Productivity Spillovers in Team Production: Evidence from Professional Basketball*

Peter Arcidiacono
Duke University
& NBER

Josh Kinsler
University of Rochester

Joseph Price
Brigham Young University
& NBER

September 24, 2013

Abstract

Workers contribute to team production through their own productivity and through their effect on the productivity of other team members. We develop and estimate a model where workers are heterogeneous both in their own productivity and in their ability to facilitate the productivity of others. We use data from professional basketball to measure the importance of peers in productivity because we have clear measures of output and members of a worker's group change on a regular basis. Our empirical results highlight that productivity spillovers play an important role in team production and accounting for them leads to changes in the overall assessment of a worker's contribution. We also use the parameters from our model to show that the match between workers and teams is important and quantify the gains to specific trades of workers to alternative teams. Finally, we find that worker compensation is largely determined by own productivity with little weight given to the productivity spillovers a worker creates, despite their importance to team production. The use of our empirical model in other settings could to lead to improved matching between workers and teams within a firm and compensation that is more in-line with the overall contribution that workers make to team production.

*We thank seminar participants at Iowa State for helpful comments.

1 Introduction

The classic economic model predicts that workers will be paid the value of their marginal product of labor. Estimating this marginal product may be complicated by team environments in which workers contribute to team production directly but also indirectly through their effect on the productivity of other team members. If firms are able to identify workers who boost peer productivity, they can leverage complementarities in team production through team and task assignments. Workers who bring out the best in others will likely be assigned to tasks essential for firm production.

Mas and Moretti (2009) provides an excellent example of spillovers in team production by looking at the placement of cashiers in a supermarket. Placing the most productive cashiers in full view of the other cashiers resulted in the other cashiers working faster. However, Mas and Moretti provide one of the few examples where actual productivity is observed and the economist has not intervened in the system.[1] Exceptions include Hamilton, Nickerson, and Owan (2003), who examine worker interactions in the garment industry, and a set of papers analyzing productivity in the academy, Azoulay, Zivin, and Wang (2010) and Waldinger (2010, 2012).

The assumption made in this literature–as well as the abundant literature on peer effects in education–is that the individuals who are most productive themselves are also the ones who will make others most productive. However, this assumption is a restriction that may be violated in many contexts. For example, within the academy there are professors who choose to focus exclusively on their own research, providing little in terms of public goods. There are other professors who are particularly adept at helping their colleagues in their research and may do so even at the expense of their own research. Similarly, a brilliant but introverted student may not be as helpful to the learning of the other students as the perhaps not-so-brilliant student who asks good questions in class.

Two papers using sports data highlight the heterogeneity in how spillovers may operate. Gould and Winter (2009) use data on baseball players to analyze how batter performance is related to the performance of other batters on the team. This paper fits perfectly with the idea that the most productive players have the largest positive peer effects: batting in front of a high-performing player results in receiving better pitches because the pitcher will not want to risk a walk prior to facing the high-performing player. Guryan, Kroft, and Notowidigdo (2009) examine how the productivity of one's golf partners affects own performance, finding no significant effects from being paired with

[1] Field experiments have also been used to examine peer effects in the workplace. See the series of papers by Bandiera, Baranakay, and Rasul (2005, 2009, 2010, and 2013) as well as Falk and Ichino (2006).

better golfers. But this may be a case where allowing individuals to be multidimensional in their abilities is important. Certain players may be very productive but are surly or disobey common golf etiquette, both of which may serve to distract their partners.

In this paper, we develop and estimate a model of team production where individuals are heterogeneous in their own productivity as well as in their ability to help others be productive. We focus on an industry, professional basketball, where the ability to help others is clearly an important part of team production. Sports data provide an excellent opportunity to study team production because the members of a team can be clearly identified and there are frequent changes in the players that compose a particular team.

We use possession-level data from games played in the National Basketball Association (NBA). We demonstrate that productivity spillovers play an incredibly important role in team production. In fact, we find that a standard deviation increase in the spillover effect of one player has nearly the same impact on team success as a standard deviation increase in the direct productivity of that player. Estimates of the model also allow us to form player rankings based on the overall contribution to team production. We compare these rankings to estimates of team production when spillovers are ignored. Players who are generally perceived by the public as selfish see their rankings fall once we account for productivity spillovers.

We also use our model to highlight how the value of a particular player can vary depending on the composition of his teammates. Most firms have various teams within their organization and have the ability to reassign workers across teams. Since individual productivity and productivity spillovers play a complementary role, the overall contribution of a player will depend on the composition of the other players already on the team.[2] We find that the assignment that produces the greatest increase in team productivity is often an assignment that does not maximize the direct productivity of the player. This suggests a tension that firms need to balance between team and player productivity, especially in firms where individual productivity has a large effect on compensation.

Given the large role spillovers play in team production in this industry, we would expect significant returns in the labor market to the ability to help others. This is not the case. Returns to own productivity are substantially higher than returns to the ability to help others, well beyond their differences in their contribution to team production. Part of the reason for this is the difficulty in measuring the ability to help others. As in the academy, direct productivity is easily observed in ways that facilitating the productivity of others is not. To the extent that own productivity and

[2] Similarly, Ichniowki, Shaw, and Prennushi (1997) find that the returns to innovative work practices (e.g. teams, incentive pay, etc.) are complementary in the steel finishing industry.

facilitating the productivity of others is endogenous, the lack of returns to the latter may result in inefficient effort allocations among workers.

2 Data

To estimate a model of player performance, we use publicly available NBA play-by-play data covering all games during the 2006-2009 regular seasons gathered from espn.com. The raw play-by-play data provides a detailed account of all the decisive actions in a game, such as shots, turnovers, fouls, rebounds, and substitutions. Plays are team specific, meaning that there is a separate log for the home team and the away team. Associated with each play are the player(s) involved, the time the play occurred, and the current score of the game. While our model of player productivity is estimated using only the play-by-play data, we augment it with additional biographical and statistical information about each player gathered from various websites which we discuss later in this section. As described in the data appendix, a number of steps are taken to clean the data. These include establishing which players are on the court, acquiring the outcomes of possessions, and matching the names of the players to data on their observed characteristics such as position and experience.

