban Economics*, 8, 47-68.
- Lucas, R.E.B (1977). Hedonic Wage Equations and Physic Wages in the Returns to Schooling. *American Economic Review*, 67, 549-558.
- McNabb, R. (1989). Compensating Wage Differentials: Some Evidence from Britain. *Oxford Economic Papers*, 41, 327-338.
- Olson, C.A. (1981). An Analysis of Wage Differentials Received by Workers on Dangerous Jobs. *Journal of Human Resources*, 16, 167-185.
- Ramsey J.B. (1969). Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. *Journal of the Royal Statistics Society, Series B*, 31, 350-371.
- Rosen, S. (1986). The Theory of Equalizing Differences. In: *Handbook of Labor Economics*, vol. 1, 641-691, New York: Elsevier.
- Showalter, H.M (1994). A Monte Carlo Investigation of the Box-Cox Model and a Non-Linear Least Squares Alternative. *Review of Economics and Statistics*, 76, 560-570.
- Viscusi, W.K. (1978). Wealth Effects and Earnings Premiums of Job Hazards. *Review of Economics and Statistics*, 60, 408-416.
- Viscusi, W.K and Moore M.J. (1987). Workers' Compensation: Wage Effects, Benefit Inadequacies and the Value of Health Losses. *Review of Economics and Statistics*, 69, 249-261.

Table 1: Variable Description

| Variable Name and Description | Mean (St. Dev.) / Sample Proportion |
|--|-------------------------------------|
| Minutes = Minutes played per game | 24.77 (10.18) |
| Offense = Points per game + Assists per game – Turnovers per game | 10.98 (6.80) |
| Defense = Rebounds per game + Steals per game + Blocks per Game | 5.85 (3.39) |
| Height = Player’s height, in centimeters | 201.22 (9.52) |
| Experience = Number of years player has played in the NBA | 3.52 (3.20) |
| Tenure = 1 if player has been with the same team for more than 3 years, 0 else | 36.1 % |
| High School = 1 if player entered the NBA directly from high school, 0 else | 2.8 % |
| Drafted = 1 if player was drafted, 0 else | 87.9 % |
| Draft no = Draft number, conditional on Drafted=1, 0 else | 16.06 (14.41) |
| Black = 1 if player is black, 0 else | 80.2 % |
| Foreigner = 1 if player is foreigner, 0 else | 7.8 % |
| Snowfall = The average monthly snowfall in inches in team’s city | 1.58 (1.71) |
| Rainfall = The average monthly rainfall in inches in team’s city | 2.87 (1.05) |
| Temperature = The average daily temperature in team’s city | 57.35 (8.23) |
| Hot = 1 if the May–October average temperature is above 80 degrees, 0 else | 52.2 % |
| Cold = 1 if the November–April average temperature is below 40 degrees, 0 else | 18.8 % |
| Log Pop = Logarithm of the city’s population | 13.56 (.92) |
| Crime = Number of all reported crimes per 100 city residents | .038 (.003) |
| Coach Playoff = 1 if coach of the team has previous playoff experience, 0 else | 72.7 % |
| Coach Ring = 1 if coach of the team has won a championship, 0 else | 17.6 % |
| Winning = 1 if the team had a winning record the season before, 0 else | 55.4 % |
| Champs = 1 if the team are the current NBA champions, 0 else | 3.1 % |

