Network Competition and Team Chemistry in the NBA

William C. Horrace*, Hyunseok Jung and Shane Sanders

1/07/2018

Abstract

We consider a heterogeneous social interaction model where agents interact with peers within their own network but also interact with agents across other (non-peer) networks. To address potential endogeneity in the networks, we assume that each network has a central planner who makes strategic network decisions based on observable and unobservable characteristics of the agents in her charge. The model forms a simultaneous equation system that can be estimated by Quasi-Maximum Likelihood or Generalized Method of Moments. We apply a restricted version of our model to data on NBA games, where agents are players, networks are individual teams organized by coaches, and at any time a player only interacts with two networks: their team and the opposing team. We find significant positive peer-effects (team chemistry) in NBA games.

Keyword Spatial Autoregressive model, Peer effects, Selectivity, Endogenous spatial weighting matrix.

*Corresponding author: whorrace@syr.edu. All authors: Syracuse University, Syracuse, NY 13244

1 Introduction

We consider a world with R independent networks where agents interact with peers within their own network but also interact with non-peers from other networks, but in different ways. For example, we can think of teams of individual agents that cooperate within their network but compete across networks. Competition between two or more firm R&D alliances comes to mind. A given firm may cooperate with an R&D ally to achieve an intellectual property discovery, but firms across alliances compete. Airline alliances (e.g., SkyTeam, Star Alliance and OneWorld) cooperate within their networks but compete across networks. In these examples, multiple networks or teams may simultaneously compete, but in some instances, such as sports, team competition is head-to-head. We restrict attention to models where an agent's single outcome (e.g., sales performance) is a function of the simultaneous outcomes of their peers and competitors. In particular, we are not concerned with the case of Liu (2014) or Cohen-Cole et al. (2017), where there is a single peer network (no competitors) with multiple outcome variables (e.g., a single network of friends each of whom allocates effort to simultaneous outcomes, such as labor and leisure hours). In the aforementioned examples, within-network interaction is complementary and cross-network interaction is competitive; however, our model allows for the converse to be true. For example, in sports competition a team's performance may be worsened or enhanced when they face a better team.

In these examples, social interaction decisions are likely to be guided by a central planner for each network (e.g., a sales mananger), and the choices of the planner may induce what Manski (1993) calls a correlated effect, where "individuals in the same group tend to behave similarly because they... face similar institutional environments." The usual solution to the correlated effects problem is to include a network-level fixed or random effect in the model specification. However, if the network consists of labor inputs to a production process, then the network itself may be endogenous in the same way that any production input may be simultaneously (and strategically) selected with output in a manager's profit maximization problem (Olley and Pakes, 1996). Following Horrace, Liu, and Patacchini (2016), we augment the outcome equation with a selection equation that models the decisions of the central planners' network choices. We consider both parametric (Lee, 1983) and semi-parametric (Dahl, 2002) approaches to the selection problem. Horrace, Liu, and Patacchini (2016) consider a network production function where a manager selects workers into a network to produce output, but they ignore cross-network competition. In this sense, our paper is a generalization of their study.

Social network interactions[1] have been studied extensively in recent decades; however, simultaneous cross-network interactions remain relatively unexplored. A few papers model simultaneous activity for a single network, and are multivariate extensions of the single equation Spatial Auto-Regressive (SAR) model of Cliff and Ord (1973, 1981) to simultaneously determined outcome variables. For example, Kelejian and Prucha (2004) generalize the SAR model to a simultaneous system. Baltagi and Deng (2015) extend the model to a panel setting with random effects, while Cohen et al. (2017) extend it to a simultaneous system with fixed effects. Yang and Lee (2017) study identification and Quasi-Maximum Likelihood (QLM) estimation of the model of Kelejian and Prucha (2004). Empirical implementation of these types of simultaneous models include the effect of peer networks on migration and housing prices (Jeanty et al., 2010); on migration, employment and income (Gebremariam et al., 2011); on rents for studio, one-bedroom and two-bedroom apartments (Baltagi and Bresson, 2011); on simultaneous fiscal policies (Allers and Elhorst, 2011); among others. All these models are clearly related, but they don't consider multiple peer networks that may be engaged in simultaneous competition around a single outcome variable.

Aside from the applied econometric contributions mentioned above, another contribution of the study is empirical in nature. We apply our model to the 2013-14, 2014-15, and 2015-16 NBA regular seasons to estimate within-network and cross-network peer-effects for each team in the league. Team chemistry (team peer-effects) receives substantial attention as a factor influencing team performance in business and sports. Unfortunately, the concept is notoriously difficult to measure. Schrage (2014) discusses team chemistry measurement as the new holy grail of performance analytics in sport and business. McCaffery and Bickart (2013) estimate team chemistry as a function of biological synchrony among players, and Kraus, Huang, and Keltner (2010) find evidence that early-season, on-court tactile communication is a predictor of later-season success at both the individual and team levels. Horrace, Liu, and Patacchini (2016) develop a network production function model that estimates peer (on-court teammate) network effects upon player productivity in men's college basketball. Due to data limitations, their peer-effects do not condition upon (confounding) effects from the network of on-court competitors. The absence of competitor effects in their model introduces omitted variable bias, and their peer-effects do not constitute (*ceteris paribus*) estimates of team-level chemistry.

We extend their peer-effect model to account for the strategic decisions and contemporaneous play of the opposing team by augmenting it with a competitor network, leading to estimation of a team's

[1]See Manski (1993), Moffitt (2001), Lee (2007a), and Bramoulle et al (2009)

"competitor-effect" in addition to its peer-effect. The interpretation of the team-level peer-effect is the same, but, in our model, the effect conditions on both teams' strategies and abilities, making for a more reliable team-chemistry measure. We find that peer-effects are generally positive for NBA teams and are moderately persistent across team-season but that negative peer-effects occur.

The rest of this paper is organized as follows. The next section introduces the econometric specification and the estimation approaches, Section 3 provides the result of the empirical exercise and Section 4 concludes.

2 Model and estimation

2.1 Econometric model

2.1.1 Outcome function

We have R networks (alliances, chains or teams), and each network, $r = 1, ..., R$ contains n_r peers with $N = \sum_r n_r$. Peers cooperate within their own network but compete with members of the other networks. The time period, t, is suppressed here. In the model that follows, all the data, weighting matrices, and the error term vary with time, and all the parameters do not. The outcome function for the r^{th} network is:

$$y_r = \lambda_{rr} W_{rr} y_r + \sum_{k \neq i} \lambda_{rk} W_{rk} y_k + x_r \beta_r + u_r, r = 1, ..., R, \tag{1}$$

where y_r is an $n_r \times 1$ outcome vector for the r^{th} network, x_r is a $n_r \times p$ exogenous input matrix, and u_r is an $n_r \times 1$ disturbance vector. W_{rr} is an $n_r \times n_r$ weight matrix for interaction within the r^{th} network, while W_{rk} for $k \neq r$ is a $n_r \times n_k$ matrix for the effect from the k^{th} network to the r^{th} network. We assume the matrices have network structure and are row-normalized, so that λ_{rr} is the average within-network effect for the r^{th} network, and λ_{rk} is the average cross-network effect for $k \neq r$ to the r^{th} network. The term β_r represents a vector of input coefficients for the r^{th} network. The existing literature assumes that $\lambda_{rk} = 0$ for all $k \neq r$[2]. We allow the within-network and cross-network effects to be positive or negative, but for convenience we will refer to W_{rr} as the peer network and W_{rk} as the competitor network.

There are other interpretations of the model. For example, we can think of the index r as representing distinct markets, where the networks might compete. If $\lambda_{rk} \neq 0$ for all r, k, then all networks

[2] When $\lambda_{rk} = 0$ for all $k \neq r$, the model reduces to the Horrace, Liu, and Patacchini (2016) model.

compete in all markets. If any network r does not compete in market $k \neq r$, then $\lambda_{rk} = 0$. In our application to the NBA, networks are individual teams, and in any game there are only ever two networks at a time on the right-hand side of the model: the peer network and the opposing teams network. We can also think of the model being for a single firm with different simultaneous projects $r = 1, ..., R$. In this case, there might only be one manager making decisions for each project (network), but that presents no additional difficulties in what follows.

Then, the outcome function of the system of R equations in a matrix form is

$$Y = \sum_{r=1}^{R} \sum_{k=1}^{R} \lambda_{rk} G_{rk} Y + XB + U \tag{2}$$

where $Y = (y_1^{'}, ..., y_R^{'})^{'}, X = Diag(x_1^{'}, ..., x_R^{'})^{'}, U = (u_1^{'}, ..., u_R^{'})^{'}$ and $B = (\beta_1^{'}, ..., \beta_R^{'})^{'}$. G_{rk} is an $N \times N$ block matrix with R row blocks and R column blocks. The blocks in G_{rk} are all blocks of zeros except for r^{th} row block in the k^{th} column block position, which equals W_{rk}. For example, if there are $R = 2$ networks, then $G_{11} = \left[\begin{smallmatrix} W_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right]$, $G_{22} = \left[\begin{smallmatrix} 0 & 0 \\ 0 & W_{22} \end{smallmatrix}\right]$, $G_{12} = \left[\begin{smallmatrix} 0 & W_{12} \\ 0 & 0 \end{smallmatrix}\right]$, and $G_{21} = \left[\begin{smallmatrix} 0 & 0 \\ W_{21} & 0 \end{smallmatrix}\right]$ where 0 is a conformable matrix of zeros. If all the networks are the same size $n_r = n$, then $G_{rk} = g_{rk} \otimes W_{rk}$, where g_{rk} is an $R \times R$ matrix of zeros except for a '1' in the r^{th} row and $k^{th} column$. In our NBA application, all the networks are the same size (5×5) with each row representing an active player. This representation is similar to the higher-order SAR model of Lee and Liu (2010) but without autoregressive errors. We assume that $I - \sum_{r=1}^{R} \sum_{k=1}^{R} \lambda_{rk} G_{rk}$ is invertible so that the model is in equilibrium. Then, the reduced form of equation (2) is $Y = (I - \sum_{r=1}^{R} \sum_{k=1}^{R} \lambda_{rk} G_{rk})^{-1}(XB + U)$.

The model forms a simultaneous SAR equation system (or network model with heterogeneous interaction effects) where the simultaneities are coming not only from the peer network but also from the competitor networks. Therefore, this model can be viewed as an extension of a single equation SAR model to a simultaneous system of multiple SAR equation models. Similar simultaneous models exist in the literature, but these models are meant for simultaneous outcome variables, rather than for a single outcome variable with simultaneous networks in a competitive setting. Kelejian and Prucha (2004) generalize the single equation SAR model due to Cliff and Ord (1973, 1981) to a system of spatially interrelated cross sectional equations, and they consider two stage least square (2SLS) and three stage least square (3SLS) estimation methods. Baltagi and Deng (2015) extend their model to a panel setting with random effects, while Cohen et al. (2017) extend it to a simultaneous equation network model with network fixed effects. They both consider 3SLS to estimate their models. Yang and Lee (2017) study identification and QML estimation for the model of Kelejian and Prucha (2004). All of these models focus on the cases where the cross interaction effects work through a single network or non-network structure, while our model contains R^2 different network matrices, allowing for a rich set of heterogeneous network interactions.[3] For empirical studies using the simultaneous SAR model, see Jeanty et al. (2010), Gebremariam et al. (2011), Baltagi and Bresson (2011), or Allers and Elhorst (2011).

Horrace et al. (2016) considers a single-equation version of this model where $R = 1$, implying no cross-network interactions. To understand the effect of ignoring the cross-networks, suppose that the correct model has $R = 2$, but we estimate the $R = 1$ model. For simplicity, assume that $\lambda_{11} = \lambda_{22} = \lambda_1$, $\lambda_{12} = \lambda_{21} = \lambda_2$, $\beta_1 = \beta_2 = \beta$, and let $G_1^* = G_{11} + G_{22}$, and $G_2^* = G_{12} + G_{21}$. Suppose we assume that $E(U|\{G_{rk}\}, X) = 0$. Consider estimating the model omitting the cross network effects using 2SLS with a set of instrumental variables $Z = [G_1^* X, X]$ without $G_2^* X$. Then, $\hat{\theta} = (\hat{\lambda_1}, \hat{\beta})' = (M' P_Z M)^{-1} M' P_Z Y$ where $M = [G_1^* Y, X]$ and $P_Z = Z(Z'Z)^{-1} Z'$. As the true model is $Y = M\theta + \lambda_2 G_2^* Y + U$, we have $\hat{\theta} = \theta + (M' P_Z M)^{-1} M' P_Z (\lambda_2 G_2^* Y + U)$. So $\hat{\theta} - \theta = \lim_{N\to\infty} (\bar{M}' P_Z \bar{M})^{-1} \bar{M}' P_Z \lambda_2 G_2^* \Big(I - \sum_{i=1}^{2} \lambda_i G_i^*\Big)^{-1} X\beta$ where $\bar{M} = [G_1 \Big(I - \sum_{i=1}^{2} \lambda_i G_i\Big)^{-1} X\beta, X]$. Therefore, as long as $\beta \neq 0$ and $\lambda_2 \neq 0$, the λ_1 and β will be estimated inconsistently.

2.1.2 Bias due to strategic interactions

Our model has endogenous variables, Y on the right hand-side, which will be handled by standard spatial econometrics techniques. In addition to this source of endogeneity, the managers' strategic actions may be correlated with the outcome. If this is the case, $E(U|\{G_{rk}\}, X) \neq 0$ in equation (2). To address this issue, we set up a static game to formulate and correct the bias.[4] We follow and adapt the basic methodologies in the game theory literature for static games of incomplete information with multiple equilibria. In particular, our arguments follow Bajari et al. (2010).[5]

Each network r has a network manager who takes actions from his finite and discrete choice set of strategies, $A = (1, ..., s, ...)$, on behalf of their members, and their decisions are based on two types of state variables, (Z, e_r), where Z is a vector of observable state variables (i.e., market characteristics) which are common to all the networks, and $e_r = (e_r(1), ..., e_r(s), ...)$ is r's unobservable action-specific

[3]Kelejian and Prucha (2004) include a footnote that their model can be generalized to the case with simultaneous weight matrices which are unique to each variable, however, this is still more restrictive than our network interaction specification. Baltagi and Deng (2015) use two different weighting matrices for each variable, but they don't have the cross-network effects.

