RAE REVIEW OF APPLIED ECONOMICS
Vol. 9, Nos. 1-2, (January-December 2013)

SPATIAL ECONOMETRICS IN NON-SPATIAL SETTINGS

*Kevin E. Henrickson** and Ryan Langrill****

ABSTRACT: *In this paper, we use data from Major League Baseball between 2000 and 2005 to illustrate how spatial econometric techniques can be used to directly model the interdependence of players. Specifically, we empirically examine the role of task specific human capital on earnings, by modeling player contracts as being determined by their performance, and the performance of other players with similar task specific human capital profiles. Our approach also provides an example to other researchers of a methodological approach to using spatial econometrics in settings where space is defined in ways other than spatial geography.*

JEL Classifications: *J31, J41, J24*

Keywords: *wages, contracting, human capital, spatial econometrics*

1. INTRODUCTION

Spatial econometric techniques have greatly enhanced the way that economists model interaction between countries and/or firms. Indeed, the spatial autoregressive model has become relatively common in applied work, (e.g. Kalnins (2003), Blonigen *et al.* (2007), McMillen *et al.* (2007) and Henrickson (2012)). However, nearly all of the spatial econometric literature measures space as geographic, largely ignoring non-geographic interdependence between economic entities. For example, foreign direct investment (FDI) could be viewed as flowing into countries with similar laws/cultures/regulations, rather than countries sharing borders. Likewise, changes in stock prices are likely influenced by changes in the stock price of firms in the same industry, while firm profits are impacted by sales of competing firms. In this paper, we illustrate the use of spatial econometrics to capture this type of non-geographic interaction, by modeling workers' wages as determined, in part, by those of workers with similar task specific human capital, i.e. workers who are spatially "close" in ability.

Task specific human capital, as described by Shaw (1984), Topel (1991), Neal (1995), and Gibbons and Waldman (2004), has not been extensively examined in labor markets due to a

---

* Department of Economics, Gonzaga University, 502 E Boone Ave, Spokane, WA 99258, U. S., *E-mail: henrickson@gonzaga.edu*

** Department of Economics, George Mason University, Enterprise Hall, 3G4, Fairfax, VA 22030, U. S., *E-mail: rlangril@gmu.edu*

lack of available data to empirically model worker differences. Instead, the literature has tended to follow the seminal works of Mincer (1958) and Becker (1964) by focusing primarily on general human capital, typically measured by education, and/or firm-specific human capital, typically measured by tenure with a firm, and the impact of these types of human capital on labor market outcomes (e.g. Altonji and Shakotko (1987); Topel (1991); Mortensen (2003); Gori and Fanti (2008); Becker and Woessmann (2009); Kambourov and Manovskii (2009); and Burdett *et al.* (2011)). However, task specific human capital is equally important in determining many of these outcomes. For example, a computer programmer with knowledge of a certain programming language will have their wages determined, in part, by the availability and wages of other programmers with knowledge of the same language. While this type of interdependence of worker wages is prevalent in many markets, empirical modeling this type of wage relationship has been difficult.

In this study, we take advantage of the availability of task specific human capital measures within Major League Baseball to assess the importance of empirically incorporating such task specific human capital into labor market outcomes. Similar previous work by Blass (1992) examines human capital accumulation within Major League Baseball, finding that older players are overpaid relative to their productivity gains. However, we illustrate the use of spatial econometric techniques to model this interdependence of outcomes, which has implications across all areas of economics. Specifically, we model baseball players as having task specific human capital that they accumulate specific to their position, and also general human capital that is not position specific. This approach of identifying task specific human capital by position is similar to Treme and Allen (2009) who examine outcomes for wide receivers in the National Football League. By isolating position in this way, Treme and Allen (2009) are focusing on players with similar task specific human capital, but this interdependence is not directly modeled such as what we do here. In particular, task specific human capital accumulation implies that a player's salary is not only based on his previous performance and general human capital, but also on the labor market conditions for players with similar task-specific human capital profiles, those 'close' to them. We account for this labor market feature through use of a spatial autoregressive model where space is defined as differences in task specific human capital levels.

Our findings suggest that task specific human capital accumulation does impact employee compensation, and that accounting for this human capital improves wage predictions in Major League Baseball. More importantly, we illustrate that spatial econometric methodology can be adapted to non-geographic spatial interdependence, allowing economists to better model interactions between economic entities, such as the interaction of stock prices across companies, FDI across countries, wages across workers, and profits across groups, all examples of situations in which spatial econometrics can be used to model non-geographic interdependence.

