CENTER FOR POLICY RESEARCH

THE MAXWELL SCHOOL

WORKING PAPER SERIES

Network Competition and Team Chemistry in the NBA

William C. Horrace, Hyunseok Jung, and Shane Sanders

Paper No. 226
March 2020

ISSN: 1252-3066

SYRACUSE UNIVERSITY
Maxwell School
Center for Policy Research

426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
(315) 443-3114/email: ctrpol@syr.edu
http://www.maxwell.syr.edu/CPR_Working_Papers.aspx

CENTER FOR POLICY RESEARCH – Spring 2020

Leonard M. Lopoo, Director
Professor of Public Administration and International Affairs (PAIA)

Associate Directors

Margaret Austin
Associate Director, Budget and Administration

John Yinger
Trustee Professor of Economics (ECON) and Public Administration and International Affairs (PAIA)
Associate Director, Center for Policy Research

SENIOR RESEARCH ASSOCIATES

Badi Baltagi, ECON
Robert Bifulco, PAIA
Leonard Burman, PAIA
Carmen Carrión-Flores, ECON
Alfonso Flores-Lagunes, ECON
Sarah Hamersma, PAIA
Madonna Harrington Meyer, SOC
Colleen Heflin, PAIA
William Horrace, ECON
Yilin Hou, PAIA
Hugo Jales, ECON
Jeffrey Kubik, ECON
Yoonseok Lee, ECON
Amy Lutz, SOC
Yingyi Ma, SOC
Katherine Michelmore, PAIA
Jerry Miner, ECON
Shannon Monnat, SOC
Jan Ondrich, ECON
David Popp, PAIA
Stuart Rosenthal, ECON
Michah Rothbart, PAIA
Alexander Rothenberg, ECON
Rebecca Schewe, SOC
Amy Ellen Schwartz, PAIA/ECON
Ying Shi, PAIA
Saba Siddiki, PAIA
Perry Singleton, ECON
Yulong Wang, ECON
Michael Wasylenko, ECON
Peter Wilcoxen, PAIA
Maria Zhu, ECON

GRADUATE ASSOCIATES

Rhea Acuña, PAIA
Mariah Brennan, SOC. SCI.
Jun Cai, ECON
Ziqiao Chen, PAIA
Yoon Jung Choi, PAIA
Dahae Choo, PAIA
Stephanie Coffey, ECON
Giuseppe Germinario, ECON
Myriam Gregoire-Zawilski, PAIA
Emily Gutierrez, PAIA
Jeehee Han, PAIA
Mary Helander, Lerner
Hyoung Kwon, PAIA
Mattie Mackenzie-Liu, PAIA
Maeve Maloney, ECON
Austin McNeill Brown, SOC. SCI.
Qasim Mehdi, PAIA
Claire Pendergrast, SOC
Jonathan Presler, ECON
Krushna Ranaware, SOC
Christopher Rick, PAIA
David Schwegman, PAIA
Saied Toossi, PAIA
Huong Tran, ECON
Joaquin Urrego, ECON
Yao Wang, ECON
Yi Yang, ECON
Xiaoyan Zhang, ECON
Bo Zheng, PAIA
Dongmei Zhu, SOC. SCI.

STAFF

Joanna Bailey, Research Associate
Joseph Boskovski, Manager, Maxwell X Lab
Katrina Fiacchi, Administrative Specialist
Michelle Kincaid, Senior Associate, Maxwell X Lab
Emily Minnoe, Administrative Assistant
Candi Patterson, Computer Consultant
Samantha Trajkovski, Postdoctoral Scholar
Laura Walsh, Administrative Assistant

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 peers in her charge. The model forms a simultaneous equation system that can be estimated by Quasi-Maximum Likelihood. We apply a restricted version of our model to data on National Basketball Association games, where agents are players, networks are individual teams organized by coaches, and competition is head-to-head. That is, at any time a player only interacts with two networks: their team and the opposing team. We find significant positive within-team peer-effects and both negative and positive opposing-team competitor-effects in NBA games. The former are interpretable as "team chemistries" which enhance the individual performances of players on the same team. The latter are interpretable as "team rivalries," which can either enhance or diminish the individual performance of opposing players.

JEL No.: C13, C31, D24

Keywords: Spatial Analysis, Peer Effects, Endogeneity, Machine Learning

Authors: William C. Horrace, Department of Economics, Center for Policy Research, Syracuse University; Hyunseok Jung, Department of Economics, University of Arkansas; Shane Sanders, Department of Sports Management, Syracuse University

1. Introduction

We consider a world with R interrelated 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 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. In most cases, within-network peer 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. 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 where peers allocates effort to simultaneous outcomes, such as labor and leisure hours).

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 planner selects the network strategically and simultaneously with output (Olley and Pakes, 1996), then the network itself may be endogenous. Following Horrace et al. (2016), we augment the outcome equation with a team 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 et al. (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 generalizes their study. It should be noted that the conceptual foundations for selection correction in this way can be traced to papers by Brock and Durlauf. Brock and Durlauf (2006) generalize the multinomial-choice model with social interactions proposed in Brock and Durlauf (2002) and suggest that multinomial control functions may be used for self-selection corrections.

Social network interactions have been studied extensively.[1] However, simultaneous cross-network interactions remain unexplored. A few papers model simultaneous activity for a single, time-invariant network, and are multivariate extensions of the single equation spatial auto-regressive model (Cliff and Ord, 1973, 1981) to simultaneous outcomes. Some are for a cross-section of data, and others are for panel data. For example, Kelejian and Prucha (2004) generalize the spatial auto-regressive (SAR) model to a simultaneous system for cross-section. Baltagi and Deng (2015) extend the model to panel data with random-effects, while Cohen-Cole et al. (2017) extend it to a simultaneous system with fixed-effects. Yang and Lee (2017) study identification and Quasi-Maximum Likelihood (QML) estimation of the model of Kelejian and Prucha (2004) for cross-section. Empirical examples of these simultaneous models include the effect of regional 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 simultaneous peer-effect models are clearly related, but none consider multiple peer networks that may be engaged in simultaneous competition around a single outcome variable, nor do they attempt to model the actions of a manager. Our model does both and in certain cases could be applied to traditional panel data, where outcomes of peers and competitors are observed in every period. Here, however, only the actions of the manager are observed in every period, and she controls which peers will compete and be observed in any period to produce the outcome (e.g., sales). Therefore, our networks are time-varying, the model is conceptualized for repeated cross-sections (not panel), and asymptotic arguments are for large n_r, the number of peers competing in network $r = 1, ..., R$, with a fixed number of networks, R, following Yang and Lee (2017). Nonetheless, our networks vary over time ($t = 1, ..., T$), so large T helps with consistent QML estimation of peer-effects, as our simulations show. Simulations also show that when the networks are small ($n_r = 5$), bootstrap bias correction improves finite sample performance of the estimator.

Team chemistry (a within-team peer-effect) receives substantial empirical attention as a factor influencing performance in sports and business. Unfortunately, team chemistry is difficult to measure.[2] McCaffery and Bickart (2013) estimate team chemistry as a function of biological synchrony among players, while 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

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

[2] See Schrage (2014) who describes it as the "new holy grail of performance analytics" in sport and business.

et al. (2016) develop a network production function model that estimates within-team peer-effects on player productivity in men's college basketball, but without competitor-effects.[3] Our research contributes to this empirical literature on team chemistry by controlling for the play and strategy of the opposing team, and may be conceptualized as a generalization of their model.

We apply a restricted version of our model (i.e., where competition is head-to-head) to the 2015-16 National Basketball Association (NBA) regular season to simultaneously estimate within-network and cross-network peer-effects for all thirty teams in the league, a total of 30×30 heterogeneous effects. We find within-team peer-effects that are statistically significant and positive in the NBA. These are performance multipliers that enhance the individual performance of players on the same team and may be interpreted as "team chemistries." We also find both positive and negative opposing-team competitor-effects. That is, the team chemistry of your opponents may either enhance or diminish the individual performance of players on your team. These effects may be loosely interpreted as "team rivalries." We correct for managerial selectivity bias using both the semi-parametric Dahl (2002) and parametric Lee (1983) approaches, and estimate the coach's selection probabilities from a selection equation in two ways: standard multinomial logit (MNL) and a random forest (RF) algorithm from the machine-learning literature.

The rest of this paper is organized as follows. The next section introduces the econometric specification and estimation approaches, and Section 3 provides a simulations study of the proposed estimator. Section 4 presents the result of the empirical exercise, and Section 5 concludes.

2. Econometric Model and Estimation Strategy

2.1 Outcome Function

We begin with a general statement of the outcome function, but consider a restricted version in our application. There are 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 around a common outcome. The outcome function for the r^{th} network is:

$$\mathbf{y}_{rt} = \lambda_{rr}\mathbf{W}_{rrt}\mathbf{y}_{rt} + \sum_{k\neq r}\lambda_{rk}\mathbf{W}_{rkt}\mathbf{y}_{kt} + \mathbf{X}_{1rt}\beta_{1r} + \iota_{n_r}\mathbf{x}_{2rt}^{'}\beta_{2r} + \mathbf{u}_{rt}, \quad r = 1, ..., R, \quad t = 1, ..., T, \tag{1}$$

where $\mathbf{y}_{rt}$ is an $n_r \times 1$ outcome vector for the r^{th} network, $\mathbf{X}_{1rt}$ is an $n_r \times p_1$ exogenous input matrix

[3] This may induce an omitted variable bias, so their peer-effects may not be *ceteris paribus* estimates of team chemistry.

that varies over peers $i = 1, .., n_r$ and t (peer-level variability), ι_{n_r} is an $n_r \times 1$ vector of ones, $\mathbf{x}_{2rt}$ is an $p_2 \times 1$ exogenous input vector that varies over t only (network-level variability), and $\mathbf{u}_{rr}$ is an $n_r \times 1$ disturbance vector. The $\mathbf{W}_{rrr}$ is an $n_r \times n_r$ weight matrix for interaction within the r^{th} network, while $\mathbf{W}_{rkr}$ 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 peer-effect for the r^{th} network, and λ_{rk} is the average cross-network competitor-effect for $k \neq r$ to the r^{th} network. The terms β_{1r} and β_{2r} are vectors of input coefficients for the r^{th} network. The existing literature assumes that $\lambda_{rk} = 0$ for $k \neq r$. We allow within-network and cross-network effects to be positive or negative. We may also refer to $\mathbf{W}_{rrt}$ as the peer network and $\mathbf{W}_{rkt}$ as the competitor network.[4]

Following Horrace et al. (2016), each network r has a manager, who populates her network with peers in each period t, selecting n_r peers from a larger group of peers at her disposal. Peers work together within their networks and compete across networks to produce $\mathbf{y}_{rt}$.[5] The strategic decisions of the managers have implications for the econometric model. First, we only observe the outcome y_{irt} (say) for peer i in period t, when he is selected into network r by his manager. As such our data are not a panel *per se* and should be considered *repeated cross-sections*. This is an important distinction between our model and other social network models for panel data, where each peer is observed in each period and networks are often fixed over time (e.g., Lee and Yu, 2010, 2014). Indeed, lack of a true panel is fundamental to the need for managerial selection correction.

Also, the simultaneous actions of the managers induce a network-level correlated effect (Manski, 1993). Without loss of generality we assume that all manager's have the same number of workers at their disposal, $n_0 > n_r$. Let d_{irt} be an indicator variable such that $d_{irt} = 1$, if worker i is assigned to network r in period t, and $d_{irt} = 0$, otherwise. Then all managerial decisions at time t can be characterized by the $n_0 \times R$ matrix $\mathbf{D}_t = (\mathbf{d}_{1t}, ..., \mathbf{d}_{Rt})$ with $\mathbf{d}_{rt} = (d_{1rt}, ..., d_{n_{r0}rt})'$. The correlated effect then implies $E(\mathbf{u}_{rt}|\mathbf{D}_t) \neq \mathbf{0}$, and $\mathbf{D}_t$ is clearly correlated with all the variables and networks on the right-hand side of (1). Ultimately, we use arguments from the game theory literature for static games of incomplete information with multiple equilibria (i.e., Bajari et al., 2010) to model managerial decisions and their effects on network peers and competitors, but for now we assume that the actions

[4]Additionally, one can include $\mathbf{W}_{rrt}\mathbf{x}_{rt}$ terms into the model to separately identify exogenous social effects from the endogenous social effects. We follow Horrace et al. (2016) and do not include exogenous network effects in the outcome function.