Table 2: Estimation Results for the Hedonic Wage Equation

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Minutes | .007*** (.002) | .008*** (.002) | .008*** (.002) | .008*** (.002) | .008*** (.002) | .008*** (.002) | .008*** (.002) |
| Offense | .015*** (.002) | .015*** (.002) | .014*** (.002) | .015*** (.002) | .014*** (.002) | .014*** (.002) | .014*** (.002) |
| Defense | .012** (.005) | .011** (.005) | .011** (.005) | .012** (.005) | .010** (.005) | .011** (.005) | .011** (.005) |
| Experience | .078*** (.008) | .078*** (.008) | .077*** (.008) | .078*** (.008) | .078*** (.008) | .077*** (.008) | .077*** (.008) |
| Experience Square | -.005*** (.001) | -.004*** (.001) | -.004*** (.001) | -.005*** (.001) | -.004*** (.001) | -.004*** (.001) | -.004*** (.001) |
| Tenure | .084*** (.020) | .084*** (.020) | .085*** (.020) | .084*** (.020) | .085*** (.019) | .086*** (.020) | .086*** (.020) |
| Drafted | .262*** (.045) | .260*** (.045) | .266*** (.045) | .265*** (.044) | .269*** (.044) | .267*** (.045) | .267*** (.045) |
| Draft Number | -.007*** (.001) | -.007*** (.001) | -.007*** (.001) | -.007*** (.001) | -.007*** (.001) | -.007*** (.001) | -.007*** (.001) |
| High School | .094*** (.033) | .096*** (.034) | .090*** (.035) | .091*** (.033) | .088*** (.035) | .090*** (.035) | .090*** (.035) |
| Height | .006*** (.002) | .006*** (.002) | .006*** (.002) | .006*** (.002) | .006*** (.002) | .006*** (.002) | .006*** (.002) |
| Black | -.005 (.017) | -.004 (.0176) | -.004 (.024) | -.003 (.017) | -.004 (.024) | -.004 (.024) | -.004 (.024) |
| Foreigner | .004 (.029) | .007 (.029) | .004 (.029) | .003 (.029) | .003 (.029) | .004 (.029) | .004 (.029) |

(continued in next page)

(continued from previous page)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|------------------|--------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Log Pop | -- | .225*** (.053) | .234*** (.059) | -- | .237*** (.056) | .219*** (.060) | .223*** (.057) |
| Crime Rate | -- | .050** (.020) | .050** (.020) | -- | .049** (.019) | .050** (.020) | .048** (.020) |
| Canada | -- | .210** (.090) | .200** (.096) | -- | .206** (.099) | .199** (.098) | .199** (.098) |
| Same Race | -- | -- | -.078** (.033) | -- | -.078** (.033) | -.075** (.033) | -.075** (.033) |
| Same Place | -- | -- | -.035 (.030) | -- | -.035 (.029) | -.035 (.030) | -.035 (.029) |
| Snowfall | -- | -- | -- | -- | -- | .075** (.039) | .076** (.030) |
| Rainfall | -- | -- | -- | -- | -- | .077*** (.025) | .076*** (.023) |
| Temperature | -- | -- | -- | -- | -- | -.010** (.005) | -.010** (.005) |
| Cold | -- | -- | -- | -- | -- | .209*** (.050) | .211*** (.048) |
| Hot | -- | -- | -- | -- | -- | .246*** (.075) | .248*** (.074) |
| Coach Playoff | -- | -- | -- | -.027** (.013) | -.026** (.012) | -- | -.026** (.012) |
| Coach Ring | -- | -- | -- | .035 (.030) | .037 (.030) | -- | .037 (.030) |
| Winning | -- | -- | -- | .026 (.023) | .025 (.022) | -- | .024 (.022) |
| Champs | -- | -- | -- | -.018 (.027) | -.019 (.029) | -- | -.019 (.029) |
| R-squared | 0.6801 | 0.6822 | 0.6834 | 0.6822 | 0.6855 | 0.6835 | 0.6856 |
| Observations | 2,065 | 2,065 | 2,065 | 2,065 | 2,065 | 2,065 | 2,065 |

Notes: Dependent variable is the logarithm of wages (mean: 14.64, standard deviation: 1.03). Year Fixed Effects and intercept included in the specification but not reported. Standard errors are clustered by team and reported in parenthesis. * = significant at the 10% level, ** = significant at the 5% level, *** = significant at the 1% level.