[4]There may be other econometric remedies to address the endogeneity issue. Recently, two categories of methodologies have been proposed to address the endogeneity in formation of spatial or network links: One is (Bayesian) One step Full information approach by Goldsmith and Imbens (2013) and Hsieh and Lee (2016), and the other is Multiple step Control Function approach by Qu and Lee (2015) and Horrace et al. (2016).

[5]Generalizing the following model to a dynamic game is left for future research.

state variable.[6] The game proceeds as follows: First, the state variables, $(Z, \{e_r\}_{r=1}^R)$, are realized. Then, the network managers simultaneously choose their actions from their choice set. Under the chosen action, networks produce single outcome, Y.

Let the r^{th} network manager's additively separable payoff function be $\pi_r(s_r, s_{-r}, Z, e_r(s)) = \pi_r^0(s_r, s_{-r}, Z) + e_r(s)$, where $s_{-r} = (s_1, ..., s_{r-1}, s_{r+1}, ..., s_R)$, the collection of all the network managers' decisions except r. We assume error terms $\{e_r(s)\}_{r=1}^R$ are distributed *iid* across s and network r with distributions $\{G_{e_r}(\cdot)\}_{r=1}^R$. We assume that Z, $\{\pi_r^0(\cdot)\}_{r=1}^R$ and $\{G_{e_r}(\cdot)\}_{r=1}^R$ are common knowledge to the managers while the $\{e_r(s)\}$ are private information. Then, we can define r's information set as $\eta_r = \{Z, e_r, \{\pi_r^0(\cdot)\}_{r=1}^R, \{G_{e_r}(\cdot)\}_{r=1}^R\}$ and r's decision as $\rho_r(\cdot) : \eta_r \to A$. Under these settings, the conditional choice probability (CCP) of r choosing $s \in A$ at a given realization of the Z is given by

$$\delta_r(s|Z) = \int 1\{\rho_r(\eta_r) = s\} dG_{e_r}(e_r) \tag{3}$$

which can be interpreted as the beliefs formed by r's opponents regarding r's decision. Since network manager r does not know the other managers' decisions at the time of her decision, her strategy is based on her expected payoff

$$\pi_r^{ex}(s, Z, e_r(s)) = \sum_{s_{-r}} \prod_{k \neq r} \delta_k(s_k|Z) \pi_r(s, s_{-r}, Z) + e_r(s) = \varphi(s, Z) + e_r(s) \tag{4}$$

We can see that the expected payoff function is similar to the standard random utility model. The only difference is that the probability distributions over other managers' actions are affecting manager r's utility. Then, it is straightforward that s will be chosen for a given realization of Z if only if

$$\pi_r^{ex}(s, Z, e_r(s)) > \max_{s' \neq s} \pi_r^{ex}(s', Z, e_r(s')) \Leftrightarrow \varphi(s, Z) + e_r(s) > \max_{s' \neq s} \varphi(s', Z) + e_r(s') \tag{5}$$

for $s, s' \in A$. It follows immediately that $\delta_r(s|Z) = Pr(\max_{s' \neq s} \varphi(s', Z) - \varphi(s, Z) + e_r(s') - e_r(s) < 0)$ in equilibrium (e.g. Bayesian-Nash Equilibrium).[7]

Formulation of the selection bias: We formulate the bias by assuming the error terms, (u_r, e_r), from the outcome equation and the payoff function (respectively) are statistically dependent, a standard approach. If there's a correlation between the two errors, the expectation of u_r conditional on the choice

[6] Time period, t, is suppressed here. Follow Bajari et al. (2010), without loss of generality we assume that the strategies of the network managers are the same

[7] In the game literature the focus is often to estimate the payoff function, and doing so requires additional structure be imposed on the function. However, this is not the focus here, so additional structure is not necessary.

of strategy, s will not have a zero mean; that is,

$$E(u_r|s) = E(u_i| \max_{s' \neq s_i} \varphi(s^{'}, Z) - \varphi(s, Z) + e_r(s^{'}) - e_r(s) < 0) \neq 0 \tag{6}$$

This correlation may exist when the two errors contain a common component, unobserved by the econometrician (e.g., a network-strategy specific fixed effect in the outcome equation and the strategy selection equation).

The bias in the outcome equation can be expressed parametrically or semi-parametrically following Lee (1983) or Dahl (2002).

1. **Lee's approach**: From (5), we see that strategy s will be chosen by manager r if only if $\epsilon_r^* < 0$, where $\epsilon_r^* = \max_{s' \neq s_i} \varphi(s^{'}, Z) - \varphi(s, Z) + e_r(s^{'}) - e_r(s)$. ϵ_r^* is a new random variable with a distribution of F_r. Following Lee (1983) and Horrace, Liu, and Patacchini (2016), we can reduce the dimensionality of F_r by the transformation $J_r(\cdot) \equiv \Phi^{-1}(F_r(\cdot))$, where Φ^{-1} is the inverse of the standard normal CDF. Then, $J_r(\epsilon_r^*)$ becomes a standard normal random variable. For notational simplicity, let $J_r(\epsilon_r^*) \equiv \epsilon_r$. We further assume ϵ_r and $u_{r,i}$ for $i = 1, ..., n_r$ are $i.i.d$ with a joint normal distribution,
$$\begin{bmatrix} u_{r,i} \\ \epsilon_i \end{bmatrix} = \text{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{bmatrix} \right),$$
where $u_{r,i}$ is the i^th element of the vector u_r. Then it can be shown that

$$E(u_r|s) = E(u_r|\epsilon_r < J_r(0), Z) = -\sigma_{12} \frac{\phi(\Phi^{-1}\delta_r(s|Z))}{\delta_r(s|Z)} \iota_{n_r} \tag{7}$$

where $\delta_r(s_r|Z)$ is the selection probability from (3) and ι_{n_r} is an n_r-dimensional vector of 1's.

2. **Dahl's approach**: Without a parametric distributional assumption on u_r and ϵ_r, we make the index sufficiency assumption of Dahl (2002) such that

$$h(\epsilon_r, u_r | \{\varphi(s, Z)\}_{s \in A}) = h(\epsilon_r, u_r | \delta_r(s|Z)) \tag{8}$$

where $h(\cdot, \cdot)$ is some bivariate distribution for u_r and ϵ_r. As Dahl points out, this assumes that the selection probability $\delta_r(s|Z)$ exhausts all the information about the behaviors of the two errors. Then, the bias will be given by $E(u_r|s) = \psi_r(\delta_r(s|Z))$, where $\psi_r(\cdot)$ is an unknown function that can be estimated nonparametrically.

With these approaches, we can identify the bias in the outcome equation by rewriting (1) as

$$y_r = \lambda_{rr} W_{rr} y_r + \sum_{k \neq r} \lambda_{r,k} W_{r,k} y_k + x_r \beta_r + a_r \iota_{n_r} + u_r^* \tag{9}$$

where $a_r \iota_{n_r} = E(u_r|s)$ are network specific fixed-effects due to the strategic actions of the network managers, and $u_r^* = u_r - E(u_r|s)$ with zero mean by construction, so the endogeneity disappears in the outcome equations conditional on a_r.[8]

2.2 Estimation

We have shown that the strategic bias can be reduced to a network specific fixed effect. If our primary interest is just to estimate network effects, it is sufficient to use the within transformation around each network to remove the network specific fixed effects, and, then, apply QML or GMM to estimate the model. We will illustrate this case in this section. However, if there are network invariant regressors (e.g., network specific characteristic) and we are also interested in estimating the strategic bias, we have to consider one more step as the within transformation eliminates them in the first step. We will briefly discuss the estimation procedure for this case at the end of this section.[9]

Here, we focus on QML to estimate the model and provide a sketch of the GMM method using the $R = 2$ networks case in the Appendix A.[10] After accounting for the strategic bias, the complete system is

$$Y = \sum_{r=1}^{R} \sum_{k=1}^{R} \lambda_{rk} G_{rk} Y + XB + A + U^* \tag{10}$$

where $A = (a_1 \iota_{n_1}^{'}, ..., a_R \iota_{n_R}^{'})^{'}$ and $U^* = (u_1^{*'}, ..., u_R^{*'})^{'}$. To remove the bias term A and avoid the incidental parameter problem (Neyman and Scott, 1948), we transform with projector $J_Q = Diag(Q_1, ..., Q_R)$

[8] The conceptual foundations for correcting endogeneity in this way for a social interactions model can be traced to a series of papers by Brock and Durlauf (2001, 2002, 2006).

[9] In our NBA application, we only estimate network peer-effects and marginal effects for network varying regressors.

[10] One may also consider GS2SLS or GS3SLS to estimate the models (Kelejian and Prucha, 2004; Lee, 2003). However, in spite of the simplicity of the 2SLS methods, they have some limitations. As discussed in Lee (2001a) and Lee (2007b), 2SLS and 3SLS are inefficient relative to the MLE, and also 2SLS may not be applicable when all the exogenous regressors in a model are really irrelevant in explaining the endogenous variables. For these reasons, we consider the QML and the GMM to estimate our models, which are computationally more challenging but more efficient than the others because they exploit not only linear but also quadratic moment conditions. This is especially effective when the linear moment conditions are weak (Lee et al, 2010b; Liu and Lee, 2010). The two methods have different advantages: QML is most efficient when the error term is normally distributed, but we have to determine the parameter space for the spatial interdependency, which could be quite complex, and we have to evaluate the Jacobian determinant in the log-likelihood, which could be intensive when the network matrices are dense. GMM, on the other hand, is generally less efficient than the MLE but does not involve the problems.

where Q_r for $r = 1, ..., R$ is the within transformation matrix, $Q_r = I_{n_r} - \frac{1}{n_r}\iota_{n_r}\iota_{n_r}^{'}$. Then, as $Q_r\iota_{n_r} = 0$ and $Q_r u_r^* = Q_r u_r$, we have

$$J_Q Y = J_Q \sum_{r=1}^{R}\sum_{k=1}^{R} \lambda_{rk} G_{rk} Y + J_Q X B + J_Q U \tag{11}$$

Extensive study of identification conditions for network models and multivariate SAR models can be found in Bramoulle et al (2009), Cohen et al (2017) and Yang and Lee (2017). In Appendix B, we discuss identification conditions for (11). The identification condition will be generally satisfied because we have multiple sets of network matrices and exogenous regressors from each group, which produces enough variation to identify the coefficients in our model.

We assume that each element $u_{r,i}$ for $i = 1, ..., n_r$ is $iid(0, \sigma_r^2)$.[11] We note that the disturbances, $Q_r u_r$, in (11) are linearly dependent because the variance matrix $\sigma_r^2 Q_r$ is singular. Following Lee et al. (2010a), we consider an equivalent but more effective transformation which can eliminate the network fixed-effects while maintaining interdependency between the disturbances. Let the orthonormal matrix of Q_r be $[P_r, \iota_{n_r}/\sqrt{n_r}]$. The columns in P_r are eigenvectors of Q_r corresponding to the eigenvalue one, such that $P_r^{'}\iota_{n_r} = 0$, $P_r^{'}P_r = I_{n_r-1}$ and $P_r P_r^{'} = Q_r$. Then, premultiplying (12) by $J_P^{'} = Diag(P_1^{'}, ..., P_R^{'})$ leads to

$$J_P^{'} Y = J_P^{'} \sum_{r=1}^{R}\sum_{k=1}^{R} \lambda_{rk} G_{rk} Y + J_P^{'} X B + J_P^{'} U \tag{12}$$

Let $\bar{Y} = J_P^{'} Y_m$, $\bar{X} = J_P^{'} X$, $\bar{U} = J_P^{'} U$, $\bar{G}_{rk} = J_P^{'} G_{rk} J_P$. Due to $J_P^{'} G_{rk} = \bar{G}_{rk} J_P^{'}$,[12] then this implies

$$\bar{Y} = \sum_{r=1}^{R}\sum_{k=1}^{R} \lambda_{rk} \bar{G}_{rk} \bar{Y} + \bar{X} B + \bar{U} \tag{13}$$

Note that $\bar{u}_r$ for $r = 1, ..., R$ is now $\bar{u}_r \sim (0, \sigma_r^2 I_{n_r-1})$. Therefore, the likelihood function[13] is

$$\ln L(\Lambda, B, \Sigma) = -\sum_{r=1}^{R} \frac{n_r - 1}{2} \ln(2\pi\sigma_r^2) + \ln|\bar{S}(\Lambda)| - \sum_{r=1}^{R} \frac{\bar{\epsilon}_r(\theta_r)^{'}\bar{\epsilon}_r(\theta_r)}{2\sigma_r^2} \tag{14}$$

where $\Lambda = (\Lambda_1^{'}, ..., \Lambda_R^{'})^{'}$ with $\Lambda_r = (\lambda_{r,1}, ..., \lambda_{r,R})$, $\Sigma = (\sigma_1^2, ..., \sigma_R^2)$, $\bar{S}(\Lambda) = I - \sum_{r=1}^{R}\sum_{k=1}^{R} \lambda_{rk}\bar{G}_{rk}$,

[11] If we assume non-normality for the error terms, then MLE is Quasi-MLE

[12] $\bar{G}_{rk} J_P^{'} = J_P^{'} G_{rk} J_Q = J_P^{'} G_{rk}(I - diag(\iota_{n_1}\iota_{n_1}^{'}/n_1, ..., \iota_{n_R}\iota_{n_R}^{'}/n_R)) = J_P^{'} G_{rk}$ because G_{rk} is row-normalized so $J_P^{'} G_{rk} diag(\iota_{n_1}\iota_{n_1}^{'}/n_1, ..., \iota_{n_R}\iota_{n_R}^{'}/n_R) = 0$.