[5]There are other interpretations of the model. For example, we can think of t as representing distinct markets, where networks (firms, teams) might compete, but not all networks may be present in each market. If any network r or k does not compete in market t, then $\lambda_{rk} = 0$.

of the managers affect all active peers in network r in the same way. Hence, the correlated effects reduce to time-varying network fixed-effects. That is, $E(\mathbf{u}_{rt}|\mathbf{D}_t) = \iota_{n_r}\alpha_{rt}$, where α_{rt} is a scalar. Let $\mathbf{u}^*_{rt} = \mathbf{u}_{rt} - \iota_{n_r}\alpha_{rt}$. Also, let $\mathbf{X}_{rt} = [\mathbf{X}_{1rt}, \iota_{n_r}\mathbf{x}'_{2rt}]$ be a $n_r \times p$ matrix with $p = p_1 + p_2$ and $\beta_r = [\beta'_{1r}, \beta'_{2r}]'$, then the entire system at time t is,

$$\mathbf{y}_t = \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rkt}\mathbf{y}_t + \mathbf{X}_t\beta + \alpha_t + \mathbf{u}^*_t, \quad t = 1, ..., T, \tag{2}$$

where $\mathbf{y}_t = (\mathbf{y}'_{1t}, ..., \mathbf{y}'_{Rt})'$, $\mathbf{X}_t = Diag(\mathbf{X}'_{1t}, ..., \mathbf{X}'_{Rt})'$, $\alpha_t = (\iota'_{n_1}\alpha_{rt}, ..., \iota'_{n_R}\alpha_{Rt})'$, $\mathbf{u}^*_t = (\mathbf{u}^{*'}_{1t}, ..., \mathbf{u}^{*'}_{Rt})'$ and $\beta = (\beta'_1, ..., \beta'_R)'$. $\mathbf{G}_{rkt}$ is an $N \times N$ block matrix with R row blocks and R column blocks. The blocks in $\mathbf{G}_{rkt}$ are all blocks of zeros except for r^{th} row block in the k^{th} column block position, which equals $\mathbf{W}_{rkt}$. For example, if there are $R = 2$ networks, then $\mathbf{G}_{11t} = \left[\begin{smallmatrix}\mathbf{W}_{11t} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{smallmatrix}\right]$, $\mathbf{G}_{22t} = \left[\begin{smallmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_{22t}\end{smallmatrix}\right]$, $\mathbf{G}_{12t}$ $= \left[\begin{smallmatrix}\mathbf{0} & \mathbf{W}_{12t}\\ \mathbf{0} & \mathbf{0}\end{smallmatrix}\right]$, and $\mathbf{G}_{21t} = \left[\begin{smallmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{W}_{21t} & \mathbf{0}\end{smallmatrix}\right]$ where $\mathbf{0}$ is a conformable matrix of zeros. We require the following regularity conditions.

Assumption 1. *Let u^*_{irt} be the i^{th} element of $\mathbf{u}^*_{rt}$. The u^*_{irt} are $iid(0, \sigma^2_r)$ and a moment of order higher than the fourth exists.*

Assumption 2. *$|\mathbf{S}_t(\mathbf{\Lambda})| > 0$ for any $\mathbf{\Lambda}$ in its parameter space, $\mathbb{L}$, where $\mathbf{S}_t(\mathbf{\Lambda}) = \mathbf{I}_N - \mathbf{M}_t(\mathbf{\Lambda})$ with $\mathbf{M}_t(\mathbf{\Lambda}) = \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rkt}$, $\mathbf{\Lambda} = (\mathbf{\Lambda}'_1, ..., \mathbf{\Lambda}'_R)'$ and $\mathbf{\Lambda}_r = (\lambda_{r1}, ..., \lambda_{rR})$. $\mathbb{L}$ is compact and the true $\mathbf{\Lambda}_0$ is in its interior.*

Assumption 3. *The $\mathbf{W}_{rkt}$ are uniformly bounded in both row and column sums in absolute value. Also, $\mathbf{S}_t(\mathbf{\Lambda})^{-1}$ is uniformly bounded, uniformly in $\mathbf{\Lambda} \in \mathbb{L}$.*

Assumption 1-3 are standard and follow Lee et al. (2010) and Yang and Lee (2017) with a few differences. Following Yang and Lee (2017), consistent QML estimation of the outcome function in (2) proceeds as $n_r \to \infty$. However, due to independence across t in Assumption 1, the time dimension helps with consistency.[6] In particular, it allows us to appeal to the arguments of Lee et al. (2010). They consider a cross-section ($T = 1$) of a large number of *independent* networks with a *homogeneous* peer-effect, and prove that the effect may be consistently estimated with QML, as either the number of peers in a network or the number of independent networks increases. Similar to Lee et al. (2010), the system-wide network $\sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rkt}$ in (2) can be seen as a single network at time t. The

[6]The *iid* assumption is restrictive, but as stated above we do not have a panel *per se*, so methods to account for possible heteroskedasticity or autocorrelation are limited. One may relax the assumption by introducing a parametric variance function or spatial autoregressive error structure, but this is left for future research.

network changes over time, as managers change peer composition, and can be seen as T networks in a single period, so asymptotic arguments similar to Lee at al. (2010) apply. That is, QML is consistent as $n_r T \to \infty$ for fixed R with u^*_{irt} independent across t, and this is borne out in our simulations study.[7] Relative to Lee et al. (2010), the only complication that QML estimation of (2) presents is that we have multiple peer-effects (λ_{rr}) and competitor effects (λ_{rk}). However, they are time invariant, so time variation in the networks $\mathbf{W}_{rrt}$ and $\mathbf{W}_{rkt}$ helps reduce the mean squared errors of their estimates.

The compact parameter space of Assumption 2 is also standard and ensures consistency and desirable asymptotic properties of QML. The implication is that $\mathbf{S}_t(\mathbf{\Lambda})$ can be inverted, ensuring stability and an equilibrium where $\mathbf{y}_t = \mathbf{S}_t(\mathbf{\Lambda})^{-1}(\mathbf{X}_t\beta + \mathbf{u}_t)$. A positive determinant ensures the likelihood function is well-defined. Following Yang and Lee (2017), since $|\mathbf{S}_t(\mathbf{\Lambda})| = \Pi_{i=1}^{N}(1-\tau_{it})$ where τ_{it} is an eigenvalue of $\mathbf{M}_t(\mathbf{\Lambda})$, a sufficient condition satisfying this assumption is $\max_i \tau_{it} < 1$. From the spectral radius theorem, $\max_i |\tau_{it}|$ is less than any of its induced matrix norms, so the condition is satisfied when $||\mathbf{M}_t(\mathbf{\Lambda})||_\infty < 1$ or $||\mathbf{M}_t(\mathbf{\Lambda})||_1 < 1$. Since $||\mathbf{M}_t(\mathbf{\Lambda})||_\infty \leq (\max_r \sum_{k=1}^{R} |\lambda_{rk}|)(\max_{r,k} ||\mathbf{W}_{rkt}||_\infty)$ and $||\mathbf{M}_t(\mathbf{\Lambda})||_1 \leq (\max_k \sum_{r=1}^{R} |\lambda_{rk}|)(\max_{r,k} ||\mathbf{W}_{rkt}||_1)$, a sufficient condition for Assumption 2 is that $\mathbb{L}$ is restricted such that either $\sum_{k=1}^{R} |\lambda_{rk}| < (\max_{r,k} ||\mathbf{W}_{rkt}||_\infty)^{-1}$ or $\sum_{r=1}^{R} |\lambda_{rk}| < (\max_{r,k} ||\mathbf{W}_{rkt}||_1)^{-1}$ for all r, k. The first condition reduces to $\sum_{k=1}^{R} |\lambda_{rk}| < 1$ when the network matrices are row-normalized.[8] This means the sum of the absolute values of the within- and cross-network effects from or to a network must be bounded, while the row or column sum of $\mathbf{W}_{rkt}$ is also bounded.

These conditions have important implications for our model. First, $\mathbf{\Lambda}$ should be sparse if R is large, so the effective number of network effects from or to a network is bounded. Second, the number of linkages for each peer in $\mathbf{W}_{rkt}$ for all r, k should be fixed as n_r increases. That is, $\mathbf{W}_{rkt}$ should be sparse as n_r increases. This is discussed in Lee (2004) and is borne out in our simulations, as the network effect cannot be consistently estimated if $\mathbf{W}_{rkt}$ is not sparse when n_r increases. Here, R is fixed. Increasing R will produce richer network interaction structure, but it increases model complexity. Moderate to small R may be appropriate in our context as network-pairwise interactions may only exist when there are few networks in the market. In a market with many networks, the effect of a single network on other networks may be negligible.

Assumption 3 is standard and limits spatial dependence to a manageable degree (e.g., Kelejian and Prucha, 1999). This assumption will be met if both the sufficient conditions for Assumption 2 (above)

[7] As we shall see, large T is necessary to consistently estimate β_{2r} after first-stage estimation of (2).

[8] These are sufficient conditions, which can be relaxed using the stationary-region search methodology proposed by Elhorst, Lacombe and Piras (2012).

are satisfied. We discuss sufficient conditions for identification of the model in the next section.

2.1.1 QML Estimation of the Outcome Function

We focus on Quasi-Maximum Likelihood (QML) estimation of the outcome function in (2).[9] Following Lee et al. (2010), we remove the bias term α_t to avoid the incidental parameter problem (Neyman and Scott, 1948) by transforming the model with projector $\mathbf{J}_Q = Diag(\mathbf{Q}_1, ..., \mathbf{Q}_R)$, where $\mathbf{Q}_r$ is the "within" transformation matrix, $\mathbf{Q}_r = \mathbf{I}_{n_r} - \frac{1}{n_r}\iota_{n_r}\iota_{n_r}^{'}$. Now, $\mathbf{Q}_r\iota_{n_r} = \mathbf{0}$ and $\mathbf{Q}_r\mathbf{u}_r^* = \mathbf{Q}_r\mathbf{u}_r$. Therefore, we have,

$$\mathbf{J}_Q\mathbf{y}_t = \mathbf{J}_Q\sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rkt}\mathbf{y}_t + \mathbf{J}_Q\mathbf{X}_{1t}\beta_1 + \mathbf{J}_Q\mathbf{u}_t, \quad t = 1, ..., T, \tag{3}$$

where $\mathbf{X}_{1t} = Diag(\mathbf{X}_{11t}^{'}, ..., \mathbf{X}_{1Rt}^{'})^{'}$.

Extensive study of identification conditions for network models and multivariate SAR models can be found in Bramoulle et al. (2009), Cohen-Cole et al. (2017) and Yang and Lee (2017). Let $\mathbf{\Xi}_r$ be a $n_r \times N$ matrix consisting of R horizontally concatenated blocks of size n_r, with I_{n_r} in the r^{th} position and zeroes in the other $R-1$ positions. Hence, $\mathbf{\Xi}_1 = (\mathbf{I}_{n_1}, \mathbf{0})$, $\mathbf{\Xi}_R = (\mathbf{0}, \mathbf{I}_{n_R})$, and $\mathbf{\Xi}_r = (\mathbf{0}, \mathbf{I}_{n_r}, \mathbf{0})$ for $r \neq 1, R$, where the $\mathbf{0}$ matrices are appropriately conformable. Let $\mathbf{\Theta}_{rkt} = \mathbf{\Xi}_r\mathbf{G}_{rkt}\mathbf{S}_{0t}^{-1}\mathbf{X}_{1t}\beta_{1,0}$ and $\mathbf{S}_{0t} = \mathbf{I}_N - \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk,0}\mathbf{G}_{rkt}$. In the Online Appendix, we show that the true parameters $\mathbf{\Lambda}_0$ (as defined in Assumption 2) and $\beta_{1,0}$ can be identified from (3) under the condition,

Assumption 4. *The matrix* $[\mathbf{Q}_r\mathbf{X}_{1rt}, \mathbf{Q}_r\mathbf{\Theta}_{r1t}, ..., \mathbf{Q}_r\mathbf{\Theta}_{rRt}]$ *have full column rank* $\forall r$ *and some* t.

The identification condition corresponds to the conditions in Liu and Lee (2010) and Yang and Lee (2017), and it will be generally satisfied here, because we have multiple network matrices and exogenous regressors from each r, which produces enough variation to identify the coefficients in our model.[10]

The disturbances in (3) are linearly dependent because the covariance matrix $\sigma_r^2\mathbf{Q}_r$ is singular. Following Lee et al. (2010, p.150), we consider "an essentially equivalent but more effective transformation" to eliminate the network fixed-effects. Let the orthonormal matrix of $\mathbf{Q}_r$ be $[\mathbf{P}_r, \iota_{n_r}/\sqrt{n_r}]$. The columns in $\mathbf{P}_r$ are eigenvectors of $\mathbf{Q}_r$ corresponding to the eigenvalue one, such that $\mathbf{P}_r^{'}\iota_{n_r} = \mathbf{0}$,

[9] One may also consider G2SLS or G3SLS 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). GMM (e.g., Lin and Lee, 2010) may be more appropriate when there is unknown heteroskedasticity at the peer-level.