Table 3: Weather Conditions and Wage Differences across NBA Cities

| | Mean Predicted Wages | Percentage Deviation from Sample Mean due to Weather Hedonics |
|--------------------|----------------------|---|
| Total Sample | \$2,279,120 | -- |
| Los Angeles, CA | \$2,097,069 | -8.0 % |
| Sacramento, CA | \$2,126,790 | -6.7 % |
| Oakland, CA | \$2,130,142 | -6.5 % |
| Charlotte, NC | \$2,191,335 | -3.9 % |
| Phoenix, AZ | \$2,197,627 | -3.6 % |
| Memphis, TN | \$2,200,526 | -3.5 % |
| Atlanta, GA | \$2,202,572 | -3.4 % |
| Indianapolis, IN | \$2,267,106 | -0.5 % |
| Portland, OR | \$2,274,174 | -0.2 % |
| San Antonio, TX | \$2,277,674 | -0.1 % |
| Miami, FL | \$2,279,562 | +0.0 % |
| Philadelphia, PA | \$2,318,103 | +0.0 % |
| Seattle, WA | \$2,284,902 | +0.3 % |
| Dallas, TX | \$2,293,316 | +0.6 % |
| Orlando, FL | \$2,298,392 | +0.9 % |
| Houston, TX | \$2,306,366 | +1.2 % |
| Washington, DC | \$2,356,864 | +2.6 % |
| New York City | \$2,339,819 | +2.7 % |
| Newark, NJ | \$2,340,315 | +2.7 % |
| Denver, CO | \$2,359,143 | +3.5 % |
| Milwaukee, WI | \$2,360,100 | +3.6 % |
| Cleveland, OH | \$2,363,131 | +3.7 % |
| Detroit, MI | \$2,365,073 | +3.8 % |
| Salt Lake City, UT | \$2,356,677 | +3.4 % |
| Chicago, IL | \$2,376,095 | +4.3 % |
| Boston, MA | \$2,378,534 | +4.4 % |
| Toronto, Canada | \$2,384,264 | +4.6 % |
| Minneapolis, MN | \$2,403,094 | +5.4 % |

Note: Mean predicted wages are calculated as described in text.

Table 4: Estimated Compensating Differences with Unobserved Player Heterogeneity

| | (1) | (2) | (3) | (4) | (5) |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Experience | .147*** (.008) | .115*** (.010) | .151*** (.007) | .120*** (.009) | .115*** (.008) |
| Experience Square | -.008*** (.001) | -.006*** (.001) | -.008*** (.001) | -.007*** (.001) | -.007*** (.001) |
| Tenure | -- | .207*** (.023) | -- | .198*** (.021) | .158*** (.022) |
| Drafted | -- | -- | -- | -- | .515*** (.043) |
| Draft Number | -- | -- | -- | -- | -.013*** (.001) |
| High School | -- | -- | .264*** (.078) | .245*** (.071) | .142** (.065) |
| Height | -- | -- | .008*** (.001) | .008*** (.001) | .004*** (.001) |
| Black | .058 (.045) | .032 (.044) | .087** (.041) | .062 (.040) | -.001 (.037) |
| Foreigner | .168*** (.051) | .148*** (.050) | .094** (.046) | .077* (.044) | .072* (.036) |
| Log Pop | .229*** (.068) | .230*** (.070) | .227*** (.071) | .229*** (.073) | .224*** (.068) |
| Crime Rate | .042* (.021) | .043* (.021) | .044* (.022) | .046* (.024) | .047* (.025) |
| Canada | .022 (.100) | .052 (.103) | .074* (.041) | .058 (.044) | .096*** (.029) |
| Same Race | -.046 (.061) | -.074 (.056) | -.027 (.055) | -.055 (.049) | -.079* (.037) |
| Same Place | -.033 (.054) | -.030 (.047) | -.019 (.055) | -.055 (.048) | -.044 (.037) |
| Snowfall | .062 (.043) | .031 (.044) | .035 (.028) | .007 (.029) | -.011 (.021) |
| Rainfall | .048 (.041) | .026 (.041) | .025 (.031) | .005 (.031) | -.027 (.024) |
| Temperature | -.011* (.006) | -.011* (.005) | -.010* (.005) | -.010* (.005) | -.010** (.005) |
| Cold | .189*** (.048) | .198*** (.050) | .168*** (.025) | .205*** (.027) | .206*** (.021) |
| Hot | .232** (.121) | .239** (.124) | .225** (.111) | .230** (.111) | .219** (.095) |
| Coach Playoff | -.013 (.018) | -.014 (.020) | -.019 (.017) | -.019 (.018) | -.028** (.013) |
| Coach Ring | .021 (.041) | .009 (.040) | .026 (.038) | .014 (.036) | .031 (.029) |
| Winning Record | -.009 (.024) | -.016 (.024) | -.009 (.022) | -.016 (.022) | .011 (.022) |
| Champs | .025 (.039) | .005 (.047) | .018 (.039) | -.001 (.047) | -.026 (.047) |
| R-squared | .3472 | .3845 | .3865 | .4204 | .5774 |