Table 1 describes our estimation sample in further detail. We use data from 905,378 possessions and 656 unique players active in the NBA from 2006-2009. On average, each player is part of 13,801 possessions, split evenly between offense and defense. The average number of possessions for each player-team-season combination is 4,507. The corresponding 25th and 75th percentile values are 1,130 and 7,470. The final four rows of Table 1 describe the typical outcomes for a possession. Slightly more than 50% of the time the offensive team scores, and conditional on scoring the offense scores on average 2.1 points.

To supplement the play-by-play data, we merge in biographical and statistical information about each player. Our primary source for this information is basketball-reference.com. The website contains basic player information such as date of birth, height, position, and college attended and a full set of statistics for each season the player is active. We also gathered information on salaries and contract years from prosportstransactions.com and storytellerscontracts.com. We also obtained additional measures of player performance from basketballvalue.com and 82games.com.

3 Model and Estimation

In this section we present a model of team production, discuss identification, and describe our estimation strategy. The innovation of the model is that the ability of an individual to influence the productivity of others is not directly tied to own productivity. We tailor the model to the NBA context, though it would be simple to expand the framework to other types of production.[3] The number of parameters to be estimated is quite large and would be computationally prohibitive using straight maximum likelihood. Consequently, we take an iterative approach as in Arcidiacono, Foster, Goodpaster, and Kinsler (2012).[4]

3.1 Model setup

Our unit of analysis is an offensive possession during an NBA game. There are five offensive and five defensive players on the court during every possession. For a given possession n, denote the set of players on the court as P_n where P_n includes the offensive players on the court O_n and the defensive players on the court D_n. For notational ease we abstract from the fact that possessions are typically observed within games, which themselves are observed within seasons. Additionally we abstract from the concept of team, even though the potential sets of offensive and defensive players will be determined by team rosters. A possession can end in one of six ways, no score or one of the five offensive players in O_n scoring at least one point. We assume that each player i on the court is fully characterized by three parameters: (i) their ability to score, o_i, (ii) their ability to help others score, s_i, and (iii) their ability to stop others from scoring d_i.

Assume for the moment that there is no heterogeneity in defensive skills. The likelihood that offensive player i scores to conclude a possession will depend on i's own ability to score and his ability to help others score as well as the similar skills of his teammates on the court. Denote $y_{in} = 1$ if the individual scores and $y_{in} = 0$ otherwise. We assume that the probability that player

[3]For example, in Mas and Moretti (2009) checkout cashiers are assumed to influence other cashiers through their own productivity. However, it would be straightforward to allow for completely separate effects.

[4]See Burke and Sass (2013) for an application of this method in education and Cornelissen, Dustmann, and Schonberg (2013) for an application in the labor market.

i scores at least one point during possession n is given by[5]

$$Pr(y_{in} = 1|P_n) = \frac{\exp\left(o_i(1 - \sum_{k \in O_n, k \neq i} s_k)\right)}{1 + \sum_{i \in O_n} \exp\left(o_i(1 - \sum_{k \in O_n, k \neq i} s_k)\right)}. \tag{1}$$

The probability that player i scores to end possession n is increasing in o_i, the offensive intercept of player i. An increase in the offensive spillover of player $k \neq i$ will have an ambiguous effect on the probability that player i scores since an increase in s_k also benefits the other offensive players in O_n. The probability that possession n ends with no points scored is simply $1 - \sum_{i \in O_n} Pr(y_{in} = 1|P_n)$.

The above model is inadequate since defenders will vary in ability. Thus, the probability that one of the players in O_n scores will depend on the composition of the players in D_n. To account for defense, we alter the above framework such that the probability that player i scores at least one point during possession n is given by

$$Pr(y_{in} = 1|P_n) = \frac{\exp\left(o_i(1 - \sum_{k \in O_n, k \neq i} s_k)\right)}{\exp\left(\sum_{j \in D_n} d_j\right) + \sum_{i \in O_n} \exp\left(o_i(1 - \sum_{k \in O_n, k \neq i} s_k)\right)} \tag{2}$$

The difference between Equation (1) and Equation (2) is that the index associated with no points being scored now varies with the abilities of the defenders in D_n. The joint defensive prowess of the players in D_n is a linear function of the defensive intercepts of each player j.

Possessions that yield positive points do not necessarily contribute equally to team success since there are a range of plausible point outcomes. To determine the expected number of points per possession for each player we scale the probability of scoring positive points by the expected number of points conditional on scoring for each player,

$$E[Points_{in}|P_n] = E[Points_i|y_i = 1] \times Pr(y_{in} = 1|P_n). \tag{3}$$

The above formulation assumes that the number of points a player scores conditional on scoring is unrelated to the identities of the other players on the court. Thus, estimates of $E[Points_i|y_i = 1]$ can be obtained separately from the estimation of the other parameters.

3.2 Normalizations and identification

The offensive intercept parameters can be readily identified if data for many possessions is available. However, to identify the offensive spillovers it is necessary to observe player i in multiple groups.

[5]Separate identification of own productivity from spillover productivity is not reliant on the logit production function. We choose this specification since possession outcomes are binary in nature. In fact, a linear model would be significantly easier to estimate.

Put differently, the spillover parameters can only be separately identified if we observe player i with different sets of teammates across possessions. If not we could simply redefine o_i such that $o_i^* = o_i(1 + \sum_{k \in O_n, k \neq i} s_k)$ for all i and estimate the o_i^*'s. Without switching there is no way to separate the offensive intercepts from the offensive spillovers.

Even when the set of teammates in O_n varies across possessions, it is still necessary to make a normalization on the spillover parameters. For any set of o_i and s_i, it is always possible to redefine the parameters such that the predicted probabilities are identical. For example, define

$$o_i^* = o_i(1 + \overline{s}) \tag{4}$$

where $\overline{s}$ is the mean spillover calculated using i's spillover and the spillover of all the players i is ever grouped with. Additionally, define

$$s_i^* = \frac{\sigma_o}{\sigma_{o^*}}(s_i - \overline{s}) \tag{5}$$

where σ_o is the standard deviation of the original offensive intercepts for i and all the players i is ever grouped with. σ_o^* is defined similarly, except over the redefined offensive intercepts. It can be shown that the predicted probability that i scores, is identical across $\{o, s\}$ and $\{o^*, s^*\}$. A natural normalization is simply to constrain $\overline{s} = 0$.