[10] An example that doesn't satisfy the condition is when the $\mathbf{W}_{rkt}$ are complete, which is the case for our empirical application. In section 4, we discuss an exclusion restriction in the weighting matrices to address the issue. When exogenous network effects ($\mathbf{W}_{rrt}\mathbf{X}_{rt}$, say) are included in the model, the R matrices in Assumption 4 include additional sub-matrices that are higher orders of $\mathbf{G}_{rrt}$.

$\mathbf{P}_r^{'}\mathbf{P}_r = \mathbf{I}_{n_r-1}$ and $\mathbf{P}_r\mathbf{P}_r^{'} = \mathbf{Q}_r$. Then, premultiplying (2) by $\mathbf{J}_P^{'} = Diag(\mathbf{P}_1^{'}, ..., \mathbf{P}_R^{'})$ leads to

$$\mathbf{J}_P^{'}\mathbf{y}_t = \mathbf{J}_P^{'}\sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rkt}\mathbf{y}_t + \mathbf{J}_P^{'}\mathbf{X}_{1t}\beta_1 + \mathbf{J}_P^{'}\mathbf{u}_t. \tag{4}$$

Let $\bar{\mathbf{y}}_t = \mathbf{J}_P^{'}\mathbf{y}_t$, $\bar{\mathbf{X}}_{1t} = \mathbf{J}_P^{'}\mathbf{X}_{1t}$, $\bar{\mathbf{u}}_t = \mathbf{J}_P^{'}\mathbf{u}_t$, $\bar{\mathbf{G}}_{rkt} = \mathbf{J}_P^{'}\mathbf{G}_{rkt}\mathbf{J}_P$. Now, $\mathbf{J}_P^{'}\mathbf{G}_{rkt} = \bar{\mathbf{G}}_{rkt}\mathbf{J}_P^{'}$.[11] This implies,

$$\bar{\mathbf{y}}_t = \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\bar{\mathbf{G}}_{rkt}\bar{\mathbf{y}}_t + \bar{\mathbf{X}}_{1t}\beta_1 + \bar{\mathbf{u}}_t. \tag{5}$$

We derive the QML function and the concentrated QML function of (5) in the Online Appendix. Call the QML estimates $\hat{\mathbf{\Lambda}}$ and $\hat{\beta}_1$ the "first-step" estimates. If all we care about is the peer- and competitor-effects, then this is the only estimation step. However, if estimation of β_{2r} and managerial selection biases are important, then there are two additional estimation steps to consider. In the "second-step" we estimate probabilities that the manager and opposing managers select teams of peers and competitors into the peer- and competitor-networks, respectively. Estimation of these "selection probability proceeds with Multinomial Logit (MNL) or Random Forest (RF), and allows us to calculate the network-level selection bias of each manager as a single index.[12] These procedures include exogenous state variables, $\mathbf{z}_{rt}$, which explain the managers' selections and will be discussed later. In a "final-step," we estimate β_2 with a regression of the residuals from the first-step on network-level covariates, $\mathbf{x}_{2t}$ and the index.

2.2 Bias Due to Strategic Interactions

Thus far, we have assumed that the endogeneity induced by the manager's choices can be controlled with a network level fixed-effect. The fixed-effect is then removed from the model with the within transformation, so that the peer-effects and competitor-effects can be consistently estimated as $n_r \to \infty$ with QMLE. We may also estimate coefficients on the explanatory variables that vary at the peer-level ($\mathbf{X}_{1t}$). Unfortunately, coefficients on the explanatory variables that vary at the network-level ($\mathbf{x}_{2t}$) are not identified. To address this issue, we set up a static game in period t to formulate and estimate the selection bias.[13] We follow and adapt the basic methodologies in the game theory literature for static

[11] $\bar{\mathbf{G}}_{rkt}\mathbf{J}_P^{'} = \mathbf{J}_P^{'}\mathbf{G}_{rkt}\mathbf{J}_Q = \mathbf{J}_P^{'}\mathbf{G}_{rkt}(\mathbf{I} - Diag(\iota_{n_1}\iota_{n_1}^{'}/n_1, ..., \iota_{n_R}\iota_{n_R}^{'}/n_R)) = \mathbf{J}_P^{'}\mathbf{G}_{rkt}$, because $\mathbf{G}_{rkt}$ is row-normalized so $\mathbf{J}_P^{'}\mathbf{G}_{rkt}Diag(\iota_{n_1}\iota_{n_1}^{'}/n_1, ..., \iota_{n_R}\iota_{n_R}^{'}/n_R) = \mathbf{0}$.

[12] Other flexible estimation methods may be used, such as kernel smoothing or local polynomial regressions (e.g. Bajari et al. (2010)). RF is considered here as is known to be effective in handling high dimensional data and in predicting nonlinear relationships.

[13] There may be other econometric remedies to address the endogeneity issue. Recently, two categories of methodologies

games of incomplete information with multiple equilibria. Arguments follow Bajari et al. (2010).[14]

Each network $r = 1, ..., R$ has a network manager who takes action a_{rt} at time t from her finite and discrete choice set of actions, $a_{rt} \in \mathbb{A} = [0, 1, ..., K]$, so $\mathbf{a}_t^{'} = (a_{1t}, ..., a_{Rt})$ is a vector of actions for all managers.[15] Ideally, $\mathbb{A}$ would be the set of all possible combinations of n_r workers from the n_0 workers at the manager's disposal (i.e., all possible values of $\mathbf{d}_{rt}$, so the set of observed actions of all managers corresponds to a mapping from the matrix $\mathbf{D}_t$). However, in practice considering all combinations may be infeasible and econometrically undesirable. One solution would be to restrict $\mathbb{A}$ to only those choices that are observed in the sample (see Horrace et al., 2016). For example, an NBA coach may never choose a lineup of his five worst players, so we can exclude this action from the choice set. Alternatively, if there is some variable that maps the characteristics of each potential combination of workers into fewer choices, then this may be more practical. This may also allow us to specify adjacency matrices that satisfy Assumption 4. For example, an NBA coach may be interested in creating positional mismatches (e.g., guards defending forwards and vice versa) between his players and the opposing team's players. Here, the choice of the number of guards may summarize the coach's actions, so $K = 6$, and it may imply that network interactions should be based on player position types. We use this action set and a network adjacency matrix based on player position in our NBA application.

Let $\mathbf{a}_{-rt}^{'} = (a_{1t}, ..., a_{(r-1)t}, a_{(r+1)t}, ..., a_{Rt})$ be the vector of actions for all managers, excluding manager r. Each manager has a vector of exogenous state variables $\mathbf{z}_{rt}$, so the vector of state variables (i.e., market conditions) for all R managers is $\mathbf{z}_t = (\mathbf{z}_{1t}^{'}, ..., \mathbf{z}_{Rt}^{'})^{'}$.[16] The state variable $\mathbf{z}_t$ is common knowledge and observable to the econometrician. Manager r is also subject to an idiosyncratic shock over her possible actions, $e_{rt}(a_{rt}) \in [e_{rt}(0), ..., e_{rt}(K)]$. These shocks, e_{rt}, are *iid* over a_{rt} and over r and t with density $G_e(e_{rt})$. The shocks are private information to manager r, and are unobservable to the econometrician. Let $\mathbf{e}_{rt}$ be the $K + 1$ vector of $e_{rt}(a_{rt})$ for all $a_{rt} \in \mathbb{A}$. Then, managers simultaneously choose their actions based on their individual information sets $\{\mathbf{z}_t, \mathbf{e}_{rt}\}$, so manager r's decision rule is a scalar function $a_{rt} = \eta_r(\mathbf{z}_t, \mathbf{e}_{rt})$. Under the chosen actions, networks produce single outcome, $\mathbf{y}_t$.

Let manager r's scalar utility function be $\mathrm{v}_r(\mathbf{a}_{rt}, \mathbf{z}_t, \mathbf{e}_{rt}) = \pi_r(a_{rt}, \mathbf{a}_{-rt}, \mathbf{z}_t) + e_{rt}(a_{rt})$. Then, the

have been proposed to address endogeneity in formation of spatial or network links: One is a (Bayesian) "One-step Full Information" approach by Goldsmith and Imbens (2013) and Hsieh and Lee (2016), and the other is a "Multiple-step Control Function" approach by Qu and Lee (2015) and Horrace et al. (2016).

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

[15] Following Bajari et al. (2010), we assume without loss of generality that $\mathbb{A}$ is the same for all managers.

[16] Per Hoshino (2019) it is desirable that $\mathbf{z}_{rt}$ contain variables excluded from (2).

conditional choice probability of r choosing a_r at a given realization of $\mathbf{z}_t$ is given by

$$\delta_r(a_{rt} = k|\mathbf{z}_t) = \int \mathbf{1}\{\eta_r(\mathbf{z}_t, \mathbf{e}_{rt}) = k\} dG_e(\mathbf{e}_{rt}), \tag{6}$$

which may be interpreted as the beliefs formed by r's opponents regarding r's decision. Here, δ_r is a scalar function. 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 for choosing action a_{rt},

$$V_r(a_{rt}, \mathbf{z}_t, \mathbf{e}_{rt}) = \sum_{\mathbf{a}_{-rt}} \prod_{k \neq r} \delta_k(a_{kt}|\mathbf{z}_t)\pi_r(a_{rt}, \mathbf{a}_{-rt}, \mathbf{z}_t) + e_{rt}(a_{rt}) = \varphi_r(a_{rt}, \mathbf{z}_t) + e_{rt}(a_{rt}). \tag{7}$$

Here, φ_r is the deterministic part of the expected payoff function and is a scalar function. 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 a_{rt} will be chosen for a given realization of $\mathbf{z}_t$, if and only if

$$V_r(a_{rt}, \mathbf{z}_t, \mathbf{e}_{rt}) > \max_{a'_{rt} \neq a_{rt}} V_r(a'_{rt}, \mathbf{z}_t, \mathbf{e}_{rt}) \Leftrightarrow \varphi_r(a_{rt}, \mathbf{z}_t) + e_{rt}(a_{rt}) > \max_{a'_{rt} \neq a_{rt}} \varphi_r(a'_{rt}, \mathbf{z}_t) + e_{rt}(a'_{rt}), \tag{8}$$

for $a_{rt}, a'_{rt} \in \mathbb{A}$. It follows that

$$\delta_r(a_{rt}|\mathbf{z}_t) = \Pr[\max_{a'_{rt} \neq a_{rt}} \varphi_r(a'_{rt}, \mathbf{z}_t) - \varphi_r(a_{rt}, \mathbf{z}_t) + e_{rt}(a'_{rt}) - e_{rt}(a_{rt}) < 0],$$

in equilibrium.[17] The equilibrium can be seen as a Bayesian-Nash Equilibrium (BNE) in a probability space with measure δ_r, $r = 1, ..., R$, and summarized by a best-response mapping that maps a compact set to itself (i.e. $[0,1]^R \to [0,1]^R$). It is also continuous in δ_r, so equilibrium existence follows from Brouwer's fixed-point theorem.[18] The next section formulates the selection bias from δ_r.

2.2.1 Formulation of the Selection Bias

A standard approach to formulate the selection bias is to assume the outcome error, $\mathbf{u}_{rt}$, and the payoff function error, $e_{rt}(a_{rt})$, are statistically dependent. This idea can be traced back to Heckman's (1979) selection model and Lee's (1983) generalization. Let $\epsilon^*_{rt} = \max_{a'_{rt} \neq a_{rt}} \varphi_r(a'_{rt}, \mathbf{z}_t) - \varphi_r(a_{rt}, \mathbf{z}_t) + e_{rt}(a'_{rt}) - e_{rt}(a_{rt})$ and let $\boldsymbol{\varphi}_r(\mathbf{z}_t) = [\varphi_r(0, \mathbf{z}_t), ..., \varphi_r(K, \mathbf{z}_t)]$ be a $K+1$ vector,[19] so $g_{ir}(u_{irt}, \epsilon^*_{rt}|\boldsymbol{\varphi}_r(\mathbf{z}_t))$

[17]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.

[18]See Bajari et al. (2010) for a simple example.