Notes: *Minutes*, *Defense*, and *Offense* are omitted from all specifications.

Table 5: Estimation of the Wage Equation using Alternative Models

| | LINEAR REGRESSION MODELS | | | BOX-COX MLE MODELS ^a | | |
|----------------------------|--------------------------|------------------------|----------------------------|---------------------------------|---------------------------------|-----------------------------------|
| | Classic (1) | Second Order (2) | Partially Linear (3) | Only LHS (4) | Both Sides Restricted (5) | Both Sides Unrestricted (6) |
| Log Pop | .223*** (.057) | .187*** (.050) | .217*** (.063) | .223** (.029) | .245** (.035) | .220** (.031) |
| Crime Rate | .048** (.020) | .043** (.019) | .049** (.021) | .054*** (.003) | .025*** (.002) | .057*** (.003) |
| Canada | .199** (.098) | .155* (.087) | .206* (.109) | .194** (.046) | .182** (.047) | .193** (.048) |
| Same Race | -.075** (.033) | -.056** (.026) | -.095** (.043) | -.073** (.034) | -.068* (.052) | -.073** (.045) |
| Same Place | -.035 (.029) | -.032 (.027) | -.035 (.030) | -.035* (.054) | -.035* (.061) | -.036* (.055) |
| Snowfall | .076** (.030) | .066** (.026) | .072** (.032) | .073* (.051) | .087** (.030) | .070** (.047) |
| Rainfall | .076*** (.023) | .063*** (.022) | .071*** (.023) | .075** (.039) | .128* (.052) | .070** (.048) |
| Temperature | -.010** (.005) | -.009* (.005) | -.010** (.005) | -.010** (.049) | -.009** (.045) | -.010** (.047) |
| Cold | .211*** (.048) | .178*** (.045) | .224*** (.052) | .213** (.019) | .198** (.030) | .212** (.024) |
| Hot | .248*** (.074) | .232** (.104) | .216** (.105) | .232*** (.001) | .288*** (.004) | .229*** (.003) |
| Coach Playoff | -.026** (.012) | -.025** (.012) | -.027** (.012) | -.026** (.047) | -.026** (.048) | -.026** (.048) |
| Coach Ring | .037 (.030) | .039 (.028) | .041 (.031) | .038 (.175) | .037 (.185) | .038 (.018) |
| Winning Record | .024 (.022) | .031 (.021) | .023 (.033) | .024 (.202) | .020 (.185) | .025 (.203) |
| Champs | -.019 (.029) | -.013 (.027) | -.016 (.033) | -.025 (.497) | -.029 (.445) | -.025 (.044) |
| R-squared | .6856 | .6934 | .6721 | -- | -- | -- |
| LR test ^b | 2.065 | 1.4455 | 7.228 | -- | -- | -- |
| Ramsey F-test ^c | 1.73 (.183) | 0.90 (.454) | 2.60 (.080) | -- | -- | -- |
| Log Likelihood | -- | -- | -- | -41.543 | -45.387 | -41.445 |
| Observations | 2,065 | 2,065 | 2,065 | 2,065 | 2,065 | 2,065 |

a= for comparison purposes, the Box-Cox coefficients are linearized, similar to Linneman (1980) and Blomquist et al (1988). The optimal transformation coefficients are 2.660 for specification 4, 2.811 for specification 5, and 2.643 (0.838 for the control variables) in specification 6. All transformation parameters are statistically significant at the 1% level.