Adding defensive heterogeneity to the scoring probabilities does not alter the identification argument regarding the offensive parameters. Note that it is not possible to identify the baseline defensive productivities unless player i is observed on defense with different sets of teammates. This is because defensive outcomes are essentially group outcomes. Thus, switching teammates is critical for identifying the defensive intercepts. Note that the mean defensive intercepts must also be normalized since the mean is not separately identified from the scale of the baseline offensive parameters.

3.3 Iterative algorithm

There are three parameters to be estimated for each player we observe in the data. As previously discussed it is not possible to estimate the offensive spillovers without an additional restriction. If all the players in the sample are connected, in other words, every player can be linked to every other player through teammates, then it would only be necessary to restrict the overall mean of the offensive spillovers to zero. In our data, all the players are in fact connected, since players switch teams both within and across seasons. However, jointly estimating all of the player parameters by maximum likelihood imposing this one restriction is not computationally feasible.

As a result, we pursue an estimation strategy that treats each team-season as independent entities. Players that switch teams within a season are treated as completely unrelated. For each team-season, we normalize the possession-weighted spillover to zero. Once we have estimated all the player parameters imposing these restrictions, we adjust team and player spillovers to be consistent with observed changes in the same player's performance across different teams and seasons. This approach has two advantages. First, it facilitates estimation since we can iteratively estimate the offensive intercepts, offensive spillovers, and defensive intercepts team by team. Second, it allows player productivity to vary across seasons as a result of random factors such as health and luck.

Once we have defined players as any unique combination of player-season-team, we estimate the offensive and defensive parameters using an iterative estimation approach. The method has three broad steps that correspond to estimating the offensive intercepts, defensive intercepts, and the offensive spillovers. The following steps outline the estimation procedure more precisely, where now players i's parameters are indexed by team-season pairs (t).

- Step 0: Make an initial guess of the parameters, denoted by $\{o_{it}^0, s_{it}^0, d_{it}^0\}$.
- Step 1: Estimate by maximum likelihood o_{it}^1 conditional on $\{s_{it}^0, d_{it}^0\}$.
- Step 2: Estimate by maximum likelihood d_{it}^1 conditional on $\{o_{it}^1, s_{it}^0\}$.
- Step 3: Estimate by maximum likelihood s_{it}^1 conditional on $\{o_{it}^1, d_{it}^1\}$.

Estimation proceeds by iterating on Steps 1-3 until convergence where the qth iteration estimates are characterized as $\{o_{it}^q, s_{it}^q, d_{it}^q\}$. Note that within Steps 1-3 estimation proceeds separately for each team-season, implying that each maximization step is searching over only approximately 15 parameters–one for each player who played on a particular team in a particular season. Our approach relies on the fact that there are no across team interactions to be concerned about when the remaining parameters are taken as fixed–when estimating the offensive parameters for one team, all the defensive parameters of the players associated with the other teams are taken as given. Note that this approach would not possible if the parameters for a particular player were constrained to be the same across seasons or across teams within a season.

3.4 Linking across years

Once the above process converges, we are left with a set of parameters describing player performance for each team-season. However, the offensive intercept and spillover parameters cannot be readily compared across teams or seasons since the average spillover is assumed to be zero for each team-season. To relax this restriction, we utilize player movement across teams both within and across seasons to adjust the mean spillovers for each team-season. This in turn allows us to adjust both the offensive spillover and offensive intercept for each player, while holding fixed any lineup's predicted probability of scoring.

The key assumption underlying this procedure is that all changes in player performance over time operate through the players themselves: there are no team or coaching effects. For example, teammates can all have above-average spillover seasons at the same time, but these positive deviations are the result of players naturally having good years in ways that are not attributed to the team or the coach.

The procedure to relax the team-season spillover normalizations effectively removes team-season effects by iterating on the following three steps:

1. Regress $\hat{s}_{it}$ on player fixed effects and team-season fixed effects (δ_t), weighting each observation by the observed number of possessions. We continue to normalize one team-season effect to zero. The estimated team-season effects are interpreted as the true average spillover for each team.[6]

2. Create adjusted offensive intercepts: $\hat{o}^*_{it} = \frac{\hat{o}_{it}}{1-4*\delta_t}$

3. Create adjusted offensive spillovers: $\hat{s}^*_{it} = \left(\frac{\hat{\sigma}_t}{\hat{\sigma}^*_t}\right)\hat{s}_{it} - \delta_t$

The adjustments in Steps 2 and 3 keep the predicted probability of scoring for any lineup the same across the $\{o, s\}$ and $\{o^*, s^*\}$ parameters. The adjusted parameters simply shift the mean spillover on a particular team to be broadly consistent with how those same players perform on other teams in other years. Iterations continue until the regression of $\hat{s}^*_{it}$ on player fixed effects and team-season fixed effects yields team-season effects that are all zero.[7]

4 Results

The estimation procedure yields three parameters for each player-season-team combination. If these parameters are capturing something permanent about player skill, they should be rather stable over time and across teams. To investigate this, we estimate separate fixed effects regressions for

[6]It is possible to control for experience or age when estimating this regression. Our results do not change if we allow for these additional explanatory variables.

[7]Only a few iterations are required in practice.

each of our skill measures. The outcome variables are the player-season-team estimates and the only explanatory variables are player fixed effects. Each player-season-team skill observation is weighted by the number of observed possessions. With these simple regressions we are able to explain approximately 83%, 50%, and 57% of the variation in the offensive intercept, offensive spillover, and defensive parameters across seasons and teams. Thus, it does appear that we are capturing something intrinsic about each player.

For evaluating players and teams, however, the parameter estimates themselves are not terribly informative. In the next few sections we demonstrate how the various player skills contribute to team performance, how to rank players using our estimates, and finally whether players and teams make decisions that are consistent with the results of our model.

4.1 Importance of the three factors

As noted above, the scale of the offensive intercept, offensive slope, and defensive intercept are not meaningful on their own. To illustrate how important each of these components are for team success, we perform the following exercise using player skill estimates from the 2009-2010 season. We first identify the four most utilized players on each team in 2009-2010 based on total possessions. We then ask how each team of four players would perform when various types of players are added. Our measure of team performance is the predicted per possession point differential against an average team.[8]

The results of this exercise are illustrated in Table 2. The first row of results shows the distribution of predicted point differentials when an average player is added to each team's top four utilized players. Across 30 teams, the average differential is slightly above zero, with a standard deviation equal to 0.087. Using this as a baseline we then explore how each team's per possession point differentials change when players with particular skills are added. We consider six different player types, altering the offensive intercept, offensive slope, and defensive intercept by one standard deviation in either direction from the average.[9]

[8] The five offensive intercepts for the average team are chosen to match the average intercepts across teams by player offensive rank. For example, the offensive intercept for the most productive scorer on the average team is the mean across all 30 teams of the best scorer's offensive intercept. Each player on the average team is assigned the overall average spillover and defensive parameter since these enter the production function linearly.