[19]The ϵ^*_{rt} is a function of a_{rt}, but it is suppressed here for notational simplicity.

is the joint distribution of u_{irt} and ϵ^*_{rt}, which depends on $\varphi_r(\mathbf{z}_t)$ and may differ across i, in general. Following Bourguignon et al. (2007), if the two errors are dependent,

$$\begin{aligned} E(u_{irt}|a_{rt}, \varphi_r(\mathbf{z}_t)) = E(u_{irt}|\epsilon^*_{rt} < 0, \varphi_r(\mathbf{z}_t)) &= \int \int_{-\infty}^{0} \frac{u_{irt} g_{ir}(u_{irt}, \epsilon^*_{rt}|\varphi_r(\mathbf{z}_t))}{Pr(\epsilon^*_{rt} < 0|\varphi_r(\mathbf{z}_t))} d\epsilon^*_{rt} du_{irt} \\ &= \alpha_{ir}(\varphi_r(\mathbf{z}_t)). \end{aligned} \tag{9}$$

The correlation may exist when the two errors contain a common component, unobserved by the econometrician but observed (or predicted) by the manager. In (9) the joint distribution $g_{ir}(\cdot)$ varies by peers in a network, which may be true if individual peers respond to the common shock differently (or the manager selects workers based on individual performances). However, in team production it may be reasonable to assume that the manager is optimizing team performance and not individual performance. Therefore, peer-level heterogeneity may be negligible, conditional on her selection. Also, note again that we do not have a true panel of data, so it is difficult to model and account for peer-level heterogeneity, in general. Therefore, we assume the following to arrive at a network-specific selection bias.

Assumption 5. *The joint distributions of u_{irt} and ϵ^*_{rt} conditional on $\varphi_r(\mathbf{z}_t)$ are identical for every peer $i = 1, ..., n_r$ in network r at time t.*

Then $E(\mathbf{u}_{rt}|a_{rt}, \varphi_r(\mathbf{z}_t)) = \iota_{n_r} \alpha_r(\varphi_r(\mathbf{z}_t))$, which are network specific fixed-effects due to the strategic actions of the network managers. This is what Manski (1993) calls a correlated effect. We simply follow the standard solution to the correlated effects problem by including a network-level fixed-effect, but we explicitly model the correlated effect by modeling a managerial selection equation. This allows us to separately identify the managerial effect from other network-level coefficients (β_{2r}) in (2).

Per Dahl (2002), estimation of the unknown control function $\alpha_r(\varphi_r(\mathbf{z}_t))$ suffers from the "the curse of dimensionality" due to the presence of a large number of alternatives and its dependence on the unknown function $\varphi_r(\mathbf{z}_t)$. To make estimation feasible, restrictions need to be imposed on the control function. We consider parametric and semi-parametric approaches following Lee (1983) and Dahl (2002), respectively.

1. **Lee's approach**: Let the distribution of ϵ^*_{rt} be F_r. Following Lee (1983) and Horrace et al. (2016), we can reduce the dimensionality of the selection bias by the transformation $J_r(\cdot) \equiv \Phi^{-1}(F_r(\cdot))$, where Φ^{-1} is the inverse of the standard normal CDF. Then, $J_r(\epsilon^*_{rt})$ becomes a standard normal random variable. For notational simplicity, let $J_r(\epsilon^*_{rt}) \equiv \epsilon_{rt}$. Per Schmertmann

(1994), Dahl (2002) and Bourguignon et al. (2007), it is implicitly assumed in Lee's approach that the joint distribution of u_{irt} and ϵ_{rt} does not depend on $\varphi_r(\mathbf{z}_t)$. That is, $E[u_{irt}|\epsilon_{rt} < J_r(0), \varphi_r(\mathbf{z}_t)] = E[u_{irt}|\epsilon_{rt} < J_r(0)]$.[20] To obtain an explicit parametric form of the bias, we further assume ϵ_{rt} and u_{irt} are *i.i.d* over i and t with joint normal distribution,

$$\begin{bmatrix} u_{irt} \\ \epsilon_{rt} \end{bmatrix} = N\left(\mathbf{0}, \begin{bmatrix} \sigma_r^2 & \sigma_{12r} \\ \sigma_{12r} & 1 \end{bmatrix}\right),$$

then it can be shown that,

$$\alpha_r(\varphi_r(\mathbf{z}_t)) = -\sigma_{12r} \frac{\phi[\Phi^{-1}\{\delta_r(a_{rt}|\mathbf{z}_t)\}]}{\delta_r(a_{rt}|\mathbf{z}_t)}, \tag{10}$$

a scalar, where $\delta_r(a_{rt}|\mathbf{z}_t)$ is the selection probability from (6).

2. **Dahl's approach**: We make the index sufficiency assumption of Dahl (2002) such that

$$\alpha_r(\varphi_r(\mathbf{z}_t)) = \psi_r(\delta_r(a_{rt}|\mathbf{z}_t)), \tag{11}$$

where $\psi_r(\cdot)$ is an unknown scalar function that may be estimated non-parametrically. Per Dahl, this assumes that the selection probability $\delta_r(a_{rt}|\mathbf{z}_t)$ exhausts all information about the behaviors of the two errors. That is, the joint distribution of u_{irt} and ϵ^*_{rt} depends on $\varphi_r(\mathbf{z}_t)$ only through $\delta_r(a_{rt}|\mathbf{z}_t)$.

2.2.2 Estimation of the Selection Bias

To estimate the strategic bias induced by the managers and the coefficients on network-varying exogenous variables ($\mathbf{x}_{2rt}$), we consider a three-step estimation procedure as follows:

1. **First Step**: Estimate $\hat{\mathbf{\Lambda}}$ and $\hat{\beta}_1$ from (5), and for $r = 1, ..., R$ and $t = 1, ..., T$ compute residuals, $\hat{\upsilon}_{rt} = \iota'_{n_r}(\mathbf{y}_{rt} - \sum_{k=1}^{R} \hat{\lambda}_{rk}\mathbf{W}_{rkt}\mathbf{y}_{kt} - \mathbf{X}_{1rt}\hat{\beta}_{1r})/n_r$.

2. **Second Step**: Estimate the selection probability $\hat{\delta}_r(a_{rt}|\mathbf{z}_t)$ for $r = 1, ..., R$ and $t = 1, ..., T$ using a parametric or nonparametric model. We use MNL for a parametric approach and RF

[20]See Schmertmann (1994), Dahl (2002) and Bourguignon et al. (2007) for more detail about the implications of the restrictions on the joint distribution of the errors in Lee's and Dahl's approaches. The assumption implies that the correlations between u_{irt} and $e_{rt}(a'_{rt}) - e_{rt}(a_{rt})$ must have the same sign for all $a'_{rt} \neq a_{rt}$, which means in our context that the random shock should have the same implication to the different actions (i.e., it can have a positive or negative effect, but should be the same for all the actions). This is relaxed slightly in the Dahl's approach where the bivariate covariances can differ in sign but only as a function of the $\delta_r(a_{rt}|\mathbf{z}_t)$.

for a nonparametric approach. The statistical properties of the MNL are well-known, and one can use the Murphy and Topel (1985) correction to account for sampling variability of the estimated probabilities in calculating standard errors in the final-step estimates below. For RF we use the bias reduction method proposed by Chernozhukov et al. (2018) which uses Neyman-orthogonal scores to remove bias in semiparametric models, where the nonparametric component is estimated with a machine learning method. The method allows the final-step estimators to be $\sqrt{T}$-consistent, even when RF converges more slowly. To ensure that RF consistently estimates $\hat{\delta}_r$, we restrict the maximum number of splits in a tree (equivalently, the depth of a tree) based on the asymptotic results of Scornet et al. (2015) and Wager and Walther (2015).[21] In simulations, we use two types of cross-validation methods to tune the maximum number of splits, and examine their effect on the mean squared error of the final-step estimates.[22]

3. **Final Step**: The selection bias and β_{2r} can be estimated from the OLS regression $\hat{\upsilon}_{rt} = \mathbf{x}_{2rt}\beta_{2r} + \gamma_r(\hat{\delta}_r(a_{rt}|\mathbf{z}_t)) + \xi_{rt}$ where ξ_{rt} is an $i.i.d.$ error term and $\gamma_r(\hat{\delta}_r(a_{rt}|\mathbf{z}_t))$ is either given by the right-hand side of (10) with $\hat{\delta}$ substituted for δ (Lee's approach) or by $\psi_r(\cdot)$ in (11) (Dahl's approach). For Dahl's approach, we use polynomial approximations of ψ_r and select polynomial order using the Akaike Information Criterion (AIC).[23] Newey (1994, 1997, 2009) and Andrew (1991) provide regularity conditions and basis functions for semi-parametric models, like Dahl's approach.[24] In our simulations and application, we use the bootstrap to calculate standard errors of the final-step estimates, thereby accounting for variability in estimates upon which they are based.

3. Simulation Experiments

Our simulation study is for $R = 3$ networks. To give a concrete example, we consider three real estate companies competing in a region. Each group has $n_r = 5$ (or $n_r = 10$) agents and each agent in a group is responsible for different parts of the region. We assume that if the agents' areas of responsibility are adjacent, then they are competing across networks or cooperating within their network. We randomly assign agents into areas within the region and use a contiguity weighting scheme to specify connections. We generate random matrices, W_{rkt} for $r, k = 1, 2, 3$, where each entry is randomly assigned with 0

[21] The consistency results in Scornet et al. (2015) and Wager and Walther (2015) are for regression trees, while we use classification trees. We conjecture the their regularity condition is appropriate here. Simulations support this conjecture.

[22] In our simulation and application, random forest is implemented in Matlab using the command $fitcensemble$.

[23] Newey et al. (1990) use the Generalized Cross-Validation (GCV) of Craven and Wahba (1979) in this context, which is approximately equivalent to AIC (Wang et al., 2007).

[24] For consistency and asymptotic normality, the number of basis functions should increase with the sample size. The number of basis functions will be selected by the econometrician in practice.

or 1, while maintaining symmetry of the interaction matrix, $\mathbf{W}_t = \begin{bmatrix} \mathbf{W}_{11t} & \mathbf{W}_{12t} & \mathbf{W}_{13t} \\ \mathbf{W}_{21t} & \mathbf{W}_{22t} & \mathbf{W}_{23t} \\ \mathbf{W}_{31t} & \mathbf{W}_{32t} & \mathbf{W}_{33t} \end{bmatrix}$. We then row-normalize each $\mathbf{W}_{rkt}$. We set $\mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} = \begin{bmatrix} 0.4 & -0.2 & 0.1 \\ 0.2 & 0 & 0 \\ -0.1 & 0 & 0.2 \end{bmatrix}$ so the first network has the strongest within network effect ($\lambda_{11} = 0.4$) but is also strongly affected by the other two networks: negatively by the second network ($\lambda_{12} = -0.2$) and positively by the third network ($\lambda_{13} = 0.1$). The second network is only affected by the first network but has no within network effect. The third network is only affected by the first, and has an intermediate within network effect ($\lambda = 0.2$), relative to others.[25]

The agents in the three networks produce outcomes over $T = 500$ (or $T = 1,000$) time periods. The data generating process for the outcome function is,

$$\begin{aligned}
\mathbf{y}_{1t} &= \lambda_{11}\mathbf{W}_{11t}\mathbf{y}_{1t} + \lambda_{12}\mathbf{W}_{12t}\mathbf{y}_{2t} + \lambda_{13}\mathbf{W}_{13t}\mathbf{y}_{3t} + \mathbf{X}_{11t}\beta_{11} + \iota_{n_r}\mathbf{x}_{21t}\beta_{21} + \iota_{n_r}\alpha_{1t} + \mathbf{u}_{1t}, \\
\mathbf{y}_{2t} &= \lambda_{21}\mathbf{W}_{21t}\mathbf{y}_{1t} + \lambda_{22}\mathbf{W}_{22t}\mathbf{y}_{2t} + \lambda_{23}\mathbf{W}_{23t}\mathbf{y}_{3t} + \mathbf{X}_{12t}\beta_{12} + \iota_{n_r}\mathbf{x}_{22t}\beta_{22} + \iota_{n_r}\alpha_{2t} + \mathbf{u}_{2t}, \qquad (12) \\
\mathbf{y}_{3t} &= \lambda_{31}\mathbf{W}_{31t}\mathbf{y}_{1t} + \lambda_{32}\mathbf{W}_{32t}\mathbf{y}_{2t} + \lambda_{33}\mathbf{W}_{33t}\mathbf{y}_{3t} + \mathbf{X}_{13t}\beta_{13} + \iota_{n_r}\mathbf{x}_{23t}\beta_{23} + \iota_{n_r}\alpha_{3t} + \mathbf{u}_{3t},
\end{aligned}$$

where $\mathbf{X}_{1rt}$ and $\mathbf{x}_{2rt}$ for $r = 1, 2, 3$ are the peer-level exogenous variables and the network-level exogenous variables, respectively. The $\mathbf{X}_{1rt}$ contains two regressors drawn from a bivariate normal distribution with a variance of 1 and a correlation of 0.5. The network-level exogenous variable, $\mathbf{x}_{2rt}$, contains an intercept and two regressors, $x_{2rt}(0), x_{2rt}(1)$ and $x_{2rt}(2)$ (respectively), independently drawn from a standard normal distribution. The network-level exogenous variables will be specified in greater detail later, so as to induce a selection bias into the system. The network-level exogenous variables will be eliminated from the model with the within transformation but will be recovered in the estimation of the network managers' strategic bias in a final step. We set all elements of β_{1r} (i.e., $\beta_{1r}(1)$ and $\beta_{1r}(2)$) and β_{2r} (i.e., $\beta_{2r}(0)$, $\beta_{2r}(1)$, and $\beta_{2r}(2)$) equal to 1 for $r = 1, 2, 3$, and each element in $\mathbf{u}_{rt}$ is distributed $iidN(0, 2)$ for $r = 1, 2, 3$.