2. EMPIRICAL MODEL & DATA

The data that we use in this study include salary data for contracts signed between 2000 and 2006 for each Major League Baseball player, collected from USA Today's Salary Database, and performance measures for the year before the contract was signed. The dependent variable is the real average annual contract value for each player, which is defined as the total contract value divided by the contract length, and is specified in hundreds of thousands of dollars. This specification, using average contract value, is used because contracts are often either front or back loaded, meaning that the actual salary changes with each year of the contract. The independent variables used consist of performance measures for each player, non-performance player preferences and the characteristics of the task specific labor market in which the player is supplying their labor. Since pitchers and hitters within major league baseball are evaluated using different performance measures, we estimate wage regressions separately for these two groups.

We include the following non-spatial variables for both pitchers and position players: a dummy variable for *left handed* players, the *age* of the player at the time they sign the contract, *age squared,* to capture any non-linearity in the impact of age on wages, and *experience,* measured as the number of years the player has appeared in a Major League baseball game.

For position players, we include the following performance measures, measured for the season before signing the contract: *runs created per game*, *homeruns per game*, *intentional walks per game*, *stolen bases per game*, *strike outs per at bat* and *fielding percentage*. We also include dummy variables for winners of the *Silver Slugger Award* and the *Gold Glove Award*, which are awarded to one player per league per position, as well as a dummy variable for players who made the *All Star* team, and the number of *wins* a player's team had. We do note that our results presented in what follows are robust to including career performance measures instead of previous year measures; however, inclusion of such measures is found statistically unwarranted as players appear to be judged primarily on their previous year's performance, which is similar to the results found by Healy (2008). Also, we note that runs created is a performance measure developed to account for the number of runs a player is responsible for, and is defined by James (2002).

For pitchers, the past performance measures are: *winning percentage,* defined as the number of games the pitcher won divided by the number of decisions that he received, the number of *games started, complete games*, *shutouts, saves*, his *strikeout to walk ratio*, the log of *earned run average*, and a dummy variable being named to the *All Star* team. We note that the log of earned run average is used because it is assumed that there is a bigger difference between pitchers who have an earned run average of 3.00 and 4.00 than there is between pitchers with earned run averages of 5.00 and 6.00. Also, we do not include team wins for pitchers, as we assume that pitchers have less control over the total number of games that their team wins since they play in fewer games than position players.

We also include several non-performance variables related to their preferences to sign with a certain team, including: a dummy variable if the contract is with a player's *previous team*, a dummy variable for a player's *hometown* team, defined as teams in states first degree contiguous to a player's birth state, and a dummy variable for players who are still under *team control*, which occurs during their first six years of Major League service, when they are not true free agents, but are still technically property of their current team. Table 1 contains summary statistics for each of these variables.

To capture the potential for task specific labor markets within Major League Baseball, we use a spatial autoregressive (lag) model, given by:

$$\text{Average Contract Value} = X\beta + \rho W \text{Average Contract Value} + \varepsilon \tag{1}$$

Table 1
Summary Statistics

Batters		*Pitchers*	
Variable	*Mean*	*Variable*	*Mean*
Average Contract Value	$25.17	Average Contract Value	$24.86
	(33.98)		(29.92)
Runs Created	55.75	Winning Percentage	0.50
	(33.92)		(0.21)
Homeruns	12.36	Games Started	13.88
	(10.44)		(13.97)
Intentional Walks	2.91	Complete Games	0.659
	(3.77)		(1.34)
Stolen Bases	6.95	Shutouts	0.25
	(9.92)		(0.63)
Silver Slugger	0.04	Saves	3.83
	(0.19)		(10.19)
Strike Outs per at Bat	0.19	Strike Outs per Walk	2.31
	(0.07)		(1.16)
Gold Glove	0.03	Earned Run Average	4.01
	(0.16)		(1.59)
All Star	0.09	All Star	0.074
	(0.29)		(0.26)
Left Handed	0.28	Left Handed	0.27
	(0.45)		(0.44)
Previous Team	0.57	Previous Team	0.55
	(0.50)		(0.50)
Home Team	0.11	Home Team	0.12
	(0.31)		(0.33)
Club Control	0.60	Club Control	0.60
	(0.49)		(0.49)
Age	28.23	Age	28.41
	(3.91)		(4.40)
Experience	4.88	Experience	4.56
	(3.83)		(3.85)
Fielding Percentage	0.86		
	(0.19)		
Team Wins	75.23		
	(24.63)		