[9] For this exercise, we define a one-standard deviation change in skill using the distribution of skills generated when we combine player estimates across seasons through the fixed effects estimation discussed at the beginning of Section 4. The standard deviation of player skills using the raw player-team-season estimates is larger, likely a reflection of measurement error in the underlying estimates. Note that the standard deviation of player skills does not change

We find that all three factors are important for team performance. Adding a one standard deviation better offensive intercept, offensive spillover, or defensive intercept player improves a team's per possession point differential by 0.032, 0.024, and 0.035 points respectively. Compared to the baseline standard deviation of point differentials of 0.087, these numbers indicate that adding a one standard deviation more skilled player increases a teams per possession point differential by 28% to 40% of a standard deviation. The largest change in team performance stems from the addition of a better defensive player. Adding a higher offensive intercept player also benefits the team, but the positive effect will be muted since adding a good scorer necessarily decreases the opportunities that the other players on the team have to score. In other words, there is a substitutability in production on offense that does not exist on defense.

The results in Table 2 also document the variability across teams in the effect different types of offensive players have on team performance.[10] Because of the complementarities in the probability of scoring, the benefit of adding a particular type of player will vary by team. For example, a high offensive intercept player may be valued more by teams that have fewer high spillover players since a productive scorer doesn't rely as much on his teammates to score. In contrast, a team with a number of productive scorers may prefer to bring in a high spillover teammate to enhance the productive skills already present.

4.2 Position comparisons

High intercept, spillover, or defensive players are often associated with particular positions on a standard NBA team.[11] For example, point guards are generally viewed as facilitators, while centers are expected to protect the basket on defense. Table 3 shows how the various skills break down by position. For each position, we show the average skill measure for each of our estimated parameters and a measure of a player's overall effectiveness (a combination of our three measures which we discuss further in the next section). For comparison purposes we also include two common measures

appreciably if we use only three seasons to run the player fixed effects regressions, suggesting that we have successfully eliminated the bulk of the measurement error.

[10]In contrast to the offensive skills, there is little variability across teams in point differential changes associated with adding a one standard deviation better (or worse) defensive player. This is primarily a reflection of the fact that we assume that there are no complementarities in defensive production.

[11]The most common lineup in professional basketball contains a point guard, shooting guard, small forward, power forward, and center. However, teams face no restrictions regarding which positions players are allowed to play at one time.

of player effectiveness, player efficiency rating (PER) and adjusted plus minus (APM).[12] All measures are standardized to have a mean of zero and standard deviation equal to one across positions.

The estimates of our model match the basic intuition about the types of skills different position players bring to a team. Point guards are by far the best spillover players but tend to be below average scorers and very poor defenders. In contrast, centers are 0.68 standard deviations better than the average defensive player and the huge defensive benefit of centers provide their most important contribution from an overall effectiveness standpoint, with centers being 0.33 standard deviations more effective than the average player. The PER's ordering of overall effectiveness by position is similar to our ranking except in the case of point guards which are ranked higher under PER. This is consistent with the criticism that the PER fails to account for defensive contributions. The APM, on the other hand, ranks centers quite differently than either PER or our measure. In the next section we discuss further how to rank players individually.

4.3 Ranking workers overall contribution

The previous sections suggests that attempts to rank players individually in a team sport is misguided since the value of each player necessarily depends on who his teammates are. However, NBA player rankings are ubiquitous and often fail to account for the team nature of the sport. We develop a player ranking that directly accounts for the complementarities present in a team setting by using estimates of each player's underlying skills from our spillover model. We first describe how we construct our rankings, and then compare them to other common rankings and rankings we generate when ignoring spillovers in team production.

There are a number of ways to measure the effectiveness of each player given our estimated parameters. We construct our preferred measure by first taking each player and pairing him with

[12]PER is a rating of a player's per-minute productivity that is generated using a complicated formula based on box score statistics. The precise formula can be found at http://www.basketball-reference.com/about/per.html. PER does not consider who each player plays with or against and is viewed largely as a measure of offensive effectiveness. PER has been criticized as a measure of player effectiveness since it emphasizes shot taking (see http://wagesofwins.com/2006/11/17/a-comment-on-the-player-efficiency-rating/). APM ratings indicate how many additional points are contributed to a team's scoring margin by a given player in comparison to the league-average player over the span of a typical game. APM is constructed using a fixed-effects regression where the dependent variable is the per-possession point differential for a given set of players on the court and the explanatory variables are player fixed effects. Further details are available at http://www.82games.com/barzilai2.htm. The advantage of APM relative to PER is that it explicitly accounts for who a player plays with and against. However, because the model is linear, it cannot capture complementarities in production.

an average team. A player's measured effectiveness is then the per possession point differential when this team plays against an average opponent.[13] The per possession point differentials are then standardized so players can be compared in standard deviation units. We create two additional measures of player effectiveness. The first takes our preferred measure and adjusts for position, since as Table 3 indicates there are significant differences in player effectiveness by position. Because teams typically field lineups with one player at each position, players who excel at underperforming positions will be valued more. Finally, rather than take each player and put him on an average team, our third measure of player effectiveness replaces each player with an average player and asks how his team's performance changes.[14] To some extent, this measure accords with how valuable each player is to their team.

Table 4 lists the top ten players in 2009-2010 according to our three rankings along with the rankings according to PER and APM.[15] Our preferred rankings indicate that Dwight Howard is the most effective player, over three standard deviations better than the average NBA player. The primary reason that Dwight Howard is so highly ranked is that he is the top ranked defensive player, almost a full standard deviation better than the next best defender. The rest of the top ten is full of names that are familiar to basketball fans, but are not necessarily the brightest stars in the game. For example, Al Horford and Chris Andersen are highly ranked because they are well above average both offensively and defensively. Horford is an above average offensive intercept, offensive spillover, and defensive player. Andersen is actually a below average offensive intercept player, but his presence on the court generates enough extra opportunities for his teammates that his offensive spillover measure is two and a half standard deviations above the mean. The rankings based on APM also pick-up Andersen's overall effectiveness.