The α_{rt} for $r = 1, 2, 3$ are the strategic bias terms, and we adapt the experiment design by Schmertmann (1994) and Bourguignon et al. (2007) to model them. For simplicity, we assume the choice specific utility function for the managers of the three networks are the same (i.e., $\varphi_r = \varphi$). For each time period t, the manager of the network r has three choices on her network, $\mathbb{A} = (1, 2, 3)$. Then her decision is made by the following rule:[26] an option $s \in \mathbb{A}$ is chosen if only if $\varphi(s, \mathbf{z}_t) + e_{rts} > \varphi(s^{'}, \mathbf{z}_t) + e_{rts'}$ for $\forall s^{'} \neq s$. The $\mathbf{z}_t$ contains an intercept and two regressors, $z_t(0), z_t(1), z_t(2)$, and the two regressors

[25] We experimented with different Λ and the results were qualitatively the same.

[26] Simulating the selection equation precedes simulating the outcome function, but for the purpose of exposition we described the outcome function first.

are independently drawn from $U(0,1)$. The e_{rts} for $s = 1,2,3$ are independently drawn from a Gumbel distribution, which naturally leads to the familiar MNL specification. The Gumbel distribution DGP provides an estimation advantage for MNL. The choice-specific utility function, $\varphi(s, \mathbf{z}_t)$, is specified in two ways, linear and non-linear, such that

1. Linear: $\varphi(1, \mathbf{z}_t) = 0$; $\varphi(2, \mathbf{z}_t) = 1 + z_t(1)$; and $\varphi(3, \mathbf{z}_t) = 2 + z_t(2)$.

2. Non-linear: $\varphi(1, \mathbf{z}_t) = 0$; $\varphi(2, \mathbf{z}_t) = Beta_{5,1}(z_t(1))$; and $\varphi(3, \mathbf{z}_t) = Beta_{5,1}(z_t(2))$,

where $Beta_{a,b}(\cdot)$ is the beta density with shape parameters a and b. Here, $\varphi(1, \mathbf{z}_t)$ is normalized to zero for identification of MNL. Given the two specifications of φ above, we generate two sets of T choices for each manager r. Next we set up a correlation between $\mathbf{z}_t$ and $\mathbf{x}_{2rt}$ such that $x_{2rt}(j) = 0.9 \times z_t(j)$ plus a draw from a $N(0, (1 - 0.9^2)/12)$ random variable for $j = 1, 2$ and $r = 1, 2, 3$.[27].

Following Schmertmann (1994), let $\zeta_{rt} = (\zeta_{rts'}, \zeta_{rts''}) = (e_{rts'} - e_{rts}, e_{rts''} - e_{rts})$, where s is the selected option, and $s^{'}$ and $s^{''}$ are the non-selected options. Also, let $F_e(\zeta_{rt}) = 1 + \exp(-\zeta_{rts'}) + \exp(-\zeta_{rts''})$. We generate the selection bias in two ways:[28]

1. Lee's model (Monotonicity)

$$\alpha_{rt}(\zeta_{rt}) = \rho \left[\Phi^{-1}\left(F_e(\zeta_{rt})\right) + \frac{F_e(\zeta_{rt})}{2\phi[\Phi^{-1}\{F_e(\zeta_{rt})\}]} \right], \tag{13}$$

   and ρ is a scalar, which we discuss below.

2. Dahl model (Non-monotonicity)

$$\alpha_{rt}(\zeta_{rt}) = c_1 \Phi\left[\zeta_{rt}\boldsymbol{\rho}_2 + \sin(4\zeta_{rt}\boldsymbol{\rho}_2)\right] - c_2 \quad \text{where } \boldsymbol{\rho}_2 = (\rho_{21}, \rho_{22})^{'}. \tag{14}$$

(13) is simply the conditional mean of the residuals of the outcome equation under Lee's (1983) approach, and (14), proposed by Schmertmann (1994), is an example of a violation of the assumptions in Lee (1983). In particular, it violates monotonicity of of the selection bias. We expect Lee's approach will work well for (13), whereas Dahl's approach will work better for (14). Following Schmertmann (1994), we adjust the experimental parameters, $\rho, \boldsymbol{\rho}_2, c_1$ and c_2, so that the bias term explains roughly 25% of the variation in the outcome. Therefore, we have a total of four different DGPs for the

[27] The correlation between $\mathbf{z}_t$ and $\mathbf{x}_{2rt}$ is to make sure selection bias works for $\mathbf{x}_{2rt}$, not only for the intercept. The large correlation is to make bias apparent so estimators are easily contrasted.

[28] See Schmertmann (1994) and Bourguignon et al. (2007) for other forms of the selection bias.

selection bias: two types of choice-specific utility functions (linear and non-linear), and two types of bias generating processes (monotonic and non-monotonic). With all variables generated, the outcome variables, y_{rt} for $r = 1, 2, 3$, are generated from the reduced form of (12) for each r and t.

For estimation, we consider several different approaches. In the first step, we estimate the network effects with and without selectivity bias correction. That is, "with selectivity bias correction" means estimation of the within-transformed model in (5), and "without selectivity bias correction" means estimation of (2) ignoring α_t. In the second stage we estimate selection probabilities, using standard MNL and RF. For RF we experiment with tuning the maximum number of splits in a tree using two cross-validation methods: one based on out-of-bag classification error and another based on five-fold cross-validated error. We compared these methods to the common practice of naively setting the maximum number of splits to $T - 1$ (a "fully grown" tree or a "tree without tuning"). To save space, a discussion of these methods and complete simulation results are reported in the Online Appendix. However, we found that five-fold cross-validation worked best in terms of bias and variance of the final-step estimates, so in the main text we only report RF results based on this tuning method.

Next, using the probability estimates we formulate the selection bias term γ_r using both the Lee approach in (10) and the Dahl approach, for which we use polynomial approximations of ψ_r in (11). Then, we estimate the network-level coefficients, β_{2r}, in two ways: with and without the estimated selection bias term γ_r included in the regression of $\hat{\upsilon}_{rt}$ on $\mathbf{x}_{2rt}$. These are reported along with ρ from Lee's model in (13) when appropriate. We perform 1,000 draws for each simulation design.

3.1. Simulated Results: Peer-effects (λ_{rr}) and Competitor-effects (λ_{rk}) Estimation

We do not report coefficients results for ($\mathbf{X}_{1rt}$), but they are available upon request. Table 1 reports the empirical mean, standard deviation (SD), and root mean squared error $\langle$RMSE$\rangle$ of the network effect estimates over 1,000 simulation draws. The first column contains the true values of the effects, and the second and third columns contain estimates without selectivity bias correction for the $\{n_r, T\} = \{5, 500\}$ and $\{10, 1000\}$ designs, respectively. These estimates are contaminated by the managers' network selections and are (mostly upward) biased. For example, in the first row $\lambda_{11} = 0.4$ is estimated with a mean of 0.651 in the $\{5, 500\}$ design (column 2), and the bias worsens to a mean of 0.7598 as $\{n_r, T\}$ increases to $\{10, 1000\}$ (column 3). Columns 4-7 are for various $\{n_r, T\}$ designs, but with selectivity bias correction (i.e., the within transformation in (5)). The last two columns employ both the selectivity bias correction and a bootstrap finite sample bias correction, discussed below.

Notice that the empirical SD of the estimates with "no bias correction" is smaller than with "selection bias correction only." For $\lambda_{11} = 0.4$ in the $\{5, 500\}$ design, compare SD (0.0138) with no bias correction (column 2) to SD (0.0192) with bias correction (column 4). This occurs because the within transformation of the bias corrected estimates reduces the effective sample size. We can also see that the number of peers in a network, n_r, and the number of time periods, T, have different effects on the bias and variance of the selection bias corrected estimates (columns 4-7). First, as T increases, SD decreases, but the opposite is true for n_r. This is because, in our setup, the network becomes more dense as n_r increases due to the increased number of effective peers or competitors. This makes it harder to estimate the network effect consistently because within-variation in a network decreases. Lee (2004) analyses this phenomenon, and Lee et al. (2007) provide simulated evidence suggesting that this may be avoided if network weights are constant as network size increases. Thus, if network sparsity is maintained, increasing n_r or T will have the same effect on the variance of the estimates.[29]

Moreover, moderately sized n_r is important for reducing bias. For most cases in Table 1, when $n_r = 10$, SD is approximately equal to RMSE, implying that the empirical distributions of the estimates are well centered on the true values. However, when $n_r = 5$, there is a sizable finite sample bias across all the network effect estimates which persists as T increases. This may be a concern in our empirical application where the network size is only 5. Therefore, we use a parametric (residual-based) bootstrap to remove the bias here and in our application. From the full sample, we obtain initial QML estimates, $\hat{\Lambda}_0$, and the residuals. The residuals contain the errors as well as selectivity bias and the effects of the group level regressors. Then we generate $b = 1, ..., B$ bootstrap samples using $\hat{\Lambda}_0$ as the true parameter in the following way: we randomly sample with replacement a set of residuals for our three networks for a given t, so that we can maintain network dependency in our data over time. Then we use the reduced form of (12) to generate bootstrap outcomes y^* and compute bootstrap estimates $\hat{\Lambda}^{*(b)}$. Bootstrap bias correction is then done by $\tilde{\Lambda} = 2\hat{\Lambda}_0 - \frac{1}{B}\sum_{b=1}^{B} \hat{\Lambda}^{*(b)}$. We set $B = 300$ in our simulations and empirical application.[30] A similar bootstrap bias correction is considered in Kim and Sun (2016) for nonlinear panel data models with fixed-effects. We also use these bootstrap samples to compute the standard errors of the estimates of the network-level coefficients, β_{2r}, in the final-step estimates below.[31] For the first step, the estimated peer- and competitor-effects in the last two columns of Table 1 are well centered on their true values even when $n_r = 5$.

[29]Additional unreported simulations confirm these sparsity results.

[30]QML estimation in the first step is computationally costly, so having a large B is not practical. We examined the effect of B on the estimation results and found that B = 300 provides reasonable results without computational burden.

[31]For these estimates, the bootstrap can be seen as a non-parametric bootstrap.

3.2 Simulated Results: Network-Level Coefficients (β_{2r}) Estimation

Table 2 reports selected estimation results of the network-level coefficients. Even though they are selected results (to save space), they are not atypical. Full results are in the Online Appendix. All these results are computed from the final-step regression of $\hat{\upsilon}_{rt}$ (the residuals from the first-step) on $\mathbf{x}_{2rt}$ with or without the selection bias correction term (γ_r) as described in section 2.2.2. We report both MNL and RF results but only RF results with five-fold cross-validation tuning.[32] The first-step estimation includes bootstrap finite sample bias correction and selectivity bias correction in all cases, corresponding to the last two columns of Table 1. Table 2 presents averages over all $R = 3$ networks of the empirical bias (Bias), root mean squared error (RMSE), and bootstrap standard error coverage rate with a target rate of 0.95 (95% Coverage Rate) for estimates of the network-level coefficients (β_{2r}) over the 1,000 simulation draws for each of four DGPs described above. For Dahl's approach, estimates of the intercept and the polynomial terms are omitted from the table.[33]

Columns 2-4 under "No Bias Term" contain the regression results without the selection bias term, γ_r. Obviously, these estimates show much larger RMSE than the other estimates in the table (columns 5-16), which include the bias correction term. For example, for "DGP1" in the $\{5, 500\}$ design, compare the RMSE for $\beta_2(1)$ of 0.42 (column 3) to 0.25 (column 6). The next four columns (5-8) and the following two columns (9 and 10) under "MNL" contain results for Lee's and Dahl's approaches, respectively, when MNL is used to estimate selection probabilities. Similarly, the last six columns (11-16) under "RF" contain a corresponding set of results when RF is used for estimation of the selection probabilities. As expected, Lee's approach with MNL outperforms the others in DGP1 in terms of bias and RMSE, whereas Dahl's approach with MNL shows the best performance in DGP2. Interestingly, when we switch between DGP1 and DGP2, performance deterioration of Lee's approach is mainly due to increased bias, while it is an increased variance for Dahl's approach. For DGP3 and DGP4, when RF is used to estimate the selection probabilities there is smaller bias when compared to MNL estimation of the probabilities. In these cases, Lee's approach with MNL exhibits similar deterioration as in DGP2, but Dahl with MNL shows a somewhat different pattern: the bias increases but variance decreases in many cases. Notice that Dahl's RMSE is improved when we move from DGP1 to DGP3. Overall, if we compare the performance of the MNL estimates to the RF estimates across the various DGPs, RF appears more robust to different designs in terms of bias, RMSE and the coverage rate.