[13] The likelihood is conditional likelihood because it is conditional on the sufficient statistic, the mean of y_r. Lee (2007a).

and $\bar{\epsilon}_r(\theta_r) = \bar{y}_r - \sum_{k=1}^{R} \lambda_{rk}\bar{W}_{rk}\bar{y}_k - \bar{x}_r\beta_r$ where $\theta_r = (\Lambda_r, \beta_r)$, and $\bar{W}_{rk}$ and $\bar{x}_r$ are defined similarly as in (13). From Lemma 1 in the Appendix C, we show $\ln|\bar{S}(\Lambda)| = -\ln f(\Lambda) + \ln|S(\Lambda)|$ where $S(\Lambda) = I - \sum_{r=1}^{R}\sum_{k=1}^{R} \lambda_{rk}G_{rk}$ and $f(\Lambda)$ is some function of Λ. For example, when $R = 2$, $f(\Lambda) = (1-\lambda_{11})(1-\lambda_{22}) - \lambda_{12}\lambda_{21}$ and when $R = 3$, $f(\Lambda) = (1-\lambda_{11})(1-\lambda_{22})(1-\lambda_{33}) - (1-\lambda_{11})\lambda_{23}\lambda_{32} - (1-\lambda_{22})\lambda_{13}\lambda_{31} - (1-\lambda_{33})\lambda_{21}\lambda_{12} - \lambda_{13}\lambda_{21}\lambda_{32}$. Using this result, we can evaluate the likelihood without P_r as

$$\ln L(\Lambda, B, \Sigma) = -\sum_{r=1}^{R} \frac{n_r - 1}{2} \ln(2\pi\sigma_r^2) - \ln f(\Lambda) + \ln|S(\Lambda)| - \sum_{r=1}^{R} \frac{\epsilon_r'(\theta_r)Q_r\epsilon_r(\theta_r)}{2\sigma_r^2} \tag{15}$$

where $\epsilon_r(\theta_r) = y_r - \sum_{k=1}^{R} \lambda_{rk}W_{rk}y_k - x_r\beta_r$. There are two things to note here. First, we need to restrict the parameter space for Λ to guarantee that $|S(\Lambda)|$ and $f(\Lambda)$ are strictly positive so that the likelihood is well defined. From Lee and Liu (2010), the parameter space for $|S(\Lambda)|$ is strictly positive, when $\sum_{r=1}^{R}\sum_{k=1}^{R} |\lambda_{rk}| < 1$, as our network matrices are row-normalized. However, Elhorst, Lacombe and Piras (2012) argue that this may be too restrictive, and suggest a stationary-region search methodology for Λ that is potentially less so, while ensuring a well-defined likelihood function. Second, it may be difficult to evaluate $|S(\Lambda)|$. The Ord (1975) eigenvalue device may used to compute the determinant; however, it may only work when the number of networks is small and all the network matrices are sparse. If the number of networks is large, then GMM may be preferred to QML, as it avoids the computational difficulties of evaluating the determinant.

To simplify estimation, we concentrate out B and Σ in (17). The QMLE of β_r and σ_r^2, given Λ_r is $\hat{\beta}_r(\Lambda_r) = (x_r'Q_rx_r)^{-1}x_r'Q_r\mu_r(\Lambda_r)$ where $\mu_r(\Lambda_r) = y_r - \sum_{k=1}^{R} \lambda_{rk}W_{rk}y_k$, and

$$\hat{\sigma_r}^2(\Lambda_r) = \frac{\epsilon_r'(\theta_r)Q_r\epsilon_r(\theta_r)}{n_r - 1} = \frac{\mu_r(\Lambda_r)'[Q_r - Q_rx_r(x_r'Q_rx_r)^{-1}x_r'Q_r]\mu_r(\Lambda_r)}{n_r - 1}. \tag{16}$$

Then the concentrated log-likelihood function in Λ is

$$\ln L^c(\Lambda) = -\sum_{r=1}^{R} \frac{n_r - 1}{2}[\ln(2\pi) + 1] - \ln f(\Lambda) + \ln|S(\Lambda)| - \sum_{r=1}^{R} \frac{n_r - 1}{2} \ln \hat{\sigma_r}^2(\Lambda_r) \tag{17}$$

Then the QMLE, $\hat{\Lambda}$, is the maximizer of the likelihood, and the QMLE of B and Σ are $\hat{\beta}_r(\hat{\Lambda}_r)$ and $\hat{\sigma_r}^2(\hat{\Lambda}_r)$ for $r = 1, ..., R$, respectively. The asymptotic distribution for these estimators can be derived from Lee et al. (2010a, Appendix B) with appropriate modifications.

Estimation of the strategic bias

The most disaggregate level of variability in this model is the individual worker-level, $i = 1, ..., n_r$, and this is the variability that ultimately identifies the peer and competitor effects, λ_{rk}. However, there may be columns in x_r that vary at the network level (or higher), and these columns will necessarily be eliminated by the within-network transformation of the data. If we are only interested in estimating the peer- and competitor-effects, this will be fine.[14] However, if we are interested in estimating the strategic bias caused by the network managers (and the coefficients on network-varying exogenous variables), we need to consider one more step as follows:

1. Estimate the selection probability $\hat{\delta}_r(s|Z)$ for $i = r, ..., R$ using a nonparametric or parametric model (kernel smoothing, local polynomial regression, or the logit model).[15]

2. Using $\hat{\Lambda}$ and $\hat{B}$ from (15), compute the residual $\hat{\upsilon}_r = \iota'_{n_r}(y_r - \sum_{k=1}^{R} \hat{\lambda}_{rk} W_{rk} y_k - x_r \hat{\beta}_r)/n_r$ for $r = 1, ..., R$. Let $\beta_{r,2} \subseteq \beta_r$ be the coefficients on network-level varying regressors, $x_{r,2} \subseteq x_r$. Then the bias and $\beta_{r,2}$ can be estimated from the OLS regression $\hat{\upsilon}_r = x_{r,2}\beta_{r,2} + \mu(\hat{\delta}_{i,m}(s|Z_)) + \xi_r$ where ξ_r is an *i.i.d.* error term and $\mu(\hat{\delta}_r(s|Z)))$ is either given by $-\sigma_{12}\frac{\phi(\Phi^{-1}(\hat{\delta}_r(s|Z)))}{\hat{\delta}_r(s|Z))}$ (Lee's approach) or some nonparametric, single-index formulation (based on Dahl's approach). See Horrace, Liu, and Patacchini (2016) for an explanation of Dahl's approach in this context.

3 Empirical application

3.1 Empirical model and variables

In this section, we apply our network competition model to NBA data for 30 teams over the 2013-14, 2014-15, and 2015-2016 regular seasons. The primitive play-by-play data were purchased and downloaded from BigDataBall.com. We then formatted the data to the player-period level, where a period represents any contiguous game period in which the same ten players are on the court. This formatting is similar to that done in the calculation of the player statistic *real plus minus.* We tabulate player box-score data to obtain *wins produced* (see, e.g., Berri, 1999) at the time-period level. The league plays 1,230 regular-season games per season (41 home games for each of 30 teams per regular season). Therefore, our data spans 3,690 regular-season games, consisting of roughly 30 time periods per game. This produces 112,204 time periods in which we observe the play of 10 players i at a time, producing a total of 1,122,040 observations. In each game a coach typically has 15 players to fill a

[14] This is our approach in the empirical exercise in the next section.
[15] Alternatively, we can simply use the cell statistics discussed in Dahl (2002).

network of five players at a time.[16] Following Horrace, Liu, and Patacchini (2016), we drop time periods less than 30 seconds and overtime periods. This results in 83,334 time periods for the league, which is equivalent to an average of about 926 periods per team-season.

Our outcome variable is *wins produced*, a continuous weighted average of individual player offensive and defensive statistics that will be defined in what follows. *Wins produced* is highly predictive of team success and is separable (i.e., measurable at the individual level). The variable is a version of the outcome variable in Horrace, Liu, and Patacchini (2016) but with different weights on the component statistics in the weighted average, as we shall see. We calculate this outcome variable for each player in each period of the data. Overall, our data and specification are designed to mimic the Horrace, Liu, and Patacchini (2016) analysis of the Syracuse Men's Basketball team.

PRODUCTION FUNCTION: With thirty teams in the league, there is scope for estimating 30^2 peer- and competitor-effects in our model. Unfortunately, the number of games between a given pair of teams is small (three or four at maximum), so there is not enough head-to-head data to estimate this many parameters. Consequently, we assume a given team's cross-network competitor-effects from the 29 other teams are the same. That is, for team r, $\lambda_{rk} = \lambda_r$ for $k \neq r$. Conceptually, our restriction on the cross-network competitor-effects is equivalent to team r playing a season long game against all the other teams in the league, consisting of subgames against individual teams (and coaches), and where at the end of each subgame, the outcome variable, y_{rt} is set to zero for each player. In other words, team r has a representative opponent, k, and opposing teams take turns being that representative opponent. Consequently, interpretation of the competitor-effect will be challenging, so we focus the analysis on the within-network effects. The unique structure of competition in this setting, requires slightly different notation than the more general model. The production function for team r and k in period t of game s is

$$\begin{aligned} y_{rts} &= \lambda_{rr} W_{rrts} y_{rts} + \lambda_r W_{rkts} y_{kts} + x_{rts}\beta_r + u_{rts} \\ y_{kts} &= \lambda_{kk} W_{kkts} y_{kts} + \lambda_k W_{krts} y_{rts} + x_{kts}\beta_k + u_{kts} \end{aligned} \tag{18}$$

where y_{rts} and y_{kts} are the 5×1 outcome vector of team r's and team k's chosen lineup in period t of game s, respectively, and W_{rrts} and W_{kkts} are the 5×5 zero diagonal and row-normalized matrices for the within-network interactions, and W_{rkts} and W_{krts} are similarly defined matrices for cross-network

[16] Understanding the effect of player injuries (or player ineligibility) on the coaches' decisions is left for future research. Sports injuries are analogous to worker absenteeism.

interactions. The x_{rts} and x_{kts} are matrices of the player-varying exogenous variables for team r and k's lineup in period t of game s, respectively.[17]. The u_{rts} and u_{kts} are 5×1 error term vectors, in which each element is assumed to be $iid\ N(0, \sigma_r^2)$ and $iid\ N(0, \sigma_k^2)$, respectively. Following the methodology in section 2.2, the log-likelihood for the system of equations (18), denoted as $\ln L_{ts}$, is,

$$\ln L_{ts} = -\sum_{h=r,k} 2\ln(2\pi\sigma_h^2) - \ln f_{ts} + \ln|S_{ts}| - \sum_{h=r,k} \frac{\epsilon_{hts}^{'}(\theta_h) Q \epsilon_{hts}(\theta_h)}{2\sigma_h^2} \tag{19}$$

where $f_{ts} = [(1-\lambda_{rr})(1-\lambda_{kk}) - \lambda_r\lambda_k]$, and $S_{ts} = I_{10} - \lambda_{rr}G_{rrts} - \lambda_r G_{rkts} - \lambda_k G_{krts} - \lambda_{kk}G_{kkts}$ with $G_{rrts} = \begin{bmatrix} W_{rrts} & 0 \\ 0 & 0 \end{bmatrix}$, $G_{rkts} = \begin{bmatrix} 0 & W_{rkts} \\ 0 & 0 \end{bmatrix}$, $G_{krts} = \begin{bmatrix} 0 & 0 \\ W_{krts} & 0 \end{bmatrix}$, $G_{kkts} = \begin{bmatrix} 0 & 0 \\ 0 & W_{kkts} \end{bmatrix}$, $Q = I_5 - \frac{1}{5}\iota_5\iota_5^{'}$, $\epsilon_{rts}(\theta_r) = y_{rts} - \lambda_{rr}W_{rrts}y_{rts} - \lambda_r W_{rkts}y_{kts} - x_{rts}\beta_r$ with $\theta_r = (\Lambda_r, \beta_r)$ and $\epsilon_{kts}(\theta_k)$ is given similarly. The log-likelihood for entire season is just the sum of the likelihood in (19) over all games, s, and time periods, t.

NETWORK: We use the same-type peer-effect weight matrix considered in Horrace, Liu, and Patacchini (2016), where "types" are the player types: *Guards* or *Forwards*, with *Forwards* including centers. That is, the same-type weight matrix is W, where $W_0 = [w_{0,ij}]$ is an adjacency matrix with $w_{0,ij} = 1$ if the i^{th} and j^{th} players are both guards or forwards. Then row-normalize W_0 so that $W1_N = 1_N$. This network specification assumes that each individual is affected only by the same type of agents in his network and the same type of agents from opposing network (exclusion restrictions).[18]

For example, let's assume the lineup for team r is $[F, F, G, F, G]^{'}$ and for team k is $[G, F, F, G, G]^{'}$ in period t of game s, where F=forward and G=guard; then, the network matrices in equation (18) are given by

$$W_{rrts} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, W_{kkts} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}, W_{rkts} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \end{bmatrix}, W_{krts} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

VARIABLES: We use the wins produced measure based on the work of sports economist David Berri (Berri, 1999; Berri et al. 2006): $y_{irts} = (0.064 * 3PT_{irts} + 0.032 * 2PT_{irts} + 0.017 * FT_{irts} +$

[17] As previously noted, the x_{rts} may contain columns that vary at the network-level (or higher), but they will be eliminated from the model with the within transformation. Their coefficients may be recovered in the estimation of the coach's strategic bias, but we ignore them in our analysis because we are only concerned with estimating the peer- and competitor-effects, which are preserved under the within transformation.

[18] Horrace, Liu, and Patacchini (2016) use both same- and cross-type weight matrices. The different-type weight matrix is W^d, where $W_0^d = [w_{0,ij}]$ is an adjacency matrix with $w_{0,ij} = 1$ if the i^{th} and j^{th} players are a guard and a forward (or vice versa). Then row-normalize W_0^d so that $W^d 1_N = 1_N$. However, they noticed that including both weight matrices into the model at the same time may lead to a multicollinearity problem. Therefore, we only include the same type weight matrix in this exercise.

$0.034 * REB_{irts} + 0.033 * STL_{irts} + 0.020 * BLK_{irts} - 0.034 * MFG_{irts} - 0.015 * MFT_{irts} - 0.034 * TO_{irts})/Mins_{irts}$, where $3PT_{irts}$, $2PT_{irts}$ and FT_{irts} are 3-point field goals made, 2-points field goals made, and free throws made, respectively, REB_{irts} is rebounds, STL_{irts} is steals, BLK_{irts} is blocks, MFG_{irts} is missed field goals, MFT_{irts} is missed free throws, TO_{irts} is turnovers, and $Mins_{irts}$ is minutes played by player i of team r in period t of game s. Wins produced per minute (or wins per minute) estimates a player's marginal win productivity based upon player-level variables related to team-winning. It represents a leading measure of NBA player production.[19] The player-varying exogenous input variables in the outcome equation (x_{rts}) are $Experience_{irts}$ and $Fatigue_{irts}$. $Experience_{irts}$ is minutes played from the start of the game to the end of period t-1, and $Fatigue_{irts}$ is minutes continuously played until the end of period t-1. We also included player dummies to control for player-specific heterogeneity. The descriptive statistics for the continuous variables are presented in the tables of Appendix E.