LeBron James is ranked number six according to our preferred method. This "low" ranking is a

[13]The average teammates a player is assigned and the average opponent are constructed using the average offensive intercepts across teams by player offensive rank. For example, the offensive intercept for the most productive scorer on the average team is the mean across all 30 teams of the best scorer's offensive intercept. Each player on the average team is assigned the overall average spillover and defensive parameter since these enter the production function linearly.

[14]For this measure we create rankings only for the top five most utilized players on each team. This allows for a straightforward determination of who the teammates will be when each player is replaced with an average player.

[15]We only consider players who accumulated at least 2,000 total possessions for any team in 2009-2010. This restriction limits the rankings to those players observed often enough to accurately estimate their underlying skills. 337 players, or 67% of active players in 2009-2010, played in 2,000 or more possessions. Each team has approximately 11 players who play more than 2,000 possessions. Among those who played at least 2,000 possession the average number of possession was 6,436. The mean for those who played fewer than 2000 possessions is 741. Rankings for 2006-2007 through 2008-2009 can be compiled in a similar manner.

reflection of the fact that in 2009-2010 LeBron is a good, but not great defender, and only an average spillover player. Based on offensive intercept alone, LeBron would be the highest ranked player, meaning that when added to an average team LeBron would have the highest scoring probability relative to adding any other player.[16] The PER measure is often criticized for over-valuing shooting and scoring and not surprisingly James comes out ahead on this measure.

When the rankings are adjusted for position or team there are slight changes. Because centers are on average the most effective players, Dwight Howard is de-valued when ranked relative to other centers and drops to the 4th best overall player. Point guards, shooting guards, and small forwards move up the ranks, with Kevin Durant now identified as the most effective player. The player rankings changed very little when players are assessed based on how their team would perform without them. Finally, many of the names on our preferred ranking list appear in the PER and APM rankings. In fact, the possession weighted correlation between our preferred measure of standardized point differential and PER is 0.42. The correlation with the APM rankings is significantly higher, equal to 0.78.[17] This is not surprising since our measure is more similar to APM since it measures a player's effectiveness controlling for the identity of the other players on the court.

4.4 Ignoring spillovers

In Table 5, we examine how our player rankings change when estimating a model that ignores spillovers, essentially ruling out any complementarity in offensive production. The first column shows the top ten players based on point differentials when playing with average teammates against an average opponent. Many of the names remain the same, such as Dwight Howard and LeBron James, but there are significant changes. In particular, players that tend to score often and also play above average defense tend to move up in the rankings. Examples include Tim Duncan, Kobe Bryant, and Chris Bosh.

The second through fourth columns look more closely at some of the changes in standardized point differentials across the spillover and no spillover models. The second column lists the changes for some notable players. For example, Carmelo Anthony, a notoriously difficult player to play with is 1.31 standard deviations better in the no spillover model than in the spillover model. In contrast, Steve Nash, a player widely believed to be one of the best offensive facilitators in the NBA

[16]Note that a ranking based strictly on each player's offensive intercept would yield a top-five of LeBron James, Dwyane Wade, Kevin Durant, Kobe Bryant, and Carmelo Anthony. These are five of the most recognizable and renowned scorers in the NBA.

[17]Again, only players with more than 2,000 possessions are considered.

is 1.30 standard deviations better in the spillover model. Columns three and four of Table 5 lists the ten players who have the largest positive and negative swings in point differentials between the spillover and no spillover models. The players that tend to improve greatly when complementarities are modeled are pass-first point guards, such as Jason Williams and Jamaal Tinsley, and players who tend not to score but generate opportunities for their teammates through offensive rebounds, screens, and passing, such as Theo Ratliff, Chris Andersen, and Anderson Varejao. The list of the ten largest negative changes is full of players who are well known to be not only bad passers, but shoot-first type players, like Chris Kaman and Carmelo Anthony.

By measuring player effectiveness within a team we are better able to identify valuable players who do not necessarily compile many of the standard statistics NBA players are so often judged by. In addition, players who tend to diminish the skills of their teammates are de-valued accordingly. However, team success depends on how players interact on the court. The next section illustrates two examples of this idea.

4.5 Allocating workers to the optimal team

For the purposes of ranking individual players we considered how each player performs with an average team. However, when actual player personnel decisions are made, success will hinge on how the various components of a team work together. We examine one of the most high profile personnel decisions in the history of the NBA when Lebron James became a free agent in the summer following the 2009-2010 season. Using our model, we evaluate Lebron's decision, examining both his own performance and the likelihood of team success. Table 6 presents the model predictions for the teams most interested in signing Lebron, Cleveland, Miami, Chicago, and New York. Cleveland provided the greatest opportunity for individual output, while Miami offered the greatest chance for team success. LeBron's predicted per possession probability of scoring declines from 0.187 with Cleveland to 0.165 with Miami, a drop of 11.8%. Interestingly, Lebron's scoring average per 36 minutes in his first year in Miami declined by 8.5% relative to his last three years in Cleveland. However, Lebron was more than willing to give up his personal statistics for an increased chance at team success. By joining Miami, Lebron increases his team's predicted per possession point differential from 0.043 to 0.219. Miami ended up winning two of the next three NBA championships.

In the summer following the 2009-2010 NBA season, another high profile free agent was on the market, Amar'e Stoudemire. Stoudemire had played for eight consecutive seasons with the Phoenix Suns and at the time was best described as an offensively skilled center with injury concerns. During his tenure in Phoenix, Stoudemire had the benefit of playing with a very high spillover player in Steve Nash, making it difficult to ascertain how well he might perform on another team. Reportedly, Phoenix was only willing to give Amar'e a four year contract at an undisclosed salary, while the New York Knicks were willing to sign Amar'e to a five year contract for 100 million dollars. Amar'e ultimately signed with New York.