[32] Five-fold cross-validation worked best in terms of RMSE of the final-step estimates. The results based on the other tuning method and non-tuning results are in the Online Appendix.

[33] Empirical standard deviation (SD) and average bootstrap standard error estimates (Avg. Bootstrap SE) for the estimates are included in the Online Appendix.

4. Application to the 2015-16 NBA Season

4.1 Empirical model and variables

We apply our network competition model to NBA data for 30 teams over the 2015-2016 regular season. 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* for each player in each period for our outcome variable.[34] 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 network of five players at a time.[35] Following Horrace et al. (2016), we drop time periods less than 30 seconds and overtime periods.[36] This results in 83,334 time periods for the league, roughly 833,000 observations in total.

Outcome Function: Since competition in sports is head-to-head, we can consider a restricted version of the general model, where there is only one peer network and one competitor network in each time period. In this case, the outcome function for team r and k in period t of game g is

$$\begin{aligned}\mathbf{y}_{rtg} &= \lambda_{rr}\mathbf{W}_{rrtg}\mathbf{y}_{rtg} + \lambda_{rk}\mathbf{W}_{rktg}\mathbf{y}_{ktg} + \mathbf{X}_{1,rtg}\beta_{1r} + \iota_5\mathbf{x}_{2,rtg}\beta_{2r} + \iota_5\alpha_{rt} + \mathbf{u}_{rtg} \\ \mathbf{y}_{ktg} &= \lambda_{kk}\mathbf{W}_{kktg}\mathbf{y}_{ktg} + \lambda_{kr}\mathbf{W}_{krtg}\mathbf{y}_{rtg} + \mathbf{X}_{1,ktg}\beta_{1k} + \iota_5\mathbf{x}_{2,ktg}\beta_{2k} + \iota_5\alpha_{kt} + \mathbf{u}_{ktg}\end{aligned} \tag{15}$$

where $\mathbf{y}_{rtg}$ and $\mathbf{y}_{ktg}$ are the 5×1 outcome vector of team r's and team k's chosen lineup in period t of game g, respectively, and $\mathbf{W}_{rrtg}$ and $\mathbf{W}_{kktg}$ are the 5×5 zero diagonal and row-normalized matrices for the within-network interactions, and $\mathbf{W}_{rktg}$ and $\mathbf{W}_{krtg}$ are similarly defined matrices for cross-network interactions. The $\mathbf{X}_{1,rtg}$ and $\mathbf{X}_{1,ktg}$ are matrices of the player-varying exogenous variables for team r and k's lineup in period t of game g, respectively. The $\mathbf{x}_{2,rtg}$ and $\mathbf{x}_{2,ktg}$ are matrices of

[34] See, e.g., Berri (1999). *Wins produced* is 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 measurable at the individual level.

[35] 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.

[36] Horrace et al. (2016) also drop time periods where the number of player types in any active lineup is less than 2, where "player types" are Guards or Forwards. Defining heterogeneous types aids in identification of the model, as we shall see. However, we found that their restriction eliminated too many of our data, so we relaxed the restriction to "less than 1," which removed only 240 out of 112,204 time periods.

the network-level exogenous variables which, as previously noted, will be eliminated from the model with the within transformation, but the coefficients will be recovered in the estimation of the coach's strategic bias. The $\mathbf{u}_{rtg}$ and $\mathbf{u}_{ktg}$ are 5×1 error term vectors, in which each element is assumed to be $iid(0, \sigma_r^2)$ and $iid(0, \sigma_k^2)$, respectively. The likelihood function of (15) is a special case of the likelihood derived in the Online Appendix.

Selection Equation: Ideally, coaches would select from all possible five-player lineups at their disposal, and each network would be specified as complete (i.e., $w_{ij} = 1/4$ for $i \neq j$, $w_{ij} = 0$ otherwise), so each selected peer would interact with all other selected peers in the lineup in the same way. Then, the set of all actions $\mathbb{A}$ would map directly in to the network specification. However, if we specify complete networks, then (5) is not identified. Identification requires sparsity (exclusion restrictions) in the networks for Assumption 4 to be satisfied. Instead, we assume that NBA coaches employ a strategy to create favorable offensive mismatches against the opposing team. That is, coaches are interested in creating easy scoring opportunities, where opposing guards are defending his taller forwards (close to the basket) or where opposing forwards are defending his quicker guards (way from the basket).[37] Therefore, we choose the number of active guards in the current period ($Nguard_{rt}$) for the dependent variable in our selection equation. Therefore, the model is specified as the multinomial logit model (MNL) such that the action set for team r in period t is $[0, 1, ..., 5]$, representing the number of guards in the current period. Since the coach's decision rule may be highly non-linear, we also estimate the selection equation using RF with five-fold CV. Selected actions are a function of state variables $\mathbf{z}_{grt}$, which we describe below.

Network: Given our assumption that coaches are interested in creating positional mismatches, we would like to specify our network adjacency matrices to reflect this. With this is mind, we use the same-type peer-effect weight matrix considered in Horrace et al. (2016), where "types" are the player positions: *Guards* or *Forwards*, with *Forwards* including Centers. That is, the same-type weight matrix is $\mathbf{W}$, where $\mathbf{W}_0 = [w_{0,ij}]$ is an adjacency matrix with $w_{0,ij} = 1$ if the i^{th} and j^{th} players are both guards or forwards.[38] Then row-normalize $\mathbf{W}_0$ so that $\mathbf{W}\iota_5 = \iota_N$. 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 the opposing network (an exclusion restriction). This restriction is required for identification of the model, in particular to separately identify the network parameters from the other

[37] The precedent for this assumption on strategy can be found in Calvo et al. (2017) and Marmarinos et al. (2016).

[38] We realize that this is a fairly restrictive network specification. An alternative would be to create network connections based on the frequency that players pass to one another, but this is left for future research.

input parameters. We can easily see that if there are no heterogeneous types, we can not distinguish the competitor-effect from the network fixed-effect induced by a manager. For more details on identification and the use of heterogeneous type restrictions, see Horrace et al. (2016) and Horrace and Jung (2018). For example, let's assume the lineup for team r is $[F, F, G, F, G]^{'}$, and for team k it is $[G, F, F, G, G]^{'}$ in period t of game g, where $F = Forward$ and $G = Guard$. Then, the network matrices in (18) are given by:

$$\mathbf{W}_{rrtg} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \mathbf{W}_{kktg} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}, \mathbf{W}_{rktg} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \mathbf{W}_{krtg} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

Variables: We use the *Wins Produced* measure based on the work of sports economist David Berri (Berri, 1999; Berri et al. 2006) for the outcome,

$$\begin{aligned} y_{irtg} =& (0.064 \cdot 3PT_{irtg} + 0.032 \cdot 2PT_{irtg} + 0.017 \cdot FT_{irtg} + 0.034 \cdot REB_{irtg} + 0.033 \cdot STL_{irtg} \\ &+ 0.020 \cdot BLK_{irtg} - 0.034 \cdot MFG_{irtg} - 0.015 \cdot MFT_{irtg} - 0.034 \cdot TO_{irtg})/Mins_{irtg}, \end{aligned}$$

where $3PT_{irtg}$, $2PT_{irtg}$ and FT_{irtg} are 3-point field goals made, 2-points field goals made, and free throws made, respectively, REB_{irtg} is rebounds, STL_{irtg} is steals, BLK_{irtg} is blocks, MFG_{irtg} is missed field goals, MFT_{irtg} is missed free throws, TO_{irtg} is turnovers, and $Mins_{irtg}$ is minutes played by player i of team r in period t of game g. 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.[39] The player-varying exogenous input variables in the outcome equation ($\mathbf{X}_{1,rtg}$) are $Experience_{irtg}$ and $Fatigue_{irtg}$. The $Experience_{irtg}$ variable is minutes played from the start of the game to the end of period t-1, and $Fatigue_{irtg}$ is minutes continuously played until the end of period t-1. We also included player dummies to control for player-specific heterogeneity. The network-level exogenous variables ($\mathbf{x}_{2,rtg}$) are the RPI_{rtg} of the opposing team from the end of the previous season (Relative Percentage Index, a measure of the opposition's power rating from the previous season), $Home_{rtg}$, an indicator variable for a home game, and $2ndHalf_{rtg}$, an indicator equal to 1 if t is in the second half.

We use nine exogenous variables for the determinants in the selection equation ($\mathbf{z}_{rtg}$). The

[39]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 Table 1.

$CurScore\ Diff_{rtg}$ variable is the "current" score differential between the two teams in the last period ($T-1$), and $CumScoreDiff_{rtg}$ is a "cumulative" version of the score differential from the beginning of the game up to and including period $t-1$. The $CumFoul_{rtg}$ ($CurFoul_{rtg}$) variable is the cumulative (current) number of fouls committed at the end of period $t-1$ (in period $t-1$). The $CurTime_{rtg}$ variable is the game time at the start of period t; the $Duration_{rtg}$ variable is the duration of period $t-1$. Because we are using the number of guards in the game as our selection equation outcome, we include predetermined measures of the number of guards as determinants of selection. The $NguardR_{rtg}$ variable is the Number of *guard*s available to the coach (on the Roster) at the beginning of the game.[40] The $NguardOPP_{rtg}$ variable is the Number of *guard*s that the OPPosing coach had in the game last period. We also include the one period lag of the dependent variable ($Nguard_{r(t-1)g}$) in the selection equation. Descriptive statistics for the variables in the outcome and selection equations are in the Online Appendix. In the next section we report the first-step structural estimates of team chemistry and team rivalries for the 2015-16 NBA regular season. All other estimates are reported and discussed in the Online Appendix.

4.3 Results

Table 3 contains the structural estimates of team chemistry (within-team peer-effects) and team rivalry (cross-team competitor-effects) from first-stage QML estimation of (5). The table contains a subset of the estimated structural parameters *by division*. That is, we do not report all competitor-effects, λ_{rk} and λ_{kr}, for all match-ups for the 30 teams. Instead, to save space we report competitor-effects for within-division rivals for each of six divisions in the table (e.g., *Atlantic Division* of the *East Conference*). These divisions (based on geography) tend to be the most competitive rivalries, and divisional teams play the most head-to-head games over the season (four games). This choice reduces the number of reported competitor-effect estimates from $30\times29 = 870$ to a manageable $6\times(5\times4) = 120$ with relatively smaller standard errors.[41] Estimates in the table incorporate bootstrap finite sample correction, and standard errors are based on the asymptotic distribution[42]

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 3 contains the peer-effect and

[40] It only varies across games, and does not change within a game, even when a guard is injured.
[41] The interaction map for the entire league is available from the authors by request.
[42] We use the asymptotic results of Lee et al. (2010) with appropriate modifications.

competitor-effect estimates for 30 NBA teams in the 2015-16 season. 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 Table 3 where Milwaukee (MIL) had the largest positive peer-effect (λ_{rr}) of 0.044. That is, when the team's average "wins produced" increases, the team's performance is enhanced by virtue of its good chemistry, conditional on coaching strategy and other environmental and performance variables. The estimates are structural parameters and not the reduced-form effects.[43] Reduced-form effects by division are presented in the Online Appendix.

In Table 3 about one third (eleven) of the teams exhibit positive and statistically significant (5% error rate) estimated peer-effects (λ_{rr}). After Milwaukee, Detroit (DET) has the second largest positive and significant peer-effect (0.036). The smallest positive and significant peer-effect is Houston (HOU) at 0.023, and only slightly better were Phoenix (PHX) and Portland (POR), tied at 0.0026. Of the eleven teams with positive and significant peer-effects, seven were in the *East Conference* and only four in the *West Conference*. The division with the most positive and significant peer-effects (three) is the *Central Division*: Cleveland (CLE), Detroit (DET) and Miluakee (MIL). Only two teams exhibit negative and significant peer-effects, Philadelphia (PHI -0.028) and Miami (MIA, -0.027), both in the *East Conference*. Finally, the bootstrap finite sample correction tended to reduce the magnitude of the peer-effects (team chemistries), while its effect on the competitor-effects was ambiguous.