3.2 Estimation strategy

We estimated the model for each season after concentrating out B and Σ. However, optimizing the entire log-likelihood with respect to 60 parameters for 30 teams at the same time may not be efficient. So, we employed the estimation strategy below:

1. **Preliminary estimation**: We include the regressor for the cross team interaction, such as $W_{rkts}y_{kts}$ or $W_{krts}y_{rts}$ in (18), but do not account for the endogeneity from the regressors. That is, we assume $\lambda_r = \lambda_k = 0$ in $\ln[(1-\lambda_{rr})(1-\lambda_{kk}) - \lambda_r\lambda_k]$ and $\ln|I_{10} - \lambda_{rr}G_{rrts} - \lambda_r G_{rkts} - \lambda_k G_{krts} - \lambda_{kk}G_{kkts}|$ of (19). Then, the entire log-likelihood for 30 teams can be separated into each individual team's log-likelihood.[20] Then, we can optimize each individual team's log-likelihood separately to get initial estimates for λ_{hh} and λ_h for $h = 1, ..., 30$. We denote the initial estimates as $\hat{\lambda}_{hh}^{pre}$ and $\hat{\lambda}_h^{pre}$ for $h = 1, ..., 30$.

2. **Updating**: We update $\hat{\lambda}_{hh}^{pre}$ and $\hat{\lambda}_h^{pre}$ for $h = 1, ..., 30$ using the original entire log-likelihood, but this update is carried out team-by-team. That is, we optimize the entire log-likelihood w.r.t one of the teams' Λ given all other teams' Λ from the preliminary estimation. We continue these

[19]See www.basketball-reference.com/about/bpm.html or wagesofwins.com/how-to-calculate-wins-produced/ for discussions of wins produced. The NBA scales this statistic to the game level by multiplying by 48 minutes per game. It is typically reported at the player level but we report it the team level in Appendix E.

[20]If λ_r and λ_k are zero, it is obvious that $\ln\left[(1-\lambda_{rr})(1-\lambda_{kk}) - \lambda_r\lambda_k\right] = \ln(1-\lambda_{rr}) + \ln(1-\lambda_{kk})$. This is also true for $\ln|I_{10} - \lambda_{rr}G_{rrts} - \lambda_r G_{rkts} - \lambda_k G_{krts} - \lambda_{kk}G_{kkts}|$ because it becomes a block diagonal matrix when λ_r and λ_k are zero.

updates from the first team to the last team (the order is arbitrary).

3. **Convergence**: We iterate the set of updates for 30 teams until the difference in $\hat{\Lambda}$ between two consecutive iterations is below some threshold.[21]

3.3 (Preliminary) Estimation results

We now present preliminary results, focusing only on the peer-effect estimates. (Estimation results for the complete system will be added in a subsequent revision.) Estimation results for the outcome equation can be found in Appendix F. We summarize the main results on team-level peer-effects in Table 1.

Peer-effects measure team chemistry conditional on strategies, abilities and opposition and do not measure team quality. Like a talented shooter can play well even with sub-optimal shot selection, a talented team can perform well even given low peer-effects. Table 1 contains the ranked peer-effects for 30 NBA teams in each of three seasons. Bounded on the unit-circle, a peer-effect close to 1 (-1) indicates good (poor) conditional team chemistry, as player performance is positively (negatively) linked to average teammate performance. Consider the results in Table 1 on the 2013-14 season, where Utah (UTA) had the largest peer-effect of 0.045. That is, when the team's average "wins produced" increases, the team gains an additional 4.5% by virtue of its good chemistry, conditional on coaching strategy and other environmental and performance variables. In this sense the peer-effect is like an output multiplier, and teams with large positive values benefit from their own team chemistry. In the 2013-14 season Utah had an average wins produced per minute of 0.0049 with a standard deviation of 0.0248. (See Appendix E, Table 2.) The distribution of this is in Figure 1. This is the distribution over active Utah players for all time periods in the 2013-14 season. Utah's 2013-14 "wins produced" is highly variable with a small mean, as the team a) played in 1,820 sampled time periods (about twice the sample average for a team-season) with an average duration of 1.8 minutes during the season and b) did not win many games overall. Over short spans, player, time-period level *wins produced* can be highly variable and sometimes quite small (non-pivotal) due to many zeros in the box score. This is reflected in the figure, where there are modes at "wins produced" of about -0.02, 0 and 0.02. During this season, Houston and San Antonio were tied for the highest average wins produced of 0.0071. Were Utah able to increase its average per period "wins produced" by 0.0022 to that of Houston, Utah's team chemistry would produce a slightly larger increase of $0.0022 * 1.045) = 0.0023$ to a wins produced

[21] In our exercise, we use a criterion of $||\hat{\Lambda} - \hat{\Lambda}^0||_2 < 10^{-8}$ where $\hat{\Lambda}$ is an estimate for Λ from the current iteration and $\hat{\Lambda}^0$ is the estimate from the previous iteration.

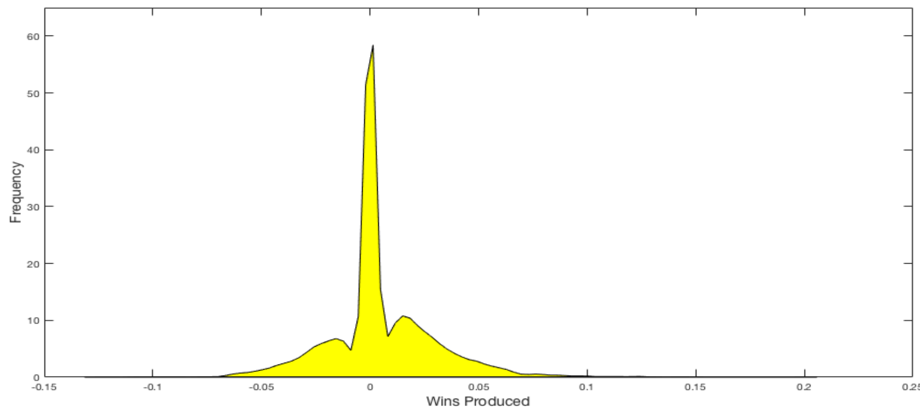
Table 1: Ranked Peer-Effects Based on Wins Produced by Season

	2013-14		**2014-15**		**2015-16**	
Rank	Team	Peer-effect	Team	Peer-effect	Team	Peer-effect
1	UTA	0.045***	IND	0.054***	CHI	0.065***
2	BOS	0.044***	PHX	0.033**	SAS	0.052***
3	PHX	0.041***	UTA	0.032***	MIN	0.049***
4	WAS	0.031**	PHI	0.023*	MIL	0.046***
5	IND	0.027**	MIN	0.020*	MEM	0.045***
6	MIN	0.026**	CLE	0.018	ORL	0.037***
7	DAL	0.022*	WAS	0.018	BKN	0.035***
8	OKC	0.021*	LAC	0.017	DET	0.030**
9	SAS	0.021*	ORL	0.014	SAC	0.028**
10	TOR	0.020*	SAC	0.011	POR	0.027**
11	BKN	0.019	SAS	0.009	BOS	0.022*
12	CLE	0.018	MEM	0.008	TOR	0.020*
13	PHI	0.017	BKN	0.007	UTA	0.016
14	NOP	0.017	NOP	0.006	CLE	0.015
15	DET	0.015	DEN	0.002	HOU	0.014
16	MIL	0.014	TOR	0.002	CHA	0.011
17	LAL	0.013	CHA	0.001	ATL	0.008
18	NYK	0.009	DAL	0.001	PHX	0.004
19	DEN	0.009	MIL	0.001	LAL	0.001
20	GSW	0.007	ATL	0.000	WAS	0.000
21	MEM	0.006	HOU	-0.001	NYK	0.000
22	CHA	0.004	LAL	-0.002	OKC	-0.002
23	SAC	0.000	POR	-0.003	LAC	-0.005
24	HOU	-0.002	CHI	-0.004	MIA	-0.007
25	ORL	-0.004	MIA	-0.005	NOP	-0.011
26	ATL	-0.005	DET	-0.006	PHI	-0.014
27	LAC	-0.009	OKC	-0.007	GSW	-0.014
28	POR	-0.010	GSW	-0.022**	DAL	-0.016
29	CHI	-0.017	NYK	-0.025**	IND	-0.017
30	MIA	-0.019*	BOS	-0.026**	DEN	-0.028**

* - significant at the 1% level. ** - significant at the 5% level.
*** - significant at the 10% level.

of 0.0072, ceteris paribus. In Table 1, most team-seasons (64 of 90) exhibit positive estimated peer-effects, but negative peer-effects occur. Median team peer-effects for each season are .015,.002, and .014. Significantly positive (negative) peer effects occur for 19 (4) of 90 team-seasons (at the 0.05 significance level). Peers may inhibit each other's marginal win product in the present environment. However, most of the teams do not exhibit significant team chemistry. In 89 of 90 team-seasons, peer-effect estimates are smaller than those of Horrace, Liu, and Patacchini (2016) in their analysis of Syracuse University Men's college basketball. This prevailing difference between sample peer-effects could be due to random variation, difference of environment, or specification of competitor-effects

Figure 1: Wins Produced per Minute per Player for Utah Jazz, 2013-14



within the present study. If it is due to competitor effects, then it seems that ignoring them causes the peer-effect to be biased.

Despite roster turnover, several teams exhibit persistence in peer-effect across seasons. Minnesota held a significantly positive peer-effect for each of the three sampled seasons. Boston, San Antonio, Utah, and Phoenix held significantly positive peer-effects in two seasons. Based on Spearman's rho, the Western Conference exhibits greater peer-effects persistence from season-to-season. If ρ_{ij} is the rank correlation of peer-effects between season i and season j for all teams in a conference, then the Western conference had significant (at the 95% level) correlations of $\rho_{12} = \rho_{23} = 0.404$, while the same statistics for the Easter conference were $\rho_{12} = 0.332$ and $\rho_{23} = -0.289$, significant at the 95% and 90% levels, respectively.

As Schrage (2014) suggests, team peer-effects are the "holy grail" of productivity analysis in sports. As we have observed, the measurement of peer-effects requires accounting for player productivity at the period level, while simultaneously controlling for the contributions of other players. Explaining variation in peer-effects across team-season also presents a challenge. Oliver (2004) develops a four-factor model of average scoring differential (between a team and their opponents). In this model the factors most highly correlated with average scoring differential are a team's shooting efficiency differential (against opponents), turnover rate differential, rebounding rate differential, and free throw rate differential. The model is highly explanatory of average score margin differentials in the NBA, while maintaining very low levels of right hand side variable dependence (Oliver, 2004). We calculated each of these four factors for each team-season in the data, and found that there is a negative and significant relationship between our estimated peer-effects and the free trow rate differential for each team, conditional on the other factors and team fixed-effects. [22] Even in our moderately-sized sample, there is evidence that a higher free throw rate differential is a moderately strong indicator of weaker peer-effects. What is the potential mechanism generating this relationship? Teams that generate a high free throw rate tend to shoot more contested shots (e.g., isolation-penetration or post entry shots). A high rate of contested shots is often an indication that a team does not rely heavily on rapid, frequent ball movement (e.g., "weak side reversals" and "penetration kicks to the corner") to generate open shots. Ball movement is intended (by its very nature) to generate peer-effects by allocating the ball away from the defense and toward open (relatively high-percentage) shots.

4 Conclusion

We estimate peer-effects using a network production model that controls for competitor-effects, strategy and overall team play and find evidence of (generally positive) peer-effects (team chemistry). In so doing, we generalize the work of Horrace, Liu, and Patacchini (2016), who develop a network production function model that estimates peer (teammate) effects upon player productivity in college basketball. Their peer-effects do not condition upon competitor effects, introducing omitted variable bias. Generalizing their model, we account for strategic decisions and contemporaneous play of opposing team via a competitor network. We apply this model to a three-year sample of NBA regular season game data. We estimate NBA team peer-effects conditioning on both teams' strategies and abilities, rendering a more reliable team-chemistry measure. While generally positive, we find examples of significantly negative peer-effects among NBA teams. We also find that team chemistry tends to persist from year-to-year and that a team's peer-effect is negatively related to its free throw rate differential (which is one of the four primary factors of score margin production in basketball), and we conjecture that teams with higher free throw rate differentials tend to take more contested shoots and are less likely to "find the open man." Based on this, it may be interesting to develop peer-effect weighting schemes based on passing and ball sharing. To do this we could develop statistics based on how often two individual players pass the ball to one another. However, this is left for future research.

[22] Results available by request.

References

[1] Allers, M. A. and Elhorst, J. P., 2011, A simultaneous equations model of fiscal policy interactions. J. Regional Science, 51: 271-291.

[2] Anselin, L., 2007. Spatial econometrics in RSUE: Retrospect and prospect. Regional Science and Urban Economics, 37(4), 450-456.

[3] Baltagi, B. H. and Bresson, G., 2015. Maximum likelihood estimation and lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: an application to hedonic housing prices in Paris. J. Urban Econ., 69 (1), 24-42

[4] Baltagi, B. H. and Deng, Y., 2015. EC3SLS estimator for a simultaneous system of spatial autoregressive equations with random effects Econometric Rev., 34, 659-694

[5] Bajari, P., Hong, H., Krainer, J. and Nekipelov, D., 2010. Estimating static models of strategic interactions. Journal of Business & Economic Statistics, 28(4), 469-482.

[6] Berri, D. J., 1999. Who is most valuable? Measuring the players production of wins in the NBA. Managerial & Decision Economics 20(8) 411-427.

[7] Berri, D., Schmidt, M., & Brook, S., 2006. The wages of wins: Taking measure of the many myths in modern sport. Stanford University Press.

[8] Boarnet, M.G. and Glazer, A., 2002. Federal grants and yardstick competition. Journal of urban Economics, 52(1), 53-64.

[9] Bramoulle, Y., Djebbari, H., Fortin, B., 2009. Identification of peer-effects through social networks. J. Econometrics 150, 41-55.

[10] Brock, W.A., Durlauf, S.N., 2001. Interactions-based models. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, Vol. 5. North-Holland, 32973380.

[11] Brock, W.A., Durlauf, S.N., 2002. A multinomial choice model with neighborhood effects. Amer. Econ. Rev. 92, 298303.

[12] Brock, W.A., Durlauf, S.N., 2006. Multinomial choice with social interactions. In: Blume, L.E., Durlauf, S.N. (Eds.), The Economy as an Evolving Complex System, Vol. III. Oxford University Press, 175206.