Should New York have been willing to give Amar'e more than Phoenix? Table 7 shows the predicted performance for both New York and Phoenix with and without Amar'e Stoudemire on the team. The final column of the table shows the change in each team's per possession point differential with Amar'e instead of an average player. Given their projected lineups, Amar'e was more valuable to the Knicks than to the Suns based on predicted team performance. Thus, our model is consistent with the Knicks' decision to offer Amar'e a more lucrative deal. From Amar'e's standpoint, the Knicks were also more attractive in terms of individual performance, as his points per possession is predicted to be 8% higher.[18]

4.6 Returns to the three factors

The previous section highlights the usefulness of our model for evaluating potential personnel decisions, but teams need to decide not only which players to obtain but also how much to pay them. Table 3 indicates that the three player skills we have identified, offensive intercept, offensive spillover, and defensive intercept, are associated with improved team performance. In this section, we examine whether player compensation correlates with our measures of player skill.

Table 8 lists the results from a series of OLS regressions where the dependent variable is contemporaneous log earnings. The skill measures we use as regressors varies across columns, allowing for us to compare the predictive power of our skill measures and standard player measures. The unit of observation in these regressions is a player-season-team combination, where we observe each player for a maximum of four seasons, 2007-2010.[19] The first column of results indicate that a one standard deviation increase in a player's offensive intercept is associated with a statistically significant 37% increase in annual earnings. A one standard deviation increase in the defensive intercept is associated with a 17% increase in earnings, while there is essentially no monetary gain to being

[18] In his first year with the Knicks, Amar'e's scoring average per 36 minutes actually increased by 3%.

[19] In the NBA, first-round draft choices are assigned salaries according to their draft position. Each contract is for two years, with a team option for the third and fourth seasons. These structured contracts may weaken the relationship between skills and earnings, since players are drafted based on potential, not performance. However, we estimated the earnings regressions using only players who have been in the league for at least four years. The results hardly change.

a better spillover player. The results are robust to controls for player position and experience.

As a point of comparison, we also estimate earnings regressions that utilize existing player effectiveness measures. Players with higher PER and APM tend to earn significantly more than other players. Our skill measures explain more of the variation in log earnings than APM, but slightly less than PER. The results across regressions suggest that teams tend to compensate players for easily measured statistics (high R^2 for PER), but fail to identify players that add to team performance in difficult to observe ways (no effect of spillover skill).

One potential reason for the apparent lack of return to the spillover factor is that this parameter is somewhat noisier than either the offensive or defensive intercept. So rather than estimate a log earnings regression using a single player-season-team skill measure, we examine how total earnings over the four seasons in our sample is related to a player's possession weighted average skill measures. Table 9 shows the results of these regressions. Similar to the results from Table 9, players with higher offensive and defensive intercepts are rewarded with higher total earnings, to the tune of 53% and 21% per standard deviation respectively. However, the results now indicate that high spillover players also earn significantly more than low spillover players. A one-standard deviation increase in a player's average offensive spillover parameter is associated with an increase in total earnings of approximately 12%. Again, the PER measures explains the greatest amount of variation in total earnings, followed by the three skill factors and APM.

5 Conclusion

Worker skills are multidimensional. One of the skills that may be important to a variety of production processes is the ability to bring out the best in others. In this paper, we use data from the NBA to identify three measures for each player: their ability to score, their ability to defend, and their ability to help others score. It is this last factor that differentiates our work and also substantially complicates estimation. Using an iterative approach along the lines of Arcidiacono et al. (2012), we show that estimating models of this type can be accomplished in a straightforward manner.

We find that all three factors are important components to overall team productivity and probability of success. Ignoring spillovers has a substantial effect on assessing the overall contribution of specific players causing previous approaches to underestimate the contribution of "team" players. We also find that there are complementarities in production between direct forms of productivity and indirect forms that operate through productivity spillovers. As such, some teams will value particular players more than others based on the current composition of their team. Players who are particularly strong at scoring but are not good facilitators will be more valued by teams that are composed of players who are not very strong at scoring themselves.

We also find that players are primarily compensated based on their direct contributions to team production with little weight given to their ability to increase the productivity of their teammates. This misalignment of incentives might reduce the incentive for players to invest in or engage in actions that increase their positive effects on the productivity of their teammates, especially in cases where compensation is based on relative performance. The use of our empirical model in other settings could to lead to improved matching between workers and and compensation that is more in-line with the overall contribution that workers make to team production.

References

[1] Arcidiacono, Peter, Foster, Jennifer, Goodpaster, Natalie, and Josh Kinsler. "Estimating spillovers using panel data, with an application to the classroom", *Quantitative Economics*, 3:3, 421-470.

[2] Azoulay, Pierre, Zivin, Joshua S. Graff, and Jilian Wang (2010): "Superstar Extinction", *The Quarterly Journal of Economics*, 125:2, 549-589.

[3] Bandiera, Oriana, Barankay, Iwan, and Imran Rasul (2005): "Social Preferences and the Response to Incentives: Evidence From Personnel Data", *The Quarterly Journal of Economics*, 120, 917-62.

[4] Bandiera, Oriana, Barankay, Iwan, and Imran Rasul (2009): "Social Connections and Incentives in the Workplace: Evidence From Personnel Data", *Econometrica*, 77:4, 1047-94.

[5] Bandiera, Oriana, Barankay, Iwan, and Imran Rasul (2010): "Social incentives in the workplace", *Review of Economic Studies*, 77:2, 417-458.

[6] Bandiera, Oriana, Barankay, Iwan, and Imran Rasul (2013): "Team Incentives: Evidence From a Firm Level Experiment", *Journal of the European Economic Association* 11:5.

[7] Burke, Mary A., and Tim R. Sass (2013): "Classroom peer effects and student achievement", *Journal of Labor Economics*, 31:1, 51-82.

[8] Cornelissen, Thomas, Dustmann, Christian, and Uta Schnberg (2013): "Peer Effects in the Workplace", working paper.

[9] Falk, Armin, and Andrea Ichino (2006): "Clean evidence on peer effects", *Journal of Labor Economics*, 24:1, 39-57.

[10] Gould, Eric D., and Eyal Winter (2009): "Interactions between workers and the technology of production: Evidence from professional baseball", *Review of Economics and Statistics*, 91(1): 188-200.

[11] Guryan, Jonathan, Kroft, Kory, and Matthew J. Notowidigdo (2009): "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments", *American Economic Journal: Applied Economics*, 1:4, 34-68.

[12] Hamilton, Barton H., Nickerson, Jack A., and Hideo Owan (2003): "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation", *Journal of Political Economy*, 111:3, 465-497.