Turing to the competitor-effects (λ_{rk}), the first row of Table 3 contains results for the Toronto Raptors (TOR). Their results are read as follows. When Toronto plays Boston (BOS), the competitor-effect of Boston on Toronto's performance is -0.272. That is, when these two teams meet, Boston's team chemistry decreases the *Wins Produced* performance of individual Toronto players on average. Looking across the row, we see that the Knicks (NYK) increase Toronto's performance (0.557). Finally, Brooklyn's (BKN, -0.159) and Philadelphia's (PHI, -0.041) team chemistries diminish the performance of Toronto. However, only the Knicks' competitor-effect is statistically significant. Looking at the second row, we see that Toronto enhances the play of Boston 0.280. The largest competitor effect is Atlanta's (ATL) effect on Charlotte (CHA) at -0.726, and the second largest is that of Cleveland (CLE) on Milwaulkee (MIL) at -0.707. Not surprisingly, when team A enhances the individual play of team B, team B almost always makes team A play worse. This is reflected in the signs changing around

[43] The same holds for the coefficients on the exogenous variables: they do not reflect the marginal effects of the exogenous variables under consideration, but they give us insight into the relative magnitudes of the marginal effects within and across teams.

the diagonal in each divisional matrix. Of course these are estimates of the structural parameters. See the Online appendix for presentation and discussion of reduced-form estimates of the peer- and competitor-effects, and other coefficient estimates from the outcome function in (2).

5. Conclusions

Adapting and extending the spatial auto-regression (SAR) model, we develop a network competition model, allowing estimation of both within-network effects (peer-effects) and cross-network interaction effects (competitor-effects). The estimates provide a more complete picture of market interactions, which may be useful in understanding and predicting how exogenous shocks to a single network translate to an entire market. We apply our network competition model to 30 teams in the 2015-16 NBA regular season and find evidence of mostly positive peer-effects (team chemistry). We find both positive and negative competitor-effects (team rivalries). That is, teams with good network play may enhance or diminish the performance of opposing teams. The model is somewhat restrictive relative to other simultaneous SAR models in the literature, so future research should explore generalization of the basic model. For example, inclusion of exogenous network effects (e.g., $\mathbf{W}_{rrt}\mathbf{X}_{1t}$) in the outcome function should be considered, as should time and/or cross-sectional dependence in its error, $\mathbf{u}_{rt}$ (relaxation of Assumption 1). It may also prove fruitful to explore formulation of managerial selection bias that varies across peers within a network (relaxation of Assumption 5). For our NBA application it may be interesting to develop peer-effect weighting schemes based on passing and ball sharing.

References

[1] Andrews, D. W. K. 1991. Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models. Econometrica, 59, 307–345.

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

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

[4] 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

[5] 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

[6] 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.

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

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

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

[10] Bourguignon, F., Fournier, M., and Gurgand, M. 2007. Selection bias corrections based on the multinomial logit model: Monte carlo comparisons. J. Econ. Surv. 21, 174–205.

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

[12] Breiman, L. 2001. Random Forests. Machine Learning, 45, 5–32.

[13] Brock, W.A., Durlauf, S.N., 2001. Interactions-based models. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, Vol. 5. North-Holland, 3297–3380.

[14] Brock, W.A., Durlauf, S.N., 2002. A multinomial choice model with neighborhood effects. Amer. Econ. Rev. 92, 298–303.

[15] 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, 175–206.

[16] Calvo, J.L., Garcia, M.A., Navandar, A. (2017). Analysis of mismatch after ball screens in Spanish professional basketball. International Journal of Performance Analysis in Sport, 17, 555-562.

[17] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018), Double/debiased Machine Learning for Treatment and Structural Parameters. The Econometrics Journal, 21, C1-C68.

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

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

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

[21] Craven, P. and Wahba, G. Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation. Numerische Mathematik, January 1979, 31, 377-403.

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

[23] 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.

[24] 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

[25] 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.

[26] Heckman, J.J. 1979. Sample selection bias as a specification error. Econometrica, 47, 153-161.

[27] Horrace,W.C., Jung, H., 2018. Stochastic frontier models with network selectivity, J. Productivity Analysis. 50: 101-116

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

[29] Hoshino,T., 2019. Two-step estimation of incomplete information social interaction models with sample selection, J. of Bus. and Econ. Stat., 37(4), 598-612.

[30] 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.

[31] James, G., Witten, D., Hastie, T. and Tibshirani, R. 2017. An introduction to statistical learning. Springer.

[32] Jeanty, 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. 343–352

[33] 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.

[34] 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

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

[36] Kim, M. and Sun, Y. 2016. Bootstrap and k-Step Bootstrap Bias Corrections for the Fixed Effects Estimator in Nonlinear Panel Data Models. Econometric Theory, 32(6), 1523-1568.

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

[38] Lee, B., Lessler, J., and Stuart, E., 2010. Improving propensity score weighting using machine learning. Stat Med. 2010 February 10; 29(3): 337-346.

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

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

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

[42] 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 .

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

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

[45] 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.

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

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

[48] Lee, L.F. and Yu, J., 2010. Estimation of spatial autoregressive panel data models with fixed effects. J. Econometrics, 154(2) 165-185.

[49] 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.

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

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

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

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

[54] Marmarinos, C., Apostolidis, N., Kostopoulos, N., Apostolidis, A. (2016). Efficacy of the "Pick and Roll" offense in top level European basketball teams. Journal of Human Kinetics, 50, 121–129.

[55] 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.

[56] Mentch, L., and Hooker, G. 2016. Quantifying Uncertainty in Random Forests via Confidence Intervals and Hypothesis Tests. Journal of Machine Learning Research, 17, 1-41.

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

[58] Murphy, K. M., and R. H. Topel. 1985. Estimation and Inference in Two-Step Econometric Models. Journal & Business and Economic Statistics, 3, 370–379.

[59] Newey, W. K. 1994. The Asymptotic Variance of Semiparametric Estimators. Econometrica, 62, 1349–1382.

[60] Newey, W. K. 1997. Convergence Rates and Asymptotic Normality for Series Estimators. J. Econometrics, 79, 147–168.

[61] Newey, W. K. 2009. Two Step Series Estimation of Sample Selection Models. Econometrics Journal. V12. 217-229.

[62] Newey, W. K. and McFadden, D. L. 1994. Handbook of Econometrics, Volume IV, Chapter 36.

[63] Newey, W. K., J. L. Powell and J. R. Walker. 1990. Semiparametric estimation of selection models: Some empirical results. AEReview P&P 80, 324–28.

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

[65] Olley, G. S., Pakes, A. 1996. The dynamics of productivity in the telecommunications equipment industry," Econometrica, 64(6), 1263–97.

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

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

[68] Schmertmann, C. P. 1994. Selectivity bias correction methods in polychotomous sample selection models. J. Econometrics 60: 101–132.

[69] 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.

[70] Scornet, E., Biau, G., and Vert, J. P. 2015. Consistency of Random Forests. The Annals of Statistics, 43, 1716–1741.

[71] 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.

[72] Wager, S. and Athey, S. 2018 Estimation and Inference of Heterogeneous Treatment Effects using Random Forests, Journal of the American Statistical Association, 113:523, 1228-1242,

[73] Wager, S. and Walther, G. 2015. Adaptive Concentration of Regression Trees with Application to Random Forests. arXiv:1503.06388.

[74] Wang, H. and Li, R. 2007. Tuning Parameter Selectors for the Smoothly Clipped Absolute Deviation Method. Biometrika, V. 94, No. 3, pp. 553-568

Tables

Table 1: Simulated Estimation of Network Peer-Effects & Competitor-Effects

			Selection Bias Corrected Estimates					
	No Bias Corrections		Selection Bias Correction Only				Bootstrap Correction	
$\{n_r, T\}$	{5,500}	{10,1000}	{5,500}	{5,1000}	{10,500}	{10,1000}	{5,500}	{5,1000}
	0.6451	0.7598	0.4178	0.4176	0.4011	0.4005	0.4002	0.4007
$\lambda_{11} = 0.4$	(0.0138)	(0.0075)	(0.0192)	(0.0140)	(0.0215)	(0.0159)	(0.0203)	(0.0139)
	⟨0.2455⟩	⟨0.3598⟩	⟨0.0262⟩	⟨0.0225⟩	⟨0.0216⟩	⟨0.0159⟩	⟨0.0203⟩	⟨0.0139⟩
	-0.1113	-0.0719	-0.1945	-0.1942	-0.2001	-0.2011	-0.2003	-0.1990
$\lambda_{12} = -0.2$	(0.0189)	(0.0099)	(0.0275)	(0.0189)	(0.0285)	(0.0193)	(0.0275)	(0.0184)
	⟨0.0907⟩	⟨0.1285⟩	⟨0.0280⟩	⟨0.0198⟩	⟨0.0285⟩	⟨0.0194⟩	⟨0.0275⟩	⟨0.0184⟩
	0.0541	0.0357	0.0936	0.0941	0.0996	0.1006	0.0998	0.0978
$\lambda_{13} = 0.1$	(0.0159)	(0.0084)	(0.0258)	(0.0188)	(0.0276)	(0.0192)	(0.0260)	(0.0184)
	⟨0.0486⟩	⟨0.0649⟩	⟨0.0266⟩	⟨0.0197⟩	⟨0.0276⟩	⟨0.0192⟩	⟨0.0260⟩	⟨0.0186⟩
	0.1274	0.0862	0.1914	0.1925	0.1997	0.1996	0.2005	0.2014
$\lambda_{21} = 0.2$	(0.0148)	(0.0079)	(0.0258)	(0.0183)	(0.0275)	(0.0198)	(0.0261)	(0.0196)
	⟨0.0740⟩	⟨0.1141⟩	⟨0.0272⟩	⟨0.0198⟩	⟨0.0275⟩	⟨0.0198⟩	⟨0.0261⟩	⟨0.0197⟩
	0.3276	0.5165	0.0126	0.0123	0.0004	0.0014	0.0000	-0.0002
$\lambda_{22} = 0.0$	(0.0187)	(0.0138)	(0.0214)	(0.0149)	(0.0205)	(0.0147)	(0.0224)	(0.0151)
	⟨0.3282⟩	⟨0.5166⟩	⟨0.0249⟩	⟨0.0193⟩	⟨0.0205⟩	⟨0.0147⟩	⟨0.0224⟩	⟨0.0151⟩
	0.0032	0.0024	-0.0119	-0.0121	-0.0152	-0.0161	0.0010	0.0003
$\lambda_{23} = 0.0$	(0.0181)	(0.0093)	(0.0268)	(0.0194)	(0.0299)	(0.0209)	(0.0279)	(0.0195)
	⟨0.0184⟩	⟨0.0096⟩	⟨0.0293⟩	⟨0.0229⟩	⟨0.0336⟩	⟨0.0264⟩	⟨0.0280⟩	⟨0.0195⟩
	-0.0595	-0.0392	-0.0958	-0.0954	-0.1021	-0.1031	-0.1005	-0.0977
$\lambda_{31} = -0.1$	(0.0137)	(0.0074)	(0.0257)	(0.0179)	(0.0283)	(0.0203)	(0.0236)	(0.0172)
	⟨0.0428⟩	⟨0.0612⟩	⟨0.0260⟩	⟨0.0185⟩	⟨0.0284⟩	⟨0.0206⟩	⟨0.0236⟩	⟨0.0173⟩
	0.0016	0.0013	0.0058	0.0045	0.0064	0.0051	-0.0025	0.0001
$\lambda_{32} = 0.0$	(0.0200)	(0.0106)	(0.0274)	(0.0189)	(0.0276)	(0.0194)	(0.0266)	(0.0190)
	⟨0.0200⟩	⟨0.0107⟩	⟨0.0280⟩	⟨0.0195⟩	⟨0.0283⟩	⟨0.0200⟩	⟨0.0267⟩	⟨0.0190⟩
	0.4987	0.6526	0.2151	0.2171	0.2012	0.2004	0.1986	0.1984
$\lambda_{33} = 0.2$	(0.0163)	(0.0114)	(0.0209)	(0.0148)	(0.0219)	(0.0159)	(0.0204)	(0.0150)
	⟨0.2991⟩	⟨0.4527⟩	⟨0.0258⟩	⟨0.0226 ⟩	⟨0.0219⟩	⟨0.0159⟩	⟨0.0204⟩	⟨0.0150⟩

For each parameter λ_{rk} we report the mean, the (Standard Deviation), and the ⟨RMSE⟩ over 1,000 simulated draws. Two types of bias correction employed: Selection Bias Correction (columns 4-9) and Bootstrap Finite Sample Bias Correction (columns 8 and 9)