[13] Cliff, A., and Ord, J., 1973. Spatial Autocorrelation. Pion, London.

[14] Cliff, A., and Ord, J., 1981. Spatial Processes, Models and Applications. Pion, London.

[15] Cohen-Cole E, Liu, X, Zenou, Y. 2017. Multivariate choices and identification of social interactions. J. Appl Econ., 1-14.

[16] Dahl, G.B., 2002. Mobility and the return to education: Testing a Roy model with multiple markets. Econometrica, 70(6), 2367-2420.

[17] Elhorst, J.P., Lacombe, D.J., Piras, G., 2012. On model specification and parameter space definitions in higher order spatial econometrics models. Reg. Sci. Urban Econom. 42, 211-220.

[18] Gebremariam G. H., Gebremedhin T. G., and Schaeffer, P. V. 2011. Employment, income, and migration in Appalachia: a spatial simultaneous equations approach. J. Reg. Sci., 51 (1), 102-120

[19] Goldsmith-Pinkham, P. and Imbens, G.W., 2013. Social networks and the identification of peer effects. Journal of Business & Economic Statistics, 31(3), 253-264.

[20] Hsieh, C.S. and Lee, L.F., 2016. A social interactions model with endogenous friendship formation and selectivity. Journal of Applied Econometrics, 31(2), 301-319.

[21] Horrace,W. C., Liu, X., Patacchini, E., 2016. Endogenous network production functions with selectivity, J. Econometrics, V. 190, Issue 2, 222-232,

[22] eanty, P.W., Partridge, M., Irwin, W. 2010. Estimation of a spatial simultaneous equation model of population migration and housing price dynamics. Regional Science and Urban Economics 40. 343352

[23] Kelejian,H. H., and Prucha, I. R., 1999. Generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review, 40, 509-533.

[24] Kelejian,H. H., and Prucha, I. R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. J. Econometrics, 104, 219-257

[25] Kelejian,H. H., and Prucha, I. R., 2004. Estimation of simultaneous systems of spatially interrelated cross sectional equations. J. Econometrics, 118, 27-50

[26] Kraus, M.W., Huang, C., and Keltner, D., 2010. Tactile communication, cooperation, and performance: an ethological study of the NBA. Emotion, 5,745-9.

[27] Lee, L.F., 1983. Generalized econometric models with selectivity. Econometrica 51, 507-512.

[28] Lee, L.F., 2001a. Generalized method of moments estimation of spatial autoregressive processes. Manuscript. Department of Economics, Ohio State University.

[29] Lee, L.F., 2001b. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. Manuscript. Department of Economics, Ohio State University.

[30] Lee , L.F., 2003. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances . Econometric Reviews 22 (4): 307-335 .

[31] Lee, L.F., 2007a. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. J. Econometrics 140, 333-374.

[32] Lee, L.F., 2007b. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. J. Econometrics, 137(2), 489-514.

[33] Lee, L.F., 2007c. The method of elimination and substitution in the GMM estimation of mixed regressive, spatial autoregressive models. J. Econometrics, 140(1), 155-189.

[34] Lee, L.F., Liu, X., 2010. Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. Econometric Theory 26, 187-230.

[35] Lee, L.F. and Yu, J., 2014. Efficient GMM estimation of spatial dynamic panel data models with fixed effects. J. Econometrics, 180(2), 174-197.

[36] Lee, L.F., Liu, X. and Lin, X., 2010a. Specification and estimation of social interaction models with network structures. The Econometrics Journal, 13(2), 145-176.

[37] Lin, X. and Lee, L.F., 2010. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. J. Econometrics 157, 34-52.

[38] Liu, X., 2014. Identification and efficient estimation of simultaneous equations network models. Journal of Business & Economic Statistics, 32(4), 516-536.

[39] Liu, X., Lee, L.F. and Bollinger, C.R., 2010b. An efficient GMM estimator of spatial autoregressive models. J. Econometrics, 159(2), 303-319.

[40] Manski, C.F., 1993. Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. 60, 531-542.

[41] McCaffery and Bickart. 2013. The Science of Team Chemistry: Measuring Team Chemistry from Human Biology in Basketball and Soccer. Presentation to the MIT Sloan Sports Analytics Conference, Boston, MA.

[42] Moffitt, R., 2001. Policy interventions low-level equilibria, and social interactions. In: Durlauf, Steven, Young, Peyton (Eds.), Social Dynamics. MIT Press.

[43] Oliver, D., 2004. Basketball on Paper: Rules and Tools for Performance Analysis. Brasey's Inc., Washington, DC.

[44] Olley, G. S., Pakes, A. 1996. The dynamics of productivity in the telecommunications equipment industry, Econometrica, 64(6), 126397.

[45] Ord, J.K., 1975. Estimation methods for models of spatial interaction. J. the American Statistical Association 70, 120-297.

[46] Qu, X. and Lee, L.F., 2015. Estimating a spatial autoregressive model with an endogenous spatial weight matrix. J. Econometrics, 184(2),209-232.

[47] Schrage, M., 2014. Team Chemistry Is the New Holy Grail of Performance Analytics. HBS Digital Article, March 05, 2014. Boston, MA: Harvard Business School Publishing.

[48] Yang, K. and Lee, L.F., 2017. Identification and QML estimation of multivariate and simultaneous equations spatial autoregressive models, J. Econometrics, V196, Issue 1, 196-214.

A GMM for the heterogeneous network interaction model

Here, we provide a sketch of the GMM method for the $R = 2$ case. With slightly modified notation for the simplicity of the case, the model is

$$Y = \sum_{r=1}^{4} \lambda_r G_r Y + X\beta + A + U^* \tag{20}$$

where $G_1 = \left[\begin{smallmatrix} W_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right]$, $G_2 = \left[\begin{smallmatrix} 0 & W_{12} \\ 0 & 0 \end{smallmatrix}\right]$, $G_3 = \left[\begin{smallmatrix} 0 & 0 \\ W_{21} & 0 \end{smallmatrix}\right]$, $G_4 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & W_{22} \end{smallmatrix}\right]$. To eliminate the network specific fixed effects, we premultiply by J_Q,

$$J_Q Y = \sum_{r=1}^{4} \lambda_r J_Q G_r Y + J_Q X\beta + J_Q U \tag{21}$$

Kelejian and Prucha (1999, 2001) introduce a method of moments (MOM) estimator for the SAR disturbance model, and Lee (2001a, 2002b, 2007b) develop it within a general GMM estimation framework and propose a GMM method which explores both IV (linear) as well as quadratic moment functions.[23] They derive a best GMM estimator (BGMME) and show it to have the same limiting distribution as the MLE or QML estimator. Lee and Liu (2010) extend the GMM methodology to the higher-order SAR model.

We assume that each element in U is $i.i.d.N(0, \sigma^2)$ to simplify the estimation.[24] Following Lee's methodology, the $i.i.d.N(0, \sigma^2)$ disturbances imply two sets of moment conditions: linear and quadratic

$$g(\theta) = \begin{pmatrix} Z' J_Q U(\theta) \\ U^{'}(\theta) J_Q F_1 J_Q U(\theta) \\ \dots \\ U^{'}(\theta) J_Q F_r J_Q U(\theta) \end{pmatrix}$$

$J_Q Z$ is a matrix instrument for the linear moments, and $F_r J_Q U(\theta)$ is for the quadratic moments, where F_r is an $n \times n$ matrix such that $tr(F_i J_Q) = 0$, and $U(\theta) = Y - \sum_{i=1}^{4} \lambda_i G_i Y + X\beta$ with $\theta = (\lambda^{'}, \beta^{'})^{'}$.[25] At θ_0, $E(g(\theta_0)) = 0$ as $E(Z' J_Q U) = 0$ and $E(U^{'} J_Q F_i j_Q U) = \sigma^2 tr(F_i J_Q) = 0$ for $i = 1...m$. We can find the intuition for the use of quadratic moments in addition to the linear moments from Lee

[23] See Lee and Lin (2010), Lee et al (2010B), Lee and Yu (2014) for additional details.

[24] The normality assumption is for the simpler forms of the Best IVs and the corresponding estimators' asymptotic distribution later. We can relax the normality and homoskedasticity assumptions in the error terms as long as there is no correlation between the error terms to apply the Lee's methodology. See Lee et al (2010b) for non-normality assumption, and Lee and Lin (2010) for the heteroskedastic specification.

[25] If there is a time dimension of $1...T$, $Z = (Z_1^{'}, ..., Z_T^{'})^{'}, U = (U_1^{'}, ..., U_T^{'})^{'}, J_{Q,T} = I_T \otimes J_Q$, and $F_{i,T} = I_T \otimes F_i$.

(2001a, 2001b). In (21), each $J_Q G_i Y$ can be expressed as $J_Q G_i(I_n - \sum_{i=1}^{4} \lambda_i G_i)^{-1} X\beta + J_Q G_i(I_n - \sum_{i=1}^{4} \lambda_i G_i)^{-1} U$, which has the deterministic part and the stochastic part. Then, when we construct moment conditions for the endogenous variables, $J_Q G_i Y$, we can use the function of the exogenous variables, X, for the approximation of the deterministic part (linear moments) while using the quadratic moments for the stochastic part as long as the $F_i J_Q U$ and the stochastic part are correlated.

The GMM estimator $\hat{\theta}$ follow from $\arg\min_\theta g^{'}(\theta) H g(\theta)$, where H is a distance matrix (or weighting matrix) for the system of equations. In practice, we need to select specific Z, F_i's, and H to implement the GMM. The choice of H would be rather easy because, following Hansen's setting and approach (1982), the optimal H will be the inverse of the variance matrix of the moment conditions with the chosen Z, F_i's.[26] For Z, F_i's, we may consider of the Lee(2007b)'s best Z and F_i's which has the least variance in the asymptotic distribution of $\hat{\theta}$. For our model, these are

- $J_Q Z = [J_Q X, J_Q C_1 X\hat{\beta}, ..., J_Q C_4 X\hat{\beta}]$ where $C_i = G_i(I_n - \sum_{i=1}^{4} \hat{\lambda}_i G_i)^{-1}$ where $\hat{\lambda_i}$ and $\hat{\beta}$ are the preliminary estimates for λ and β.
- $F = (F_i)$ for $i = 1...4$ where $F_i = C_i - \frac{tr(C_i J_Q)}{n-2} J_Q$

We denote this as GMM$_{Lee}$. The derivation of the best Z, F_i's and the asymptotic distribution of the estimators can be found in Appendix D. GMM$_{Lee}$ requires the preliminary estimates of $\hat{\lambda_i}$'s and $\hat{\beta}$ as well as the estimate of the variance matrix of the moment conditions. So, alternatively, we may also think of simpler Kelejian and Prucha (2004)'s linear moments and their corresponding quadratic moments (GMM$_{KP}$) as below[27]

- $J_Q Z = [J_Q X, J_Q G_1 X, J_Q G_2 X, ..., J_Q G_1^{s_1} G_2^{s_2} G_3^{s_3} G_4^{s_4} X]$ $s_1, s_2, s_3, s_4, \geq 0$
- $F = (F_i)$ for $i = 1...m$ where $F_i = H_i - \frac{tr(H_i J_Q)}{n-2} J_Q$ where $H_i = G_1^{s_1} G_2^{s_2} G_3^{s_3} G_4^{s_4}, s_1, s_2, s_3, s_4, \geq 0$

In order to reduce the number of parameters in $g^{'}(\theta)\hat{\Omega}^{-1} g(\theta)$ where $\theta = (\lambda^{'}, \beta^{'})^{'}$, we may replace β with consistent estimator, $\hat{\beta}(\lambda) = [X^{'} J_Q X]^{-1} X^{'} J_Q (I_n - \sum_{i=1}^{4} \lambda_i G_i) Y$. Lee (2007c) analyzes the effect of this modification on the asymptotic distribution of the estimators.

[26]This should be estimated from the estimated residuals of U from an initial consistent estimate of θ.
[27]The asymptotic distribution for this estimator can be established similarly as in the Appendix C

B Identification condition for the equation (13)

Proposition 1. *The equation (13) is identified if* $[J_{Q,m}X_{k,m}, J_{Q,m}\Theta_{k,1,m}, ..., J_{Q,m}\Theta_{k,n_m,m}]$ *has a full rank for* $\forall k = 1, ..., n_m$ *in some* m *where* $\Theta_{i,j,m} = G_{i,j,m}S_m^{-1}X_mB$.

Proof. *Denote* $S_m = I_{N_m} - \sum_{k=1}^{n_m}\sum_{l=1}^{n_m}\lambda_{k,l}G_{k,l,m}$. *As* $S_m^{-1} = I_{N_m} + \sum_{k=1}^{n_m}\sum_{l=1}^{n_m}\lambda_{k,l}G_{k,l,m}S_m^{-1}$ *and* $J_{Q,m}Y = J_{Q,m}S_m^{-1}X_mB + J_{Q,m}S_m^{-1}U_m^*$,[28] *the equation (13) can be written as*

$$\begin{aligned} J_{Q,m}Y_m &= J_{Q,m}S_m^{-1}X_mB + J_{Q,m}S_m^{-1}U_m^* \\ &= \sum_{k=1}^{n_m}\sum_{l=1}^{n_m}\lambda_{k,l}J_{Q,m}G_{k,l,m}S_m^{-1}X_mB + J_{Q,m}X_mB + J_{Q,m}S_m^{-1}U_m^* \end{aligned} \tag{22}$$

Denote $\Theta_{i,j,m} = G_{i,j,m}S_m^{-1}X_mB$, *then, we can see (24) will be identified as long as*

$$[J_{Q,m}X_m, J_{Q,m}\Theta_{1,1,m}, ..., J_{Q,m}\Theta_{1,n_m,m}, ..., J_{Q,m}\Theta_{n_m,1,m}, ...J_{Q,m}\Theta_{n_m,n_m,m}] \tag{23}$$

has a full rank. Due to the structures of the network matrices, the rank condition is equivalent to that

$$[J_{Q,m}X_{k,m}, J_{Q,m}\Theta_{k,1,m}, ..., J_{Q,m}\Theta_{k,n_m,m}] \tag{24}$$

has a full rank for $\forall k = 1, ..., n_m$ *in some* m *where* $X_{k,m}$ *is a* $N_m \times p$ *matrix whose entries in the position* $(\sum_{k=1}^{i-1}n_{k,m} + 1 : \sum_{k=1}^{i}n_{k,m}, 1 : p)$ *is given by* $x_{i,m}$ *but the rest entries are zeros so that* $\sum_{k=1}^{n_m}X_{k,m} = (x_{1,m}^{'}, ..., x_{n_m,m}^{'})^{'}$ *where* $(a : b, c : d)$ *means entries from the* a^{th} *row to* b^{th} *row and from* c^{th} *column to* d^{th} *column. This condition is generally satisfied because we have multiple sets of network matrices and exogenous regressors for each group which produces enough variations to identify the coefficients in our model.* ■

C Lemma 1

Lemma 1. *Let the orthonormal matrix of* $Q_{i,m}$ *be* $[P_{i,m}, l_{n_{i,m}}/\sqrt{n_{i,m}}]$. *The columns in* $P_{i,m}$ *are eigenvectors of* $Q_{i,m}$ *corresponding to the eigenvalue one, such that* $P_{i,m}^{'}l_{n_{i,m}} = 0$, $P_{i,m}^{'}P_{i,m} = I_{n_{i,m}-1}$ *and* $P_{i,m}P_{i,m}^{'} = Q_{i,m}$. *Denote* $J_{P,m}^{'} = Diag(P_{1,m}^{'}, ..., P_{n_m,m}^{'})$ *and* $\bar{G}_{i,j,m} = J_{P,m}^{'}G_{i,j,m}J_{P,m}$, *then,* $\ln|\bar{S}_m| = -\ln f(\Lambda) + \ln|S_m|$ *where* $\bar{S}_m = I_{N_m-n_m} - \sum_{k=1}^{n_m}\sum_{l=1}^{n_m}\lambda_{k,l}\bar{G}_{k,l,m}$, $S_{k,m} = I_{N_m} - \sum_{k=1}^{n_m}\sum_{l=1}^{n_m}\lambda_{k,l}G_{k,l,m}$ *and* $f(\Lambda)$ *is some function of* Λ.