[13] Mas, Alexandre and Enrico Moretti (2009): "Peers at Work", *American Economic Review*, 99:1, 112-145.

[14] Price, Joseph and Justin Wolfers (2010): "Racial Discrimination Among NBA Referees", *The Quarterly Journal of Economics*, 125:4, 1859-1887.

[15] Waldinger, Fabian (2010): "Quality Matters: The Expulsion of Professors and the Consequenses for PhD Student Outcomes in Nazi Germany", *Journal of Political Economy*, 118:4, 787-831.

[16] Waldinger, Fabian (2012): "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany", *Review of Economic Studies*, 79, 838-861.

A Data Appendix

There were a number of intermediate steps required to transform the raw play-by-play data that we gathered from espn.com into our final dataset. The first step was to determine which players were on the court at each point during the game. Since the play-by-play data does not provide a running list of who is on the court over the course of the game, we have to infer who is on the court based on the players we observe in the data.

Teams can freely substitute players at the start of each quarter and none of these substitutions appear in the data. However, any substitutions that occur during the quarter, we observe both the players coming in and the players going out. We combine this with information on the names of the players that record some action in the data to construct the set of five players on each time for every play-level observation in the data.

The second step is to transform our play-level data into a single observation for each possession. An offensive possession begins anytime a team obtains the ball and switches from defense to offense. Possessions can end in many ways, such as a made basket, a missed shot, a turnover, or the end of a quarter. For possession that result in positive points, we also capture which player on the offensive team scores and how many points they scored. For this study, any play by the offensive team that extends a possession (such as an offensive rebound) does not create new possession but just becomes the continuation of the possession already going. The one exception is when an offensive rebound occurs fouling foul shots, since it is very common for substitutions to occur during foul shots.

If in the middle of a possession there is a substitution, the player entering the game is the one considered on the court for that possession. The one exception is substitutions that occur during fouls shots in which case the players coming out are considered part of the possession that resulted in foul and any points scored from the foul shots are credited to that possession.

From the defensive standpoint, the only relevant outcome is whether the offensive teams scores positive points. Steals, blocks, and defensive rebounds will get reflected in an increased probability that the offensive team does not score.

At the end of this process we are left with 915,580 unique possessions. We drop about 1% of these possessions either because we could not identify all of the players on the court or identify the player who shot the basket. A possession that either has too few players on the court or a player on the court more than once typically indicates a data entry error in the play-by-play data. Often this implies that active lineups for other possessions during that quarter are likely to be incorrect. As a result, any quarter that has a possession either too few or too many players is dropped.

Finally, our empirical strategy requires us to estimate the model separately by season. If during a season a player never scores nor is ever part of a defensive unit that keeps the other team from scoring, then that player's offensive and defensive parameters are not identified. Thus, we identify who these players are, and then eliminate all possessions during which these players are on the court. Typically there are about five to ten players per season who fall into this category.

Table 1: Sample Statistics

Seasons Covered	2006-2007 through 2009-2010
Total Possessions (Involving 5 Offensive and Defensive Players)	915,580
Utilized Possessions	905,378
Fraction of Possessions Discarded	0.01
Unique Players	656
Average Possessions Per Player	13,801
SD Possessions Per Player	12,882
25th Percentile Of Possession Distribution	2,342
75th Percentile Of Possession Distribution	22,609
Average Possessions Per Player-Season	5,081
SD Possessions Per Player-Season	3,557
25th Percentile Of Possession Distribution	1,697
75th Percentile Of Possession Distribution	8,083
Average Possessions Per Player-Season-Team	4,507
SD Possessions Per Player-Season-Team	3,570
25th Percentile Of Possession Distribution	1,130
75th Percentile Of Possession Distribution	7,470
Proportion of Possessions with Positive Points	50.8
Avg. Points Per Possession	1.06
SD Points Per Possession	1.11
Avg. Points Per Possession | Points>0	2.1

Table 2: Skills and Winning, 2009-2010

Take the 4 most utilized players on each team and...	Point Differential	SD Point Differential	Minimum	Maximum
Add average player	0.0048	0.0873	-0.1536	0.257

	Δ Point Differential	SD ΔPoint Differential	Minimum	Maximum
Add 1 SD Better Intercept Player	0.0320	0.0026	0.0257	0.0363
Add 1 SD Worse Intercept Player	-0.0248	0.0020	-0.0288	-0.0203
Add 1 SD Better Spillover Player	0.0237	0.0017	0.0206	0.0285
Add 1 SD Worse Spillover Player	-0.0234	0.0016	-0.0280	-0.0205
Add 1 SD Better Defensive Player	0.0346	0.0004	0.0329	0.0348
Add 1 SD Worse Defensive Player	-0.0345	0.0003	-0.0348	-0.0334

Unit of observation is an NBA team in the 2009-2010 season.

Table 3: Average Skills by Position

	Point Guard	Shooting Guard	Small Forward	Power Foward	Center
	(All measures standardized at the population level)				
Offensive Intercept	-0.147	0.081	-0.013	0.069	0.006
Offensive Spillover	0.168	-0.091	-0.012	-0.073	0.012
Defensive Intercept	-0.511	-0.305	-0.097	0.315	0.683
Overall Rank	-0.298	-0.108	-0.016	0.134	0.327
PER	-0.039	-0.112	-0.107	0.065	0.218
APM	-0.101	0.039	0.057	0.015	-0.011
Observations	395	410	349	438	417

Unit of observation is a player-season-team combination. Means are constructed by weighting the total number of possessions for a player-season-team combination. Across all positions the mean of each measure is zero with a standard deviation of 1. Note that the number of observations for the APM measure is smaller since players with few possessions are not assigned a rating.