Notes: Average contract value is defined is the total amount of the contract divided by the contract length, and is specified in hundreds of thousands of dollars. We also note that while all players have general characteristics which they are all evaluated on (e.g. experience, age, etc.), there are inherent differences in batters and pitchers. As such, we model wages separately for pitchers and batters, with each model incorporating the same general characteristics of the players, but with batter/pitcher specific variables used depending on the player's position. Our batter specific variables, shown above, include: runs created, home runs, intentional walks, stolen bases, strikeouts, team wins and whether the player won either the Gold Glove Award or Silver Slugger Award for their position. Similarly, for pitchers we include: winning percentage, games started, complete games, shutouts, saves, strikeouts per walk and the log of the pitcher's earned run average (ERA).
Standard deviations are in parentheses.

where X is the vector of player characteristics, player performance measures, and non-performance factors discussed above. Appendix A provides a more detailed account of the spatial autoregressive model, but the model given by equation (1) above is a typical OLS specification with an additional term to capture the spatial nature of task specific human capital on salary, ρ*WAverage Contract Value*.

Specifically, W is a block diagonal spatial weighting matrix. Each diagonal block of this weighting matrix accounts for one year, with the off diagonal elements within the block, $\gamma_{i,j}$, accounting for the skill proximity between players i and j, and the on diagonal elements set to zero, to prevent a player's salary from being regressed on itself. The specification of $\gamma_{i,j}$ is intended to capture the labor markets for task specific human capital, and as such $\gamma_{i,j}$ is set to zero for any players who play different positions as their human capital is associated with different tasks. The two exceptions to this rule are right fielders and left fielders, who are considered to play the same position, corner outfield, and utility players, who are defined as players who did not spend at least 75% of their time at one position. For players who do have the same task specific human capital, $\gamma_{i,j}$ is defined as the absolute value of the difference between player i's value over a replacement level player (VORP) and player j's VORP, where a player's VORP is a statistic calculated by Baseball Prospectus designed to capture each player's value by comparing their performance to the performance that could be expected from the 'average' player at each position.

As with all spatial autoregressive models, the specification of the weighting matrix is chosen based on one's belief about the relationship between spatially related observations. In the present context, 'space' is defined based on the proximity of players' human capital rather than geographic. As such, it is assumed that each player's salary is based, in part, on the salary of individuals with similar task specific human capital, but that this relationship decreases as players' performance based ratings diverge. However, as is common in the spatial econometrics literature, the ad-hoc nature of the weighting matrix was examined by using numerous alternative specifications of the weighting matrix, including equal weights for all players at the same position, and weights which allow a player's salary to be a function of the average of players at several positions which require similar task specific human capital were examined. However, we note that the results of these robustness checks are qualitatively similar to the results presented here. It is worth noting that one could create a weighting matrix taking into account other player characteristics such as where the batter hits in the batting order. We do not include this type of analysis in our weighting matrix as teams will often have different strategies in hiring players, but all teams need to field a team with one player at each position, hence our focus on position/task specific human capital in our weighting matrix.

As is common in the spatial econometric literature, the spatial weighting matrix, W, is row standardized and estimated via maximum likelihood. Row standardizing means that *WAverage Contract Value* can be interpreted as the weighted average contract value of players with the same task specific human capital, and ρ, our estimated term in equation (1), can be interpreted as estimating the effect of the weighted average contract value of players with similar task specific human capital profiles on salary. This structure implies that a positive estimate of ρ indicates that the labor market in Major League Baseball is dependent upon on task specific human capital, as players with similar human capital profiles' salaries are positively correlated.

3. RESULTS

The results of estimating equation (1) by both OLS and maximum likelihood are presented in Table 2 for batters and Table 3 for pitchers. The OLS results of estimating equation (1) without the endogenous spatially weighted average contract value are included in these tables to assess the stability of the parameter estimates; however, all tests for the appropriateness of the spatial autoregressive model indicate that wages are spatially correlated by task specific human capital.

(a) Player Characteristics & Past Performance

There are four variables included as explanatory variables in equation (1) that are intended to capture player characteristics. The first of these, the dummy variable for *left handed* pitchers and batters is insignificant across all four specifications, indicating that players do not receive more or less pay due to being left handed. The other three player characteristic variables, however, are all statistically significant, indicating that players receive higher pay as both their *experience* and *age* increase, while the *age squared* variable indicates that the effect of age on salary peaks and then reverses. We note that this model was also estimated with experience squared, but that the estimated coefficient was insignificant, likely because of the inclusion of age squared. The positive coefficient of the *experience* variable is consistent with our *a priori* expectations, and indicates the existence of human capital accumulation, with one year of additional experience accounting for a $173,000 increase in annual salary for batters and a $362,000 increase for pitchers.