[28] Note that $J_{Q,m}G_{k,l,m}J_{Q,m} = J_{Q,m}G_{k,l,m}$ ∀ k,j, which leads that $J_{Q,m}S_m^{-1}J_{Q,m} = J_{Q,m}S_m^{-1}$

Proof. *Here, we show that the Lemma holds for two networks case. From this, we can easily see that the Lemma holds for any number of networks case. We suppress m here. Define* $H = \begin{bmatrix} [P_1, \frac{l_{n_1}}{\sqrt{n_1}}] & 0 \\ 0 & [P_2, \frac{l_{n_2}}{\sqrt{n_2}}] \end{bmatrix}$.

Then, we can show that $|H^{'}(I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l})H| = |H^{'}H||I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l}| = |I_N - \sum_{k=1}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l}|$ *as* $|H^{'}H| = 1$. *Next, we show*

$$H^{'}(I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l})H = \begin{bmatrix} [P_1, \frac{l_{n_1}}{\sqrt{n_1}}]^{'} & 0 \\ 0 & [P_2, \frac{l_{n_2}}{\sqrt{n_2}}]^{'} \end{bmatrix} \begin{bmatrix} I_{n_1} - \lambda_{1,1}W_{1,1} & -\lambda_{1,2}W_{1,2} \\ -\lambda_{2,1}W_{2,1} & I_{n_2} - \lambda_{2,2}W_{22} \end{bmatrix} \begin{bmatrix} [P_1, \frac{l_1}{\sqrt{n_1}}] & 0 \\ 0 & [P_2, \frac{l_2}{\sqrt{n_2}}] \end{bmatrix}$$

as $P_i^{'}W_{i,j}l_j = 0$ *and* $l_{n_i}^{'}W_{i,j}l_{n_j} = n_i$ *for i,j=1,2*

but $l_{n_i}^{'}W_{i,j}P_j$ *may not be zero because* $W_{i,j}$ *is not necessarily symmetric*

$$= \begin{bmatrix} P_1^{'}(I_{n_1} - \lambda_{1,1}W_{1,1})P_1 & 0 & -\lambda_{1,2}P_1^{'}W_{1,2}P_2 & 0 \\ \frac{l_{n_1}^{'}}{\sqrt{n_1}}(I_{n_1} - \lambda_{1,1}W_{11,})P_1 & 1 - \lambda_{1,1} & \frac{l_{n_1}^{'}}{\sqrt{n_1}}(-\lambda_{1,2}W_{1,2})P_2 & -\sqrt{\frac{n_1}{n_2}}\lambda_{1,2} \\ -\lambda_{2,1}P_2^{'}W_{2,1}P_1 & 0 & P_2^{'}(I_{n_2} - \lambda_{2,2}W_{2,2})P_2 & 0 \\ \frac{l_{n_2}^{'}}{\sqrt{n_2}}(-\lambda_{2,1}W_{2,1})P_1 & -\sqrt{\frac{n_2}{n_1}}\lambda_{2,1} & \frac{l_{n_2}^{'}}{\sqrt{n_2}}(I_{n_2} - \lambda_{2,2}W_{2,2})P_2 & 1 - \lambda_{2,2} \end{bmatrix} \quad (25)$$

Then, from the Laplace's formula, $|H^{'}(I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l})H|$ *is given by*

$$
\begin{aligned}
&|H^{'}(I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l})H| \\
&= (1-\lambda_{1,1})\begin{vmatrix} P_1^{'}(I_{n_1}-\lambda_{1,1}W_{1,1})P_1 & -\lambda_{1,2}P_1^{'}W_{1,2}P_2 & 0 \\ -\lambda_{2,1}P_2^{'}W_{2,1}P_1 & P_2^{'}(I_{n_2}-\lambda_{2,2}W_{2,2})P_2 & 0 \\ \frac{l_{n_2}^{'}}{\sqrt{n_2}}(-\lambda_{2,1}W_{2,1})P_1 & \frac{l_{n_2}^{'}}{\sqrt{n_2}}(I_{n_2}-\lambda_{2,2}W_{2,2})P_2 & 1-\lambda_{2,2} \end{vmatrix} \\
&\qquad + (-1)^{n_2}(-\sqrt{\frac{n_2}{n_1}}\lambda_{2,1})\begin{vmatrix} P_1^{'}(I_{n_1}-\lambda_{1,1}W_{1,1})P_1 & -\lambda_{1,2}P_1^{'}W_{1,2}P_2 & 0 \\ \frac{l_{n_1}^{'}}{\sqrt{n_1}}(I_{n_1}-\lambda_{1,1}W_{1,1})P_1 & \frac{l_{n_1}^{'}}{\sqrt{n_1}}(-\lambda_{1,2}W_{1,2})P_2 & -\sqrt{\frac{n_1}{n_2}}\lambda_{1,2} \\ -\lambda_{2,1}P_2^{'}W_{2,1}P_1 & P_2^{'}(I_{n_2}-\lambda_{2,2}W_{2,2})P_2 & 0 \end{vmatrix} \\
&= (1-\lambda_{1,1})(1-\lambda_{2,2})\begin{vmatrix} P_1^{'}(I_{n_1}-\lambda_{1,1}1W_{1,1})P_1 & -\lambda_{1,2}P_1^{'}W_{1,2}P_2 \\ -\lambda_{2,1}P_2^{'}W_{2,1}P_1 & P_2^{'}(I_{n_2}-\lambda_{2,2}W_{2,2})P_2 \end{vmatrix} \\
&\qquad - \lambda_{1,2}\lambda_{2,1}\begin{vmatrix} P_1^{'}(I_{n_1}-\lambda_{1,1}W_{1,1})P_1 & -\lambda_{1,2}P_1^{'}W_{1,2}P_2 \\ -\lambda_{2,1}P_2^{'}W_{2,1}P_1 & P_2^{'}(I_{n_2}-\lambda_{2,2}W_{2,2})P_2 \end{vmatrix} \\
&= \Big((1-\lambda_{1,1})(1-\lambda_{2,2})-\lambda_{1,2}\lambda_{2,1}\Big)\underbrace{\begin{vmatrix} P_1^{'}(I_{n_1}-\lambda_{1,1}W_{1,1})P_1 & -\lambda_{1,2}P_1^{'}W_{1,2}P_2 \\ -\lambda_{2,1}P_2^{'}W_{2,1}P_1 & P_2^{'}(I_{n_2}-\lambda_{2,2}W_{2,2})P_2 \end{vmatrix}}_{\#}
\end{aligned} \tag{26}
$$

As $\# = |J_P^{'}(I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l})J_P| = |I_{N-2} - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}\bar{G}_{k,l}|$, $|I_N - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}G_{k,l}| = \Big((1-\lambda_{1,1})(1-\lambda_{2,2})-\lambda_{1,2}\lambda_{2,1}\Big)|I_{N-2} - \sum_{k=1}^{2}\sum_{l=1}^{2}\lambda_{k,l}\bar{G}_{k,l}|$. *The Lemma holds for two network group case. From this (in particular, from the matrix (26)), we can easily see that the Lemma holds for any number of networks case. For example, when there are three network groups,* $f(\Lambda) = (1-\lambda_{1,1})(1-\lambda_{2,2})(1-\lambda_{3,3}) - (1-\lambda_{1,1})\lambda_{2,3}\lambda_{3,2} - (1-\lambda_{2,2})\lambda_{1,3}\lambda_{3,1} - (1-\lambda_{3,3})\lambda_{2,1}\lambda_{1,2} - \lambda_{1,3}\lambda_{2,1}\lambda_{3,2}$■

D Derivation of the best IV and the F_is, and its asymptotic distribution

For the best IV and the corresponding F_i, we first derive the covariance matrix for the moment conditions. From the Lemmas 3 in Appendix B and $tr(AB) = vec(A^{'})^{'}vec(B)$ for any conformable matrices A and B, where $vec(A)$ is the column vector formed by stacking the columns of A, we have that $var(g(\theta)) = \Omega$ where

$$\Omega = \begin{bmatrix} \sigma^2 Z' J_Q Z & \mu_3 Z' J_Q \omega_D \\ \mu_3 \omega_D' J_Q Z & (\mu_4 - 3\sigma^4)\omega_D' \omega_D + \sigma^4 \Delta \end{bmatrix} \tag{27}$$

with $\omega_D = [vec_D(J_Q F_1 J_Q)...vec_D(J_Q F_m J_Q)]$, and $\Delta = [vec(J_Q F_1' J_Q)...vec(J_Q F_m' J_Q)]' [vec(J_Q F_1^s J_Q)... vec(J_Q F_m^s J_Q)]$. Because, in our model, the error terms are normally distributed which implies $\mu_3 = 0$ and $\mu_4 = 3\sigma^4$ so simply $\Omega = \begin{bmatrix} \sigma^2 Z' J_Q Z & 0 \\ 0 & \sigma^4 \Delta \end{bmatrix}$. It should be noted that the two types of moments, linear and quadratic, are not correlated when the error are normally distributed, which allows us to choose the best linear moments and the quadratic moments independently; in other words, when we have an optimal Z, they are still optimal given the quadratic moments with F_is.

We know that following the Hansen setting and approach (1982), the asymptotic distribution of the GMM estimator with the inverse of the variance matrix of the moment conditions, Ω^{-1}, as for its distance (weighting) matrix, is given by

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \xrightarrow{D} N(0, (\lim_{n\to\infty} \frac{1}{n} D' \Omega^{-1} D)^{-1}) \tag{28}$$

where $D = \frac{\partial E(g(\theta_0))}{\partial \theta'}$. In our model, as $J_Q G_i Y = J_Q G_i (I_n - \sum_{i=1}^4 \lambda_i G_i)^{-1} X\beta + J_Q G_i (I_n - \sum_{i=1}^4 \lambda_i G_i)^{-1} U$,

$$\begin{aligned} D &= \frac{\partial E(g(\theta_0))}{\partial \theta'} \\ &= \begin{bmatrix} Z' J_Q G_1 (I_n - \sum_{i=1}^4 \lambda_i^0 G_i)^{-1} X\beta_0 & ... & Z' J_Q G_4 (I_n - \sum_{i=1}^4 \lambda_i^0 G_i)^{-1} X\beta_0 & Z' J_Q X \\ \sigma_0^2 tr(J_Q F_1^s J_Q G1 (I_n - \sum_{i=1}^4 \lambda_i^0 G_i)^{-1}) & ... & \sigma_0^2 tr(J_Q F_1^s J_Q G_4 (I_n - \sum_{i=1}^4 \lambda_i^0 G_i)^{-1}) & 0 \\ \vdots & ... & \vdots & \vdots \\ \sigma_0^2 tr(J_Q F_m^s J_Q G1 (I_n - \sum_{i=1}^4 \lambda_i^0 G_i)^{-1}) & ... & \sigma_0^2 tr(J_Q F_m^s J_Q G_4 (I_n - \sum_{i=1}^4 \lambda_i^0 G_i)^{-1}) & 0 \end{bmatrix} \\ &= \begin{bmatrix} Z' J_Q D_{11} & Z' J_Q X \\ \sigma_0^2 D_{21} & 0 \end{bmatrix} \end{aligned} \tag{29}$$

where $D_{11} = [C_1^0 X\beta_0, ..., C_4^0 X\beta_0]$ with $C_i^0 = G_i(I_n - \sum_{i=1}^4 \lambda_i^0 G_i)$

and $D_{21} = \begin{bmatrix} tr(J_Q F_1^s J_Q C_1^0) & ... & tr(J_Q F_1^s J_Q C_4^0) \\ \vdots & ... & \vdots \\ tr(J_Q F_m^s J_Q C_1^0) & ... & tr(J_Q F_m^s J_Q C_4^0) \end{bmatrix}$

Then, under certain regularity conditions as in Lee (2007b) and Lee and Liu (2010), the asymptotic distribution of the estimators, θ, derived from $\min_\theta g(\theta)^{'}\hat{\Omega}^{-1}g(\theta)$, is given by

$$\sqrt{n}(\hat{\theta}_{GMM} - \theta_0) \xrightarrow{D} N(0, (\lim_{n\to\infty}\frac{1}{n}\Sigma)^{-1})$$

$$\begin{aligned} \text{where} \quad \Sigma &= D^{'}\Omega^{-1}D \\ &= \begin{bmatrix} D_{11}^{'}J_Q Z & \sigma_0^2 D_{21}^{'} \\ X^{'}J_Q Z & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma^2}(Z^{'}J_Q Z)^{-1} & 0 \\ 0 & \frac{1}{\sigma^4}\Delta^{-1} \end{bmatrix} \begin{bmatrix} Z^{'}J_Q D_{11} & Z^{'}J_Q X \\ \sigma_0^2 D_{21} & 0 \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} D_{21}^{'}\Delta^{-1}D_{21} & 0 \\ 0 & 0 \end{bmatrix}}_{(1)} + \underbrace{\frac{1}{\sigma^2}[D_{11}, X]^{'}J_Q Z(Z^{'}J_Q Z)^{-1}Z^{'}J_Q[D_{11}, X]}_{(2)} \end{aligned} \tag{30}$$