Table 4: Player Rankings based on Standardized Point Differentials, 2009-2010

Point Differential w/ Average Team				Δ in Team Point Differential					
Preferred		Position Adjusted		when Replaced with Average		PER		APM	
Dwight Howard	3.195	Dwyane Wade	3.017	Dwight Howard	3.266	LeBron James	3.866	Dwyane Wade	3.873
Kevin Durant	2.907	Deron Williams	2.889	Kevin Durant	2.807	Dwyane Wade	3.119	LeBron James	3.657
Dwyane Wade	2.629	Kevin Durant	2.772	Dwyane Wade	2.484	Kevin Durant	2.686	Dwight Howard	3.283
Al Horford	2.501	Dwight Howard	2.648	Al Horford	2.283	Chris Bosh	2.396	Steve Nash	2.839
Deron Williams	2.496	Andre Miller	2.267	Deron Williams	2.256	Tim Duncan	2.324	Chris Andersen	2.397
LeBron James	2.187	Dirk Nowitzki	2.074	Dirk Nowitzki	2.010	Dwight Howard	2.156	Ray Allen	2.291
Dirk Nowitzki	2.029	LeBron James	2.062	LeBron James	1.968	Chris Paul	2.083	Kobe Bryant	2.243
Marc Gasol	1.950	Al Horford	2.023	Marc Gasol	1.909	Dirk Nowitzki	1.891	Kevin Durant	2.136
Chris Andersen	1.915	Chris Andersen	1.958	Andre Miller	1.874	Pau Gasol	1.891	Deron Williams	2.045
Andre Miller	1.897	Stephen Jackson	1.728	Chuck Hayes	1.542	Amar'e Stoudemire	1.818	Nene Hilario	1.935

Table 5: Player Rankings Ignoring Spillovers, 2009-2010

Point Differential w/ Average Team		Spillover Differential - No Spillover Differential: Notable Δ's		Largest Positive Δ's		Largest Negative Δ's	
Dwight Howard	3.126	Carmelo Anthony	-1.287	Jason Williams	2.788	James Singleton	-1.898
Kevin Durant	2.828	Vince Carter	-1.083	Jon Brockman	2.145	Andray Blatche	-1.781
Andrew Bogut	2.345	Allen Iverson	-1.086	Jamaal Tinsley	2.128	Eric Maynor	-1.518
LeBron James	2.334	Tim Duncan	-1.033	Mike Miller	2.062	Chris Kaman	-1.499
Tim Duncan	2.121	Kobe Bryant	-0.963	Theo Ratliff	1.704	Mo Williams	-1.481
Vince Carter	2.092	Shane Battier	1.036	Chris Andersen	1.703	Luis Scola	-1.329
Kobe Bryant	2.071	Deron Williams	1.130	Anderson Varejao	1.617	Sam Young	-1.306
Dwyane Wade	2.035	Al Horford	1.148	DeShawn Stevenson	1.598	Carmelo Anthony	-1.287
Chris Bosh	1.882	Steve Nash	1.334	Shaun Livingston	1.480	Sergio Rodriguez	-1.245
Dirk Nowitizki	1.855	Chris Andersen	1.703	Earl Watson	1.375	Mario Chalmers	-1.211

Table 6: LeBron James and "The Decision"

If LeBron joins ...	Projected Teammates	Probability LeBron Scores per Possession	Team per Possession Points	Point Differential
Chicago Bulls	Boozer, Deng, Noah, Rose	0.175	1.204	0.136
Cleveland Cavaliers	Williams, Hickson, Parker, Varejao	0.187	1.142	0.043
Miami Heat	Chalmers, Bosh, Ilgauskas, Wade	0.165	1.244	0.219
New York Knicks	Chandler, Felton, Gallinari, Stoudemire	0.172	1.155	0.126

To generate predicted outcomes for each team we utilize player estimates based on all four years of data.

Table 7: Amar'e Stoudemire's Free Agency

	Lineup	Amar'e Points Per Possession	Team per Possession Points	Point Differential	Δ in Team Differential w/ Amar'e
If Knicks sign Amar'e...	Amar'e, Chandler, Felton, Gallinari, Jeffries	0.352	1.088	0.068	0.026
If Knicks don't sign Amar'e...	Average Player, Chandler, Felton, Gallinari, Jeffries		1.066	0.042	
If Suns sign Amar'e...	Amar'e, Nash, Hill, Frye, Richardson	0.326	1.205	0.102	0.016
If Suns don't sign Amar'e...	Average Player, Nash, Hill, Frye, Richardson		1.193	0.086	

To generate the predicted outcomes for each team we utilize player estimates based on all four years of data.

Table 8: Skills and Wages, Contemporaneously

Dependent Variable: Log Earnings

Offensive Intercept	0.373*	0.374*	0.343*		
	(0.023)	(0.023)	(0.016)		
Offensive Slope	-0.004	-0.004	-0.031		
	(0.023)	(0.023)	(0.016)		
Defensive Intercept	0.169*	0.158*	0.078*		
	(0.022)	(0.024)	(0.018)		
Player Efficiency Rating				0.367*	
				(0.014)	
APM					0.208*
					(0.018)
Experience			0.359*	0.354*	0.366*
			(0.016)	(0.016)	(0.019)
Experience^2			-0.016*	-0.016*	-0.018*
			(0.001)	(0.001)	(0.002)
Position Effects	N	Y	Y	Y	Y
R^2	0.178	0.18	0.604	0.611	0.521
N	2009	2009	2009	2009	1539

Unit of observation is a player-season-team combination. Robust standard errors in parentheses. * Indicates a coefficient that is statistically significant at a 5% level. Coefficients are estimated by OLS weighting each observation by the total number of possessions associated with that player-season-team combination. All skill measures are standardized to have a mean of zero and a variance of one.

Table 9: Skills and Wages over Career

Log Total Earnings 2007-2010					
Average Offensive Intercept	0.525*	0.522*	0.461*		
	(0.044)	(0.044)	(0.028)		
Average Offensive Slope	0.120*	0.116*	0.064*		
	(0.045)	(0.044)	(0.027)		
Average Defensive Intercept	0.213*	0.220*	0.101*		
	(0.039)	(0.048)	(0.031)		
Average Player Efficiency Rating				0.480*	
				(0.025)	
Average APM					0.291*
					(0.028)
Average Experience			0.503*	0.498*	0.500*
			(0.027)	(0.026)	(0.034)
Average Experience^2			-0.027*	-0.027*	-0.028*
			(0.002)	(0.002)	(0.003)
Position Effects	N	Y	Y	Y	Y
R^2	0.248	0.258	0.708	0.734	0.616
N	656	656	656	656	494

Unit of observation is a player. Robust standard errors in parentheses. * Indicates a coefficient that is statistically significant at a 1% level. Coefficients are estimated by OLS weighting each observation by the total number of possessions associated with that player over all four seasons. Average skill measures are also constructed as a possession weighted average of the player-season-team measures. The average skill measures are then standardized to have a mean of zero and a variance of one.