For batters, the variable intended to capture the bulk of each players previous year's performance is *runs created per game*, which is positive and significant across both specifications, with each additional run created per game increasing annual salary by $5,031,300. In addition to the *runs created per game,* we also have several other statistical measures of performance to capture specific skills not adequately captured by *runs created per game* in the determination of wages. The first of these performance measures, the number of *homeruns per game* the player hit in the previous season, is positive and statistically significant indicating that each additional *homerun per game* amounts to an estimated salary increase of $6,558,500. This result implies that either home runs' offensive contribution is underestimated by *runs created per game*, or that teams value homeruns more than for their pure offensive value, perhaps because of marketing or fan interest considerations.

Intentional walks per game are included as an explanatory variable in our estimation not because of the productive impact of an intentional walk, but because intentional walks capture opposing teams' perception of the batter's ability. In particular, a pitcher and/or manager will only intentionally walk a batter if they fear the batter's ability so much that they will concede first base to the hitter rather than attempt to get that hitter out; thus, only the elite superstar players should receive more than an occasional intentional walk. Indeed, the results in Table 3 illustrate this impact with each additional *intentional walk per game* in the previous season accounting for an estimated $14,651,500 increase in a player's average contract value. *Stolen Bases per game* is included because, in addition to the pure offensive importance of stealing a base, *stolen bases* also offer an imperfect measure of speed, which will impact the player's ability to score runs, defend in the field and impact the course of the game. Our estimates support this idea that stolen bases are valuable beyond what *runs created per game* measures,

Table 2
OLS and Spatial Lag Wage Regression Results for MLB Batters

Variable	*OLS*	*Spatial Lag*
All Star	11.019***	10.206***
	(3.200)	(3.167)
Runs Created per Game	53.964***	50.313***
	(7.885)	(7.891)
Home Runs per Game	68.725***	65.585***
	(18.840)	(18.592)
Intentional Walks per Game	144.731***	146.515***
	(30.665)	(30.206)
Stolen Bases per Game	35.093***	32.186***
	(12.646)	(12.503)
Strike Outs per At Bat	-22.736*	-22.767*
	(12.240)	(12.054)
Silver Slugger Award	23.125***	23.610***
	(4.908)	(4.837)
Gold Glove Award	24.927***	24.667***
	(4.710)	(4.640)
Fielding Percentage	0.410	0.750
	(4.163)	(4.101)
Team Wins	0.053*	0.052*
	(0.031)	(0.031)
Previous Team	-1.065	-1.167
	(1.610)	(1.586)
Home Team	4.945**	4.835**
	(2.358)	(2.323)
Club Control	-18.029***	-18.011***
	(2.397)	(2.361)
Left Handed	-0.821	-1.019
	(1.710)	(1.685)
Experience	1.736***	1.730***
	(0.479)	(0.472)
Age	10.105***	10.000***
	(1.851)	(1.823)
Age Squared	-0.194***	-0.192***
	(0.031)	(0.031)
Constant	-140.144***	-138.730***
	(29.060)	(28.623)
Average Salary of Batters with the Same Task Specific Human Capital		0.096***
		(0.037)
Observations	875	875
Adjusted R^2	.61	

Notes: We model player wages separately for pitchers and batters. In addition to the spatial lag term, we include, for all players: experience, age, age squared, left-handedness, whether they were an All-Star and whether they signed with their previous team, hometown team, or were still under club control. We also include batter specific variables including: runs created, home runs, intentional walks, stolen bases, strikeouts, team wins and whether the player won either the Gold Glove Award or Silver Slugger Award for their position. Standard errors are in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

with results indicating that an additional *stolen base per game* is worth an extra $3,218,600 per year. Interestingly, while *stolen bases per game*, our indirect measure of defensive ability through speed is statistically significant, our direct measure of defensive ability, *fielding percentage*, is statistically insignificant, a result that is likely due to the relatively little variation that this variable offers.

Another statistic not fully captured in *runs created per game* is strikeouts. A strikeout is often seen as a more damaging out than either a groundout or a flyout, because it offers no opportunity for a base runner to advance. It also offers no chance of forcing a fielder to commit an error, another way a batter can reach base. Our measure of strikeouts, *strikeouts per at bat*, is used to normalize strikeouts across playing time. According to Table 2, one extra strikeout for every ten at bats results in a $2,276,700 reduction in the player's average annual contract value.