We can notice that the asymptotic variance is the sum of the two parts, (1) that is affected by the choice of F_is, and (2) that is affected by the choice of Z, and the two are independent which is due to the normality assumption in the error terms. Then, by the generalized Schwartz inequality, the best linear IV will be $J_Q Z_B = J_Q[D_{11}, X] = J_Q[C_1^0 X\beta_0, ..., C_4^0 X\beta_0, X]$, which reduce (2) to $\frac{1}{\sigma^2}(Z_B^{'}J_Q Z_B)$. And for the best quadratic moments, as $vec^{'}(A^{'})vec(B^s) = tr(AB^s) = \frac{1}{2}tr(A^s B^s) = \frac{1}{2}vec^{'}(A^s)vec(B^s)$ for any conformable matrix A and B, and $tr(J_Q F_i^s J_Q C_i^0) = tr(J_Q F_i^s J_Q C_i^0 J_Q) = tr(J_Q F_i^s J_Q \underbrace{(C_i^0 - \frac{tr(C_i^0 J_Q)}{n-2}J_Q)}_{(3)} J_Q) = \frac{1}{2}tr(J_Q F_i^s J_Q(C_i^0 - \frac{tr(C_i^0 J_Q)}{n-2}J_Q)^s J_Q) = \frac{1}{2}vec^{'}(J_Q F_i^s J_Q)vec(J_Q(C_i^0 - \frac{tr(C_i^0 J_Q)}{n-2}J_Q)^s J_Q)$, Δ and D_{21} can be rewritten as $\Delta = \frac{1}{2}[vec(J_Q F_1^s J_Q)...vec(J_Q F_m^s J_Q)]^{'}[vec(J_Q F_1^s J_Q)...vec(J_Q F_m^s J_Q)]$ and $D_{21} = \frac{1}{2}[vec(J_Q F_1^s J_Q)...vec(J_Q F_m^s J_Q)]^{'}[vec(J_Q(C_1^0 - \frac{tr(C_1^0 J_Q)}{n-2}J_Q)^s J_Q), ..., vec(J_Q(C_4^0 - \frac{tr(C_4^0 J_Q)}{n-2}J_Q)^s J_Q)]$. We need to note that (3) allow us to transform the D_{21} to the form where we can use the generalized Schwartz inequality while satisfying $tr(F_i J_Q) = 0$ when we set $F_i = (3)$. Then, by the generalized Schwartz inequality, the best F_B is given by $F_B = (F_{B,i}) = (C_i^0 - \frac{tr(C_i^0 J_Q)}{n-2}J_Q)$ for $i = 1, ..., 4$, which leads to the maximum value for $D_{21}^{'}\Delta^{-1}D_{21} = \frac{1}{2}[vec(J_Q F_{B,1}^s J_Q)...vec(J_Q F_{B,4}^s J_Q)]^{'}$ $[vec(J_Q F_{B,1}^s J_Q)...vec(J_Q F_{B,4}^s J_Q)] = \frac{1}{2}(tr(J_Q F_{B,i}^s J_Q F_{B,j}^s))_{i,j} = (tr(J_Q F_{B,i}^s J_Q F_{B,j}))_{i,j} = (tr(J_Q F_{B,i}^s J_Q C_j^0))_{i,j}$ for $i, j = 1, ..., 4$. Then, the asymptotic distribution of the Best GMM (BGMM) with the

F_B, Z_B and the distance matrix of Ω^{-1} is given by

$$\sqrt{n}(\hat{\theta}_{BGMM} - \theta_0) \xrightarrow{D} N(0, (\lim_{n\to\infty} \frac{1}{n}\Sigma)^{-1})$$
$$\text{where} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\sigma^2}(Z_B^{'} J_Q Z_B) \quad \text{and} \quad \Sigma_{11} = (tr(J_Q F_{B,i}^s J_Q C_j^0))_{i,j} \quad \text{for } i, j = 1, ..., 4 \tag{31}$$

We can show that, under the regularity conditions, this asymptotic variance is equivalent with the one from MLE (Lee, 2007b).

E Descriptive Statistics for NBA Data

Descriptive statistics for each team in each season are in Table 2-4. Teams are ranked based on the estimated peer-effect in Table 1. For each team we present the mean and standard deviations of wins produced per minute ($Wins$), player experience ($Experience$), and fatigue ($Fatigure$) in Tables 2-4. $Experience$ is minutes played from the start of the game to the end of the last period, and $Fatigue$ is minutes continuously played until the end of the last period. Each team played 82 games in each season. We also report the number of periods($Periods$), observations ($Observations = 5 * Periods$) and average duration per period (ADP).

Table 2: Season 2013-14

Team	Wins (wins/min)		Experience (mins)		Fatigue (mins)		Games	Periods	Observations	ADP
	Mean	SD	Mean	SD	Mean	SD				
UTA	0.0049	0.0248	13.794	9.0010	5.2439	4.7098	82	1,826	9,130	1.8223
BOS	0.0050	0.0251	13.860	9.2391	4.6564	4.3956	82	1,858	9,290	1.7856
PHX	0.0064	0.0257	13.735	8.9724	5.0906	4.5023	82	1,820	9,100	1.8456
WAS	0.0055	0.0250	14.594	9.6572	4.9988	4.6697	82	1,763	8,815	1.8678
IND	0.0058	0.0243	14.287	9.4832	5.1662	4.8740	82	1,726	8,630	1.9308
MIN	0.0062	0.0257	13.739	9.4149	5.4394	5.0011	82	1,736	8,680	1.9257
DAL	0.0063	0.0256	13.389	8.9545	3.6121	3.6704	82	2,021	10,105	1.6514
OKC	0.0070	0.0257	13.146	9.0519	5.1941	5.0077	82	1,817	9,085	1.8307
SAS	0.0071	0.0258	12.257	8.1465	4.0379	3.8646	82	1,887	9,435	1.7790
TOR	0.0061	0.0253	14.235	9.4367	4.7076	4.3403	82	1,873	9,365	1.7760
BKN	0.0055	0.0237	12.717	8.5106	4.7137	4.2582	82	1,780	8,900	1.9225
CLE	0.0050	0.0247	13.731	9.1616	4.8408	4.5234	82	1,821	9,105	1.8339
PHI	0.0049	0.0259	13.491	9.0089	4.0930	3.9932	82	1,908	9,540	1.7361
NOP	0.0061	0.0253	13.406	9.0509	4.7968	4.5722	82	1,878	9,390	1.7740
DET	0.0061	0.0254	14.905	9.6595	5.4940	5.0559	82	1,811	9,055	1.8297
MIL	0.0051	0.0249	13.876	9.0502	4.8941	4.6383	82	1,835	9,175	1.8135
LAL	0.0058	0.0253	13.916	9.0277	5.1118	4.6862	82	1,783	8,915	1.8878
NYK	0.0054	0.0243	14.462	9.4261	5.0768	4.7384	82	1,810	9,050	1.8432
DEN	0.0062	0.0262	13.251	8.7821	4.8769	4.5826	82	1,909	9,545	1.7564
GSW	0.0067	0.0262	14.840	9.9523	5.3323	5.5644	82	1,816	9,080	1.8174
MEM	0.0057	0.0239	13.626	9.0649	4.9568	4.6083	82	1,779	8,895	1.8466
CHA	0.0055	0.0247	13.286	8.7265	4.7021	4.4664	82	1,780	8,900	1.8933
SAC	0.0059	0.0251	14.008	9.3787	5.4816	5.3535	82	1,821	9,105	1.8184
HOU	0.0071	0.0259	14.571	9.6148	5.3114	5.0911	82	1,768	8,840	1.8993
ORL	0.0052	0.0245	13.862	9.1897	5.0372	4.7521	82	1,711	8,555	1.9501
ATL	0.0058	0.0254	13.558	8.8973	4.1517	3.8550	82	1,848	9,240	1.8068
LAC	0.0068	0.0263	14.621	9.8490	5.0150	4.7810	82	1,867	9,335	1.7628
POR	0.0062	0.0255	14.455	9.5653	4.6440	4.5626	82	1,752	8,760	1.9005
CHI	0.0053	0.0237	15.545	9.8489	6.5129	5.6085	82	1,644	8,220	2.0183
MIA	0.0064	0.0245	13.522	8.8854	4.7217	4.3458	82	1,814	9,070	1.8506

ADP: the average duration per period

Table 3: Season 2014-15

Team	Wins (wins/minute)		Experience (mins)		Fatigue (mins)		Games	Periods	Observations	ADP
	Mean	SD	Mean	SD	Mean	SD				
IND	0.0058	0.0256	13.037	8.2898	4.5181	4.0489	82	1,824	9,120	1.8057
PHX	0.0060	0.0256	14.016	9.0446	4.9225	4.3733	82	1,876	9,380	1.7743
UTA	0.0059	0.0261	13.820	8.8752	3.9650	3.8336	82	2,039	10,195	1.6217
PHI	0.0048	0.0265	13.140	8.4847	3.6735	3.6020	82	2,000	10,000	1.6325
MIN	0.0054	0.0252	14.878	9.6813	5.3879	4.9777	82	1,882	9,410	1.7740
CLE	0.0064	0.0254	14.379	9.4663	4.8025	4.5130	82	1,871	9,355	1.7734
WAS	0.0059	0.0252	13.231	8.7861	4.8437	4.4499	82	1,797	8,985	1.8423
LAC	0.0068	0.0258	14.355	9.6817	4.9482	4.6120	82	1,800	9,000	1.8181
ORL	0.0052	0.0244	14.011	9.4634	4.7217	4.4842	82	1,808	9,040	1.8398
SAC	0.0057	0.0254	13.775	9.1679	4.8073	4.4546	82	1,885	9,425	1.7640
SAS	0.0066	0.0254	12.364	8.2696	3.7030	3.6057	82	1,968	9,840	1.6986
MEM	0.0058	0.0240	13.292	8.8446	4.9134	4.4669	82	1,836	9,180	1.8094
BKN	0.0052	0.0245	13.785	9.1974	5.0943	4.6681	82	1,869	9,345	1.7900
NOP	0.0062	0.0255	14.415	9.4324	5.0235	4.6812	82	1,837	9,185	1.7978
DEN	0.0056	0.0260	13.238	9.0719	4.9736	4.7908	82	1,880	9,400	1.7684
TOR	0.0064	0.0258	13.799	8.6910	4.5766	4.1246	82	1,874	9,370	1.7762
CHA	0.0055	0.0247	13.205	8.7309	4.6170	4.2395	82	1,775	8,875	1.8785
DAL	0.0063	0.0265	13.168	8.7553	3.9218	3.8520	82	2,040	10,200	1.6504
MIL	0.0057	0.0247	13.076	8.4892	4.8779	4.3510	82	1,857	9,285	1.8245
ATL	0.0062	0.0263	13.310	8.7269	3.6930	3.5976	82	1,954	9,770	1.6949
HOU	0.0064	0.0263	14.511	9.3455	4.6074	4.4417	82	1,862	9,310	1.7910
LAL	0.0055	0.0251	13.234	8.4778	5.2972	4.9816	82	1,692	8,460	1.9895
POR	0.0063	0.0251	13.768	9.3065	4.3383	4.1583	82	1,764	8,820	1.8839
CHI	0.0064	0.0252	15.114	9.4816	5.4462	4.9676	82	1,728	8,640	1.9041
MIA	0.0054	0.0251	13.764	8.9606	4.7052	4.4468	82	1,934	9,670	1.7300
DET	0.0056	0.0253	13.792	8.7494	5.0442	4.4047	82	1,783	8,915	1.8634
OKC	0.0064	0.0265	13.714	9.2057	5.1044	4.6884	82	1,933	9,665	1.7424
GSW	0.0076	0.0261	13.260	8.7507	4.7493	4.3759	82	1,845	9,225	1.7982
NYK	0.0048	0.0244	12.849	8.5494	4.3007	4.1936	82	1,906	9,530	1.7674
BOS	0.0055	0.0254	13.221	8.4947	4.2091	3.9081	82	1,963	9,815	1.6984

ADP: the average duration per period

Table 4: Season 2015-16

Team	Wins (wins/min)		Experience (mins)		Fatigue (mins)		Games	Periods	Observations	ADP
	Mean	SD	Mean	SD	Mean	SD				
CHI	0.0059	0.0249	13.443	8.8153	4.6841	4.4133	82	1,793	8,965	1.8500
SAS	0.0072	0.0253	12.260	8.2964	3.8246	3.6986	82	1,878	9,390	1.7312
MIN	0.0058	0.0253	13.162	8.9522	4.8360	4.3706	82	1,891	9,455	1.7825
MIL	0.0055	0.0242	14.327	9.5646	5.5877	5.3570	82	1,786	8,930	1.8766
MEM	0.0056	0.0252	13.569	8.8102	4.8904	4.4279	82	1,855	9,275	1.7359
ORL	0.0060	0.0249	13.366	9.1810	5.1182	4.9035	82	1,815	9,075	1.8481
BKN	0.0053	0.0252	13.583	8.9497	4.5742	4.2478	82	1,838	9,190	1.7807
DET	0.0059	0.0257	14.666	9.4099	4.8601	4.5260	82	1,754	8,770	1.9050
SAC	0.0064	0.0261	14.171	9.0768	4.8797	4.5854	82	1,957	9,785	1.7061
POR	0.0066	0.0261	13.415	8.7288	4.6371	4.1711	82	1,839	9,195	1.8099
BOS	0.0061	0.0266	13.568	8.7688	4.5319	4.1359	82	1,894	9,470	1.6840
TOR	0.0068	0.0254	13.991	8.9852	4.8002	4.2214	82	1,802	9,010	1.8506
UTA	0.0059	0.0256	13.565	8.9888	3.8109	3.7618	82	2,035	10,175	1.6233
CLE	0.0066	0.0264	13.731	9.0387	4.5803	4.2359	82	1,883	9,415	1.7611
HOU	0.0063	0.0269	13.822	9.4650	4.4653	4.7234	82	1,964	9,820	1.7123
CHA	0.0068	0.0254	13.590	8.9516	4.9244	4.5149	82	1,741	8,705	1.9243
ATL	0.0063	0.0264	13.145	8.6175	3.3534	3.3129	82	2,021	10,105	1.6391
PHX	0.0053	0.0261	13.941	9.4149	5.2119	5.0957	82	1,821	9,105	1.8327
LAL	0.0056	0.0256	12.758	8.3083	4.9857	4.4327	82	1,752	8,760	1.9016
WAS	0.0061	0.0257	13.714	9.0765	4.7863	4.4908	82	1,914	9,570	1.7293
NYK	0.0058	0.0251	13.252	8.9107	4.4921	4.3276	82	1,847	9,235	1.8075
OKC	0.0073	0.0265	13.545	8.9972	4.5932	4.3073	82	1,857	9,285	1.7843
LAC	0.0065	0.0263	13.064	8.7264	4.6467	4.2978	82	1,862	9,310	1.7728
MIA	0.0065	0.0253	14.409	8.8739	4.7797	4.2856	82	1,847	9,235	1.8138
NOP	0.0060	0.0252	14.227	9.1684	5.0658	4.5300	82	1,875	9,375	1.7772
PHI	0.0051	0.0265	13.129	8.5196	3.6186	3.5643	82	2,061	10,305	1.6193
GSW	0.0081	0.0268	13.157	9.1074	4.2704	4.0395	82	1,892	9,460	1.7745
DAL	0.0058	0.0257	13.475	8.8809	4.0182	3.9327	82	1,968	9,840	1.7049
IND	0.0061	0.0255	13.900	8.9143	4.6639	4.1757	82	1,839	9,195	1.8162
DEN	0.0058	0.0257	13.450	8.7122	5.0957	4.6100	82	1,843	9,215	1.8071

ADP: the average duration per period

F Estimation results

The estimation results for the main equation are in Tables 5-7 for each team in each season. As the focus is the peer-effects and not the marginal effects of other important variables, we only estimate the main equation after the within transformation. Therefore, only the marginal effects for variables that vary at the player-level are identified. Teams in tables are ranked based on the estimated peer-effect in Table 1.