Table 2: Simulated Estimation of Network-Level Coefficients

		No Bias Term γ_r			MNL						RF + Five-Fold Cross-Validation					
					Lee				Dahl		Lee				Dahl	
$\{n_r, T\}$		$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	ρ	$\beta_2(1)$	$\beta_2(2)$	$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	ρ	$\beta_2(1)$	$\beta_2(2)$
DGP1: Lee + Linear Utility Function																
	Bias	-1.40	-0.32	0.51	-0.01	-0.01	0.00	0.01	-0.10	0.12	-0.08	-0.06	0.07	0.04	-0.14	0.18
$\{5, 500\}$	RMSE	1.42	0.42	0.59	0.23	0.25	0.25	0.17	0.57	0.58	0.26	0.27	0.27	0.19	0.35	0.38
	95% Coverage Rate	0.00	0.78	0.54	0.95	0.95	0.95	0.97	0.99	0.99	0.93	0.95	0.95	0.94	0.96	0.96
	Bias	-1.40	-0.34	0.50	-0.01	0.00	0.00	0.00	-0.07	0.09	-0.06	-0.04	0.06	0.03	-0.14	0.17
$\{5, 1000\}$	RMSE	1.41	0.39	0.55	0.16	0.18	0.18	0.11	0.37	0.35	0.19	0.20	0.20	0.14	0.28	0.30
	95% Coverage Rate	0.00	0.59	0.26	0.95	0.94	0.94	0.96	0.97	0.98	0.93	0.95	0.95	0.94	0.94	0.94
DGP2: Dahl + Linear Utility Function																
	Bias	-1.48	-0.45	0.67	-0.41	-0.20	0.30	0.40	-0.03	0.12	-0.47	-0.23	0.35	0.42	-0.17	0.30
$\{5, 500\}$	RMSE	1.49	0.53	0.74	0.47	0.35	0.41	0.42	0.42	0.36	0.53	0.37	0.45	0.45	0.37	0.44
	95% Coverage Rate	0.00	0.67	0.36	0.57	0.87	0.80	0.21	0.99	0.99	0.50	0.85	0.76	0.28	0.94	0.91
	Bias	-1.47	-0.46	0.67	-0.42	-0.20	0.30	0.40	0.00	0.07	-0.46	-0.23	0.34	0.42	-0.17	0.27
$\{5, 1000\}$	RMSE	1.48	0.51	0.71	0.45	0.28	0.35	0.41	0.25	0.25	0.49	0.30	0.39	0.43	0.30	0.37
	95% Coverage Rate	0.00	0.36	0.10	0.28	0.81	0.67	0.03	0.99	0.99	0.24	0.78	0.63	0.05	0.92	0.86
DGP3: Lee + Nonlinear Utility Function																
	Bias	-2.07	0.77	0.75	-0.16	0.17	0.18	0.06	0.16	0.17	-0.02	-0.05	-0.03	-0.03	-0.11	-0.09
$\{5, 500\}$	RMSE	2.08	0.82	0.81	0.31	0.32	0.33	0.19	0.33	0.34	0.31	0.30	0.29	0.23	0.34	0.33
	95% Coverage Rate	0.00	0.25	0.30	0.90	0.91	0.90	0.93	0.92	0.92	0.94	0.94	0.95	0.95	0.95	0.96
	Bias	-2.07	0.76	0.76	-0.15	0.19	0.18	0.04	0.21	0.20	0.02	-0.05	-0.06	-0.04	-0.08	-0.09
$\{5, 1000\}$	RMSE	2.08	0.79	0.79	0.25	0.27	0.27	0.14	0.30	0.29	0.21	0.22	0.22	0.16	0.26	0.26
	95% Coverage Rate	0.00	0.06	0.06	0.87	0.84	0.85	0.93	0.84	0.85	0.95	0.94	0.94	0.95	0.93	0.92
DGP4: Dahl + Nonlinear Utility Function																
	Bias	-2.50	1.05	1.06	-0.83	0.57	0.56	0.27	0.49	0.49	-0.69	0.37	0.36	0.17	0.16	0.16
$\{5, 500\}$	RMSE	2.50	1.10	1.10	0.87	0.63	0.62	0.31	0.57	0.58	0.74	0.47	0.47	0.27	0.37	0.38
	95% Coverage Rate	0.00	0.06	0.06	0.10	0.46	0.49	0.62	0.65	0.64	0.31	0.76	0.76	0.83	0.95	0.94
	Bias	-2.50	1.06	1.07	-0.81	0.55	0.56	0.25	0.48	0.49	-0.64	0.34	0.35	0.16	0.10	0.11
$\{5, 1000\}$	RMSE	2.50	1.08	1.09	0.83	0.59	0.59	0.27	0.52	0.53	0.67	0.40	0.41	0.21	0.27	0.26
	95% Coverage Rate	0.00	0.01	0.00	0.01	0.20	0.18	0.45	0.39	0.36	0.10	0.61	0.60	0.75	0.94	0.94

All these results are based on the first-step estimates which include selection bias correction and bootstrap finite sample bias correction.
95% Coverage Rate means bootstrap standard error coverage rate with a target rate 0.95.

Table 3: Network Peer-Effect (λ_{rr}) and Competitor-Effect (λ_{rk}) Estimates within Division

EAST CONFERENCE

Atlantic Division						**t-statistics**				
	TOR	BOS	NYK	BKN	PHI	TOR	BOS	NYK	BKN	PHI
TOR	0.013	-0.229	0.557	-0.159	-0.041	1.107	-1.361	3.026	-0.928	-0.270
BOS	0.280	0.013	-0.652	0.569	-0.524	1.604	1.110	-4.785	5.405	-3.336
NYK	-0.254	0.490	0.031	-0.330	0.200	-1.770	4.258	2.570	-2.074	1.319
BKN	-0.007	-0.524	0.368	0.034	-0.331	-0.049	-4.858	2.400	2.626	-2.822
PHI	0.060	0.245	-0.189	0.443	-0.028	0.373	2.005	-1.259	3.533	-2.468
Central Division						**t-statistics**				
	CLE	IND	DET	CHI	MIL	CLE	IND	DET	CHI	MIL
CLE	0.029	0.272	-0.480	-0.383	0.814	2.476	1.427	-3.436	-3.814	5.801
IND	-0.346	-0.018	0.332	-0.250	0.129	-1.575	-1.527	2.228	-1.175	0.965
DET	0.498	-0.523	0.036	-0.369	-0.761	3.340	-2.794	2.739	-1.557	-5.295
CHI	0.532	0.207	0.182	0.017	-0.569	3.743	0.818	0.991	1.144	-2.937
MIL	-0.707	-0.234	0.635	0.420	0.044	-5.763	-1.482	4.759	2.512	3.306
Southeast Division						**t-statistics**				
	MIA	ATL	CHA	WAS	ORL	MIA	ATL	CHA	WAS	ORL
MIA	-0.027	-0.016	0.099	-0.201	-0.092	-2.265	-0.087	0.695	-1.483	-0.577
ATL	-0.099	-0.009	0.417	0.576	-0.296	-0.517	-0.761	3.188	4.642	-1.715
CHA	-0.062	-0.726	0.028	-0.034	0.832	-0.333	-4.191	2.206	-0.188	5.429
WAS	0.151	-0.387	0.057	0.007	-0.183	1.059	-3.405	0.338	0.606	-1.014
ORL	0.434	0.355	-0.563	0.266	0.029	2.069	2.165	-4.599	1.187	2.061

WEST CONFERENCE

Northwest Division						**t-statistics**				
	OKC	POR	UTA	DEN	MIN	OKC	POR	UTA	DEN	MIN
OKC	0.013	0.066	0.381	-0.452	0.122	1.135	0.461	2.442	-2.445	0.752
POR	0.034	0.024	0.245	-0.240	-0.385	0.262	2.011	2.087	-1.425	-2.264
UTA	-0.305	-0.184	0.014	0.591	0.018	-2.082	-1.213	1.155	4.367	0.111
DEN	0.160	0.137	-0.542	-0.018	-0.185	1.309	0.800	-4.831	-1.482	-0.853
MIN	-0.139	0.359	0.175	0.073	0.013	-0.705	1.979	1.096	0.302	0.979
Pacific Division						**t-statistics**				
	GSW	LAC	SAC	PHX	LAL	GSW	LAC	SAC	PHX	LAL
GSW	-0.011	0.026	-0.192	-0.587	-0.294	-0.952	0.099	-0.999	-3.083	-1.197
LAC	0.036	-0.013	0.159	0.064	-0.226	0.138	-1.156	0.805	0.465	-0.863
SAC	-0.006	-0.214	0.026	-0.499	0.423	-0.035	-0.854	2.029	-3.886	2.761
PHX	0.347	-0.109	0.509	0.024	0.575	2.251	-0.803	3.850	2.007	2.214
LAL	0.327	0.220	-0.296	-0.359	0.007	1.444	0.845	-2.262	-1.802	0.534
Southwest Division						**t-statistics**				
	SAS	DAL	MEM	HOU	NOP	SAS	DAL	MEM	HOU	NOP
SAS	0.013	0.331	0.064	-0.134	0.416	0.939	2.517	0.277	-0.904	2.295
DAL	-0.424	-0.018	0.115	0.444	-0.425	-3.371	-1.566	0.636	3.647	-3.118
MEM	-0.037	-0.055	0.025	0.029	0.069	-0.148	-0.294	1.709	0.122	0.395
HOU	0.266	-0.349	0.152	0.023	0.445	1.967	-2.929	0.467	2.009	3.139
NOP	-0.284	0.331	0.016	-0.447	-0.015	-1.884	2.632	0.111	-3.617	-1.300

This is a subset of results from estimation of the complete model.
The interaction map for the entire league is available from the authors by request.
Estimates include finite sample bootstrap bias correction.
Asymptotic standard errors and t-statistics.

A Sufficient Conditions for Identification of Λ_0 and $\beta_{1,0}$

Proposition 1. *Let $\mathbf{\Lambda}_0$ and $\beta_{1,0}$ be the true parameter values. Let $\mathbf{\Xi}_r$ be a $n_r \times N$ matrix consisting of R horizontally concatenated blocks of size n_r, with I_{n_r} in the r^{th} position and zeroes in the other $R-1$ positions. Hence, $\mathbf{\Xi}_1 = (\mathbf{I}_{n_1}, \mathbf{0})$, $\mathbf{\Xi}_R = (\mathbf{0}, \mathbf{I}_{n_R})$, and $\mathbf{\Xi}_r = (\mathbf{0}, \mathbf{I}_{n_r}, \mathbf{0})$ for $r \neq 1, R$, where the $\mathbf{0}$ matrices are appropriately conformable. Suppressing t, let $\mathbf{\Theta}_{rk} = \mathbf{\Xi}_r \mathbf{G}_{rk} \mathbf{S}_0^{-1} \mathbf{X}_1 \beta_{1,0}$ and $\mathbf{S}_0 = \mathbf{I}_N - \sum_{r=1}^R \sum_{k=1}^R \lambda_{rk0} \mathbf{G}_{rk}$. Then, the true parameters $\mathbf{\Lambda}_0$ and $\beta_{1,0}$ for the system of equation 3 (in the main text) can be identified if the R matrices $[\mathbf{Q}_r \mathbf{X}_{1r}, \mathbf{Q}_r \mathbf{\Theta}_{r1}, ..., \mathbf{Q}_r \mathbf{\Theta}_{rR}]$, $r = 1, ..., R$ have full column rank.*

Proof. *As $\mathbf{S}_0^{-1} = \mathbf{I}_N + \sum_{r=1}^R \sum_{k=1}^R \lambda_{rk0} \mathbf{G}_{rk} \mathbf{S}_0^{-1}$ and $\mathbf{J}_Q \mathbf{y} = \mathbf{J}_Q \mathbf{S}_0^{-1} \mathbf{X}_1 \beta_{1,0} + \mathbf{J}_Q \mathbf{S}_0^{-1} \mathbf{u}^*$,*[44] *equation 3 (in the main text) evaluated at the true parameters is written as*

$$\mathbf{J}_Q \mathbf{y} = \mathbf{J}_Q \mathbf{X}_1 \beta_{1,0} + \mathbf{J}_Q \sum_{r=1}^R \sum_{k=1}^R \lambda_{rk0} \mathbf{G}_{rk} \mathbf{S}_0^{-1} \mathbf{X}_1 \beta_{1,0} + \mathbf{J}_Q \mathbf{S}_0^{-1} \mathbf{u}^* \tag{16}$$

Then, $\mathbf{\Lambda}_0$ and $\beta_{1,0}$ in (16) *will be identified as long as the R matrices $[\mathbf{Q}_r \mathbf{X}_{1r}, \mathbf{Q}_r \mathbf{\Theta}_{r1}, ..., \mathbf{Q}_r \mathbf{\Theta}_{rR}]$, $r = 1, ..., R$ have full column rank. This condition will be generally satisfied because we have