In order to adequately capture the pay that superstars in the league receive, we included several awards that are bestowed upon players who excel in particular areas. We also note that earlier results included additional award categories, but because most other awards only go to one player per year, there is almost no variation for many of the other awards, hence our focus on only those awards which are given to multiple players every year. The first of these honors is being selected to play in the annual *All Star* game, which is bestowed on the best players at each position in each the National League and the American League, as well as a number of reserve players chosen from the game's elite. Being named to the *All Star* team in the year prior to signing a contract is estimated to increase a player's annual contract value by $1,020,600. Other annual awards include the *Silver Slugger* Award, which is given to the best hitter at each position in each league, and the *Gold Glove* Award, which is voted on by major league coaches to determine the best fielder at each position in each league. These two awards are estimated to increase a player's average annual contract value by $2,361,000 and $2,466,700, respectively. Examining these results, we note that, as was discussed previously, *fielding percentage* is statistically insignificant; however, the result on winning the *Gold Glove Award*, which is tied to fielding, captures the increased salary associated with being the best fielder at one's position.

Our final measure of past performance for batters is their team's win total from the previous year. This is included because batters, who play every day, have greater opportunity to impact the outcome of every game. Thus, a player can be known as a "winner" for doing all of the little things to help his team win, a skill that our included performance measures may not adequately capture. Indeed, our estimates indicate that for each extra game that a player's team wins in the year before he signs a new contract, the player receives $5,200 in extra compensation.

For pitchers, we include fewer performance measures as there are fewer tasks associated with pitching. In particular, we exclude *team wins* because a pitcher plays in a smaller portion of a team's games. We also excluded the *Silver Slugger* and *Gold Glove* awards as a pitcher's batting and fielding performances are secondary to his pitching.

Of the variables used to estimate the impact of pitcher performance on average contract value, shown in Table 3, one of the primary measures of pitcher quality is their *winning percentage*. This variable, measured as the number of games that the pitcher won divided by his decisions (wins plus losses), indicates that a 10% increase in *winning percentage* leads to an estimated increase in average annual contract value of $60,380. However, this measure is applied

Table 3
OLS and Spatial Lag Wage Regression Results for MLB Pitchers

Variable	*OLS*	*Spatial Lag*
All Star	8.019**	5.780*
	(3.309)	(3.118)
Winning Percentage	6.162*	6.038*
	(3.662)	(3.490)
Games Started	0.819***	0.576***
	(0.075)	(0.078)
Complete Games	3.549***	2.953***
	(0.904)	(0.851)
Shutouts	-2.484	-2.043
	(1.769)	(1.661)
Saves	0.517***	0.292***
	(0.092)	(0.091)
Strike Outs per Walk	3.087***	3.072***
	(0.725)	(0.688)
Log Earned Run Average	-8.144***	-5.842***
	(2.244)	(2.265)
Previous Team	1.207	0.677
	(1.610)	(1.521)
Home Team	5.389**	4.055*
	(2.316)	(2.188)
Club Control	-9.766***	-8.779***
	(2.343)	(2.231)
Left Handed	-0.002	-0.112
	(1.731)	(1.629)
Experience	3.351***	3.620***
	(0.479)	(0.474)
Age	6.816***	7.205***
	(1.895)	(1.789)
Age Squared	-0.126***	-0.136***
	(0.032)	(0.031)
Constant	-90.228***	-100.442***
	(29.131)	(27.436)
Average Salary of Pitchers with the Same Task Specific Human Capital		0.313***
		(0.042)
Observations	734	734
Adjusted R^2	.55	

Notes: We model player wages separately for pitchers and batters. In addition to the spatial lag term, we include, for all players: experience, age, age squared, left-handedness, whether they were an All-Star and whether they signed with their previous team, hometown team, or were still under club control. We also include batter specific variables including: winning percentage, games started, complete games, shutouts, saves, strikeouts per walk and the log of the pitcher's earned run average (ERA). Standard errors are in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

to all pitchers, and pitchers with different roles often receive vastly different average contract values. Thus *games started, complete games, shutouts* and *saves* are included to capture differences in each pitcher's "role" on his team.