Table 5: 2013-14 Season Estimates.

	Parameter values					T statistics				
Team	λ_{rr}	λ_{rk}	Exper.	Fatigue	σ^2	λ_{rr}	λ_{rk}	Exper.	Fatigue	σ^2
UTA	0.0446	0.0565	0.0000	0.0001	0.0006	3.529	2.285	-0.390	1.826	57.032
BOS	0.0442	0.0009	0.0000	0.0002	0.0007	3.531	0.038	-0.216	2.541	57.536
PHX	0.0410	-0.0519	0.0000	0.0001	0.0007	2.807	-1.925	-0.100	1.083	55.746
WAS	0.0314	-0.0052	0.0002	0.0002	0.0006	2.219	-0.204	2.056	1.840	55.166
IND	0.0274	0.0317	0.0003	0.0000	0.0006	2.189	1.292	2.798	0.392	55.514
MIN	0.0259	-0.0399	0.0002	0.0001	0.0007	2.228	-1.700	2.211	1.576	56.086
DAL	0.0223	-0.0269	0.0000	0.0001	0.0007	1.846	-1.267	-0.163	0.896	59.684
OKC	0.0208	-0.0119	0.0000	0.0001	0.0007	1.656	-0.505	-0.281	1.529	56.664
SAS	0.0206	0.0214	0.0002	0.0001	0.0007	1.880	0.944	2.098	0.659	58.502
TOR	0.0198	0.0279	-0.0001	0.0002	0.0007	1.708	1.176	-0.870	3.007	57.956
BKN	0.0186	-0.0008	0.0001	0.0001	0.0006	1.484	-0.037	1.565	0.840	56.151
CLE	0.0175	0.0355	0.0001	0.0001	0.0006	1.531	1.669	0.780	1.607	57.284
PHI	0.0171	0.0544	0.0000	0.0000	0.0007	1.511	2.438	0.224	-0.202	58.547
NOP	0.0165	-0.0103	0.0001	0.0001	0.0006	1.515	-0.459	1.248	1.492	58.348
DET	0.0150	-0.0376	-0.0001	0.0001	0.0006	1.305	-1.628	-1.320	1.542	57.084
MIL	0.0142	-0.0351	0.0000	-0.0001	0.0006	1.262	-1.604	0.472	-0.821	57.553
LAL	0.0133	0.0037	-0.0002	0.0003	0.0006	1.169	0.167	-1.942	3.631	56.738
NYK	0.0089	-0.0057	-0.0001	0.0001	0.0006	0.801	-0.253	-1.233	0.750	57.172
DEN	0.0089	-0.0072	0.0000	0.0000	0.0007	0.818	-0.311	-0.514	0.519	58.687
GSW	0.0069	0.0106	0.0000	0.0000	0.0007	0.563	0.429	0.246	-0.396	56.670
MEM	0.0056	-0.0188	0.0000	0.0000	0.0006	0.493	-0.846	0.284	0.616	56.570
CHA	0.0044	-0.0080	0.0000	0.0000	0.0006	0.393	-0.363	0.265	-0.260	56.612
SAC	0.0004	0.0087	-0.0001	0.0000	0.0006	0.037	0.376	-1.319	0.630	56.879
HOU	-0.0020	-0.0272	0.0001	0.0001	0.0007	-0.180	-1.209	0.808	1.433	56.330
ORL	-0.0038	0.0320	0.0000	0.0002	0.0006	-0.333	1.368	0.494	2.137	55.395
ATL	-0.0052	0.0205	0.0001	0.0001	0.0006	-0.463	0.943	0.761	0.808	57.438
LAC	-0.0090	0.0163	0.0001	0.0000	0.0007	-0.823	0.729	0.677	0.097	57.822
POR	-0.0102	-0.0184	0.0001	0.0002	0.0006	-0.879	-0.781	0.787	2.684	55.849
CHI	-0.0171	0.0057	-0.0002	0.0001	0.0006	-1.438	0.249	-3.000	1.413	54.035
MIA	-0.0193	-0.0214	-0.0002	0.0002	0.0006	-1.713	-0.953	-2.037	2.822	56.740

λ_{rr} is the peer-effect; λ_{rk} is the competator effect. Exper. is Experience.

Table 6: 2014-15 Season Estimates.

Team	Parameter values					T statistics				
	λ_{rr}	λ_{rk}	Exper.	Fatigue	σ^2	λ_{rr}	λ_{rk}	Exper.	Fatigue	σ^2
IND	0.0541	-0.0025	0.0000	0.0001	0.0007	3.702	-0.099	0.501	0.637	55.924
PHX	0.0328	-0.0040	-0.0002	0.0002	0.0007	2.501	-0.169	-1.854	1.815	57.282
UTA	0.0321	0.0440	0.0000	0.0001	0.0007	2.746	1.903	-0.424	1.329	60.277
PHI	0.0232	-0.0017	0.0000	0.0002	0.0007	1.711	-0.073	-0.155	1.981	58.458
MIN	0.0204	0.0022	0.0000	0.0000	0.0007	1.805	0.099	0.546	-0.094	58.243
CLE	0.0182	-0.0051	0.0002	0.0001	0.0007	1.570	-0.219	2.855	1.785	57.901
WAS	0.0181	-0.0383	-0.0001	0.0001	0.0007	1.411	-1.594	-0.762	0.742	56.184
LAC	0.0167	-0.0212	0.0002	0.0001	0.0007	1.530	-0.911	1.598	1.609	57.229
ORL	0.0140	0.0209	0.0001	0.0000	0.0006	1.264	0.929	0.893	-0.241	57.247
SAC	0.0107	0.0235	-0.0001	0.0002	0.0007	0.946	1.047	-1.032	1.988	58.183
SAS	0.0088	-0.0144	0.0001	0.0000	0.0007	0.808	-0.654	0.796	-0.072	59.502
MEM	0.0076	-0.0176	-0.0001	0.0002	0.0006	0.686	-0.808	-1.344	2.308	57.579
BKN	0.0066	-0.0169	-0.0001	0.0002	0.0006	0.607	-0.824	-1.832	2.329	58.122
NOP	0.0056	0.0505	0.0000	0.0001	0.0006	0.505	2.277	0.299	1.581	57.514
DEN	0.0021	-0.0202	0.0001	0.0001	0.0007	0.188	-0.898	1.038	0.914	57.984
TOR	0.0017	-0.0103	0.0001	0.0001	0.0007	0.153	-0.456	1.648	1.147	58.061
CHA	0.0014	0.0112	0.0002	0.0001	0.0006	0.123	0.486	2.568	0.590	56.524
DAL	0.0011	0.0425	0.0001	0.0001	0.0007	0.101	1.931	0.885	0.706	60.477
MIL	0.0005	-0.0092	0.0001	0.0000	0.0006	0.049	-0.429	1.655	-0.286	57.768
ATL	0.0000	0.0225	0.0000	0.0000	0.0007	0.003	0.977	0.068	0.256	58.461
HOU	-0.0011	0.0383	0.0000	0.0002	0.0007	-0.098	1.662	0.078	2.659	57.834
LAL	-0.0022	0.0281	-0.0001	0.0001	0.0006	-0.180	1.061	-0.721	1.722	54.723
POR	-0.0031	0.0449	0.0000	-0.0002	0.0006	-0.274	1.941	-0.113	-1.771	56.223
CHI	-0.0038	-0.0012	0.0000	0.0001	0.0006	-0.321	-0.049	-0.314	0.820	55.533
MIA	-0.0054	-0.0070	0.0000	0.0000	0.0006	-0.490	-0.319	-0.242	0.608	58.726
DET	-0.0064	-0.0734	0.0000	0.0004	0.0006	-0.550	-3.163	0.011	4.188	56.296
OKC	-0.0071	0.0082	0.0002	-0.0001	0.0007	-0.645	0.368	2.787	-1.089	58.760
GSW	-0.0219	0.0569	0.0001	0.0000	0.0007	-2.009	2.378	0.766	0.250	57.356
NYK	-0.0254	-0.0028	0.0002	-0.0001	0.0006	-2.333	-0.129	2.362	-1.215	58.137
BOS	-0.0261	-0.0067	0.0000	0.0000	0.0006	-2.411	-0.313	0.239	0.376	58.933

λ_{rr} is the peer-effect; λ_{rk} is the competator effect. Exper. is Experience.

Table 7: 2014-15 Season Estimates.

	Parameter values					T statistics				
Team	λ_{rr}	λ_{rk}	Exper.	Fatigue	σ^2	λ_{rr}	λ_{rk}	Exper.	Fatigue	σ^2
CHI	0.0648	-0.0635	0.0000	0.0001	0.0006	4.120	-2.598	0.580	1.143	54.955
SAS	0.0522	0.0112	0.0000	0.0002	0.0007	3.607	0.460	-0.555	2.302	56.670
MIN	0.0489	0.0049	0.0000	0.0001	0.0007	3.592	0.212	0.457	1.238	57.368
MIL	0.0457	0.0135	0.0001	0.0000	0.0006	3.418	0.578	1.065	-0.104	56.079
MEM	0.0452	0.0174	0.0000	0.0001	0.0007	2.975	0.729	-0.136	0.789	55.853
ORL	0.0365	0.0195	0.0001	0.0000	0.0006	2.606	0.812	0.840	-0.476	55.983
BKN	0.0350	0.0006	0.0001	-0.0001	0.0007	2.747	0.028	0.717	-1.356	57.035
DET	0.0296	0.0107	-0.0001	0.0002	0.0007	2.233	0.428	-0.593	2.227	55.515
SAC	0.0277	-0.0221	0.0000	0.0001	0.0007	2.188	-0.963	0.299	0.769	58.562
POR	0.0269	-0.0150	0.0001	0.0001	0.0007	2.430	-0.686	0.695	1.103	57.871
BOS	0.0218	-0.0008	0.0002	0.0000	0.0007	1.802	-0.035	1.891	0.262	57.979
TOR	0.0200	0.0006	0.0000	0.0002	0.0007	1.761	0.026	0.470	2.054	57.081
UTA	0.0157	0.0008	0.0001	0.0003	0.0007	1.301	0.038	1.094	2.809	59.773
CLE	0.0155	0.0273	-0.0002	0.0004	0.0007	1.368	1.184	-2.002	4.074	58.195
HOU	0.0137	0.0060	0.0000	0.0002	0.0007	1.223	0.264	0.250	2.353	59.343
CHA	0.0113	-0.0126	0.0000	0.0001	0.0007	0.943	-0.490	-0.226	1.570	55.771
ATL	0.0084	0.0495	0.0003	-0.0001	0.0007	0.734	2.318	2.763	-0.485	59.825
PHX	0.0037	-0.0365	-0.0001	0.0000	0.0007	0.319	-1.564	-0.752	0.151	57.007
LAL	0.0011	0.0548	0.0000	0.0001	0.0007	0.091	2.233	0.415	1.523	55.910
WAS	0.0000	-0.0075	0.0000	0.0001	0.0007	0.001	-0.344	-0.309	0.689	58.429
NYK	-0.0001	0.0506	0.0001	0.0001	0.0006	-0.011	2.300	0.773	1.489	57.557
OKC	-0.0020	-0.0506	-0.0001	0.0002	0.0007	-0.179	-2.257	-0.869	1.866	57.668
LAC	-0.0045	0.0170	-0.0002	0.0002	0.0007	-0.409	0.767	-1.990	2.648	57.780
MIA	-0.0071	0.0267	0.0000	0.0001	0.0006	-0.590	1.199	0.051	1.191	57.004
NOP	-0.0111	0.0085	0.0001	0.0001	0.0006	-0.920	0.387	1.281	1.563	57.313
PHI	-0.0139	-0.0246	0.0001	0.0001	0.0007	-1.208	-1.128	0.981	1.278	59.984
GSW	-0.0144	-0.0203	0.0001	0.0000	0.0007	-1.336	-0.858	0.899	0.403	58.161
DAL	-0.0163	-0.0089	0.0001	-0.0001	0.0007	-1.415	-0.412	1.531	-0.650	58.744
IND	-0.0175	-0.0005	0.0001	0.0001	0.0007	-1.457	-0.024	0.665	0.792	56.736
DEN	-0.0282	-0.0123	0.0001	0.0001	0.0006	-2.247	-0.552	1.329	1.233	56.321

λ_{rr} is the peer-effect; λ_{rk} is the competator effect. Exper. is Experience.