b= The LR test statistic is distributed χ^2_3 (Critical values: 9.49 at the 5% level, 13.28 at the 1% level). Null Hypothesis: No higher order terms are omitted in specification.

c= The F-statistic and p-value for the RESET misspecification test (Ramsey, 1969) are reported. The null hypothesis is that there is no omission of higher order or interaction terms in the specification. This statistic is derived by taking the fitted values from the model being tested and producing higher order terms of its fitted values. These terms are included in the base model and a standard F-test is performed to determine whether they are jointly significantly different from zero.

Table 6: Estimation of the Wage Equation with Unobserved Player Heterogeneity

| | LINEAR REGRESSION MODELS | | BOX-COX MLE MODELS | | |
|----------------|--------------------------|----------------------|--------------------|---------------------------|-----------------------------|
| | Classic (1) | Partially Linear (2) | Only LHS (3) | Both Sides Restricted (4) | Both Sides Unrestricted (5) |
| Log Pop | .229*** (.073) | .175*** (.021) | .182 (.191) | .271 (.198) | .262 (.192) |
| Crime Rate | .046* (.024) | .047* (.028) | .084 (.001) | .026*** (.000) | .046** (.032) |
| Canada | .058 (.044) | .079*** (.025) | .065 (.745) | .063 (.752) | .059 (.750) |
| Same Race | -.055 (.049) | -.113** (.046) | -.059 (.211) | -.058 (.219) | -.055 (.217) |
| Same Place | -.055 (.048) | -.036 (.037) | -.018 (.472) | -.018 (.481) | -.017 (.477) |
| Snowfall | .007 (.029) | -.023 (.021) | .005 (.953) | .025 (.799) | .089 (.801) |
| Rainfall | .005 (.031) | -.035 (.024) | .004 (.963) | .130 (.744) | .142 (.837) |
| Temperature | -.010* (.005) | -.011** (.005) | -.020 (.243) | -.010 (.237) | -.013 (.234) |
| Cold | .205*** (.027) | .210*** (.021) | .225* (.069) | .212* (.086) | .156* (.087) |
| Hot | .230** (.111) | .277*** (.031) | .207*** (.006) | .198*** (.008) | .220*** (.007) |
| Coach Playoff | -.019 (.018) | -.030* (.015) | -.020 (.331) | -.020 (.335) | -.019 (.332) |
| Coach Ring | .014 (.036) | .030 (.027) | .017 (.571) | .017 (.559) | .019 (.543) |
| Winning Record | -.016 (.022) | .011 (.022) | -.014 (.485) | -.014 (.476) | -.014 (.482) |
| Champs | -.001 (.047) | -.017 (.038) | -.001 (.904) | -.006 (.914) | -.004 (.909) |
| R-Squared | .4204 | .5676 | -- | -- | -- |
| Log Likelihood | -- | -- | -680.515 | -677.584 | -674.748 |
| LR Test | 5.369 | 3.924 | -- | -- | -- |
| Ramsey F-Test | 0.91 (0.451) | 0.83 (0.487) | -- | -- | -- |

Notes: Dependent variable is the logarithm of wages (2065 observations). Not included in the specifications are *Minutes*, *Offense*, *Defense*, *Draft Number*, and *Drafted*. For a description of the statistical models and the specification tests, see notes of Table 5. Estimated coefficients for personal characteristics are not reported. In parenthesis, are reported the clustered standard errors by team in columns (1)-(3) and the p-values of the LR test in columns (4)-(6). *= significant at the 10% level, **= significant at the 5% level, ***=significant at the 1% level. The optimal transformation coefficients for the dependent variable are 2.672 for specification 3, 2.785 for specification 4, and 7.068 (2.736 for the control variables) in specification 6.