Games started is used to identify starting pitchers, who typically have the most innings pitched of any pitcher. Each additional game a pitcher started in the previous season is estimated to increase his annual contract by $57,600. *Complete games* are relatively rare, and are typically only accomplished by dominant pitchers with both the talent and the stamina to remain in the game for the entire nine innings. Our results on this variable indicate that there exists a $295,300 increase in annual salary for each additional complete game in the previous season. *Shutouts* are even rarer than complete games, as they require the pitcher to pitch nine scoreless innings; however, our results for this performance measure are not statistically significant, most likely because of the relatively few observations with *shutouts*. Our final measure, included to account for different pitcher roles within their team, is *saves,* which is used to identify closers, who pitch at the end of the game in the most stressful situations. The averages presented previously in Table 1 indicated that these pitchers receive more than the average pitcher, and our results in Table 3 confirm this result, with each additional *save* being worth $29,200 in annual salary.

Two measures of pitcher quality that do not depend on a pitcher's role are his ratio of *strikeouts per walk* and his *earned run average*. The *strikeouts per walk* variable captures the pitcher's ability to either overpower or fool hitters, causing them to not put the ball into play, while the *earned run average* variable captures the pitcher's ability to prevent runs from scoring. Both of these measures of pitcher effectiveness are positive and significant, with an increase of one additional *strikeout per walk* leading to a $307,200 increase in average annual contract value, and an increase in a pitcher's *earned run average* leading to lower pay, but not linearly.

Our final measure of pitcher performance is a dummy variable for pitchers who were named to the *All Star* team. As with batters, this is an honor bestowed upon the best pitchers during the first half of a season. This honor is estimated to increase the pitcher's average contract value by $578,000. We excluded other awards given to pitchers, as they are only given to one pitcher per year and were therefore not applicable to our analysis due to the few observations and the limited variation in the recipients of these awards.

(b) Non-Performance Factors

In addition to the aforementioned player characteristics and past performance measures, we include three different non-performance factors that could impact a player's salary. Specifically, given the structure of the free agent market within Major League Baseball, we are interested in the impact of signing a contract while still under *club control* (before reaching free agency, a time when the team holds complete monopsony power over that player's services), the effect of signing with one's *previous team* and the impact of signing with ones *hometown team* on average annual salary.

Current free agency rules in Major League Baseball require players to have six years of experience prior to becoming a free agent. Before reaching this level of service, a player must sign a new contract with their current team. While these players under club control do begin to gain some bargaining power following their third (sometimes second) year of service, as they become arbitration eligible, they are not able to market their talent to the entire market. Since teams do not have to compete with one another for the services of a player who is not eligible for free agency, the average contract value should be lower for players under club control, though as Marburger (2004) shows, arbitration eligible players tend to receive arbitration awards which lie between the player's salary in the previous season and the average free agent salary of comparable players. Our estimates in Tables 2 and 3 corroborate this for both batters and pitchers, with each receiving $1,801,100 and $877,900 less, respectively, in annual salary.

The impact of a player signing with his previous team has a theoretically ambiguous result, as players are likely to take lower wages in order to remain on the same team and avoid the psychic costs of moving; however, teams are apt to offer higher wages to keep the players that both they, are their fans, are familiar with. Our results on players signing with the *previous team* in Tables 2 and 3 show no statistical evidence that either batters or pitchers salaries are influenced by signing with their previous team.

Likewise, a player signing with his hometown team has an ambiguous affect on salary. While players are likely to accept less money to play for their hometown team, teams are willing to pay more for these players as they offer additional marketing and fan attendance opportunities, given the fans' regional connections to such players. The results in Tables 2 and 3 indicate that the incentive for teams to pay more in order to obtain "home grown" talent outweighs the player's incentive to stay near his birthplace, with our estimates indicating that batters who sign with their *hometown team* receive $483,500 more per year, while pitchers receive $405,500 more.

(c) Task Specific Human Capital

The spatial autoregressive model outlined in the previous section allows for an examination of the existence of task specific human capital, since it directly models each player's salary as being dependent upon the salary of other players with similar task specific human capital profiles. The statistical significance of the ρ term from equation 1 in both Table 2 and Table 3 indicates the appropriateness of the spatial autoregressive model, and that task specific human capital labor markets do exist within Major League Baseball. This also implies that any previous work examining wage determination in baseball, and perhaps in many other industries, suffers from biased results due to omitted variables, and that task specific human capital should be identified within any labor market in which division of labor exists. Comparing our results via OLS and the spatial lag model indicate that, although OLS is subject to an omitted variable bias, it is not especially severe as none of the non-spatial variables change signs or significance; however, when comparing the results from OLS and from the spatial lag model, many of the coefficients for pitchers are statistically different from one another. Examining the coefficient on the spatially weighted *average annual salary of players with the same task specific human capital*, we find that salaries of players with the same task specific human capital are positively correlated, with an increase of $100,000 in a player with similar task specific human capital's salary increasing average contract values by $9,600 for batters and $31,300 for pitchers. This result may seem small in terms of economic significance, but given that players regularly sign contracts for tens of millions of dollars, failure to account for such interdependence can greatly bias any predictions formed from such wage regressions.

These results indicate that free agents have the alue of their contract determined not only by their characteristics, performance measures and non-performance measures, but also by the market for players at the same position. The market for players with task specific human capital is in turn determined by both the supply of players with each type of task specific human capital and the demand for such task specific human capital by teams. If a player has control over the length of his contract, our results show that he should choose to be a free agent in years in which there is a limited supply of players with similar task specific human capital profiles, and/or years in which the other players with similar task specific human capital can be expected to receive high salaries, as their salary levels are correlated within these labor markets.

4. CONCLUSIONS

In this study we illustrate the methodological use of spatial econometric techniques in settings where space can be defined based on non-geographic distances. Specifically, we examine the effect of task specific human capital on labor market outcomes within Major League Baseball. In this context, we use a spatial autoregressive specification to model player salaries as being dependent upon: player characteristics, player preferences and also on the salaries of similarly skilled players (e.g. on the salaries of players with similar task specific human capital profiles). Our results of this analysis suggest that contract values are affected by 'spatially close' players, and that the spatial autoregressive model based on worker characteristic 'space', rather than geographic 'space', improves our estimation results.

These results add to the literature on wage determination, contracting, and returns to human capital accumulation, as well as presenting a methodological approach to modeling economic relationships through spatial econometrics, where space is defined as non-geographic. To our knowledge, this is the first study that directly models task specific labor markets, with the results having important implications for previous studies of human capital accumulation and wage differentials, which conclude that wage differentials arise because of non-competitive factors. Instead, our results point to the potential for wage differentials stemming from task specific human capital. In addition, the methodological approach used here is general enough to be used in numerous situations in which the observations are interdependent, by using non-geographic measures of space. For example, space could be defined as similarities in cultures, skills, economies, resources and the like to better model economic outcomes.

Appendix A

Anselin (1988) provides the theoretical underpinnings of modern day spatial econometrics. In general, one can model spatial interdependence between economic entities through either a spatial autoregressive (lag) model or a spatial error model (or a hybrid of both models). Typically, the spatial error model is used if one suspects that entities may be spatially correlated, but has no economic rationale for this dependence, while the lag model is used when there is a direct economic explanation for the spatial linkage. Regardless of the model used, the spatial dependence between observations is defined econometrically by use of a spatial weighting matrix, *W*. This matrix is user specified, and empirically defines the relationship between any two observations, with each element of the matrix representing the proximity of the observations. As such, the matrix is symmetric, with all of the on-diagonal elements set to zero to prevent an observation from being regressed upon itself. For panel data, such as that used in this paper, this implies a spatial matrix of the form:

$$W = \begin{bmatrix} W_{2000} & 0 & 0 & 0 & 0 & 0 \\ 0 & W_{2001} & 0 & 0 & 0 & 0 \\ 0 & 0 & W2002 & 0 & 0 & 0 \\ 0 & 0 & 0 & W2003 & 0 & 0 \\ 0 & 0 & 0 & 0 & W2004 & 0 \\ 0 & 0 & 0 & 0 & 0 & W2005 \end{bmatrix}$$

Where each block of this matrix defines the spatial proximity of the observations within a given year. The elements within each of these blocks are then defined, in the present paper, as:

$$W_t = \begin{bmatrix} 0 & \frac{1}{\gamma_{i,j}} & \frac{1}{\gamma_{i,k}} \\ \frac{1}{\gamma_{j,i}} & 0 & \frac{1}{\gamma_{j,k}} \\ \frac{1}{\gamma_{k,i}} & \frac{1}{\gamma_{k,j}} & 0 \end{bmatrix}$$

Where $\gamma_{i,j}$ is intended to capture the skill proximity of any two players, with $\gamma_{i,j}$ increasing as player proximity decreases. In general, the spatial weighting matrix can take many forms, one example being represented above, but requires the researcher to specify weights that link any two observations, something that is typically based on geographic proximity, with larger weights placed on "close" observations. However, as demonstrated here, this proximity can be non-geographic, and instead be based on observed similarities in the characteristics of the observations. Regardless of the weights used, it is common to row standardize the spatial weighting matrix, to make interpretation of the estimated coefficient easier, as the spatial lag term in equation (1) becomes the weighted average of the dependent variable for an observation's neighbors. Not row standardizing the weighting matrix implies that the weights themselves would impact the dependent variable as observations with closer neighbors would have higher weights than observations with further neighbors.

Although others have these models using instrumental variables techniques (see Anselin (1988) for a description of this methodology or Henrickson (2012) for an empirical example) or Bayesian techniques (e.g. LeSage (1997)), the standard approach is to use maximum likelihood estimation (MLE). Assuming that the error term in equation (1) is normally distributed with constant variance, the log-likelihood function is:

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} u_i^2 - \frac{n}{2}\log\sigma^2 + \log\left|I - \rho W\right| \quad (2)$$

However, because of the popularity of spatial modeling, this maximum likelihood estimation procedure has been incorporated into many modern econometrics software packages, such as Stata, with the user only having to specify the version of the model to be run (spatial lag or error) and provide the program with a spatial weighting matrix.

REFERENCES

Altonji, Joseph G. and Robert A. Shakotko (1987), "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54 (3): 437–459.

Anselin, Luc. (1988), *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic Publishers.

Becker, Gary S. (1964), *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. New York: National Bureau of Economic Research.

Becker, Sascha O. and Ludger Woessmann. (2009), "Was Weber Wrong? A Human Capital Theory of Protestant Economic History." *The Quarterly Journal of Economics*, 124 (2): 531–596.

Blass, A. (1992), "Does the Baseball Labor Market Contradict the Human Capital Model of Investment?" *The Review of Economics and Statistics* 74 (2): 261-268.

Blonigen, Bruce, Ronald Davies, Glen Waddell and Helen Naughton. (2007), "FDI in Space: Spatial Autoregressive Relationships in Foreign Direct Investment." *European Economic Review* 51: 1303-1325.

Burdett, Kenneth, Carlos Carrillo-Tudela and Melvyn G. Coles. (2011), "Human Capital Accumulation and Labor Market Equilibrium." *International Economic Review*, 52 (3): 657–677.

Gibbons, Robert and Michael Waldman (2004), "Task-Specific Human Capital." *The American Economic Review* 94 (2): 203-207.

Gori, Luca and Luciano Fanti. (2008), "Human Capital, Income, Fertility and Child Policy." *Economics Bulletin*, 9 (7): 1–7.

Healy, Andrew (2008), "Do Firms Have Short Memories? Evidence From Major League Baseball." *Journal of Sports Economics*, 9 (4): 407–424.

Henrickson, Kevin E. (2012), "Spatial Competition and Strategic Firm Relocation." *Economic Inquiry*, 50(2): 364-379.

James, Bill. (2002), *Wins Shares*. (Stats, Inc Publishing).

Kalnins, Arturs. (2003), "Hamburger Prices and Spatial Econometrics." *Journal of Economics and Management Strategy* 12: 591-616.

Kambourov, Gueorgui and Iourii Manovskii. (2009), "Occupational Specificity of Human Capital." *International Economic Review*, 51 (1): 63–115.

LeSage, James P. (1997), "Bayesian Estimation of Spatial Autoregressive Models." *International Regional Science Review* 20: 113-129.

Marburger, Daniel. (2004), "Arbitrator Compromise in Final Offer Arbitration: Evidence from Major League Baseball." *Economic Inquiry* 42 (1): 60-68.

McMillen, Daniel, Larry Singell and Glen Waddell. (2007), "Spatial Competition and the Price of College." *Economic Inquiry* 45 (4): 817-833.

Mincer, Jacob. (1958), "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy*, 66 (4): 281–302.

Mortensen, Dale T. (2003), *Wage Dispersion: Why are Similar Workers Paid Differently?* Cambridge, MA: MIT Press.

Neal, Derek. (1995), "Industry-Specific Human Capital: Evidence from Displaced Workers." *Journal of Labor Economics* 13 (4): 653-677.

Shaw, Kathryn. (1984), "A Formulation of the Earnings Function Using the Concept of Occupational Investment." *The Journal of Human Resources* 19 (3): 319-340.

Topel, Robert. (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *The Journal of Political Economy* 99 (1): 145-176.

Treme, Julianne and Samuel K. Allen. (2009), "Widely Received: Payoffs to Player Attributes in the NFL." *Economics Bulletin*, 29 (3): 1631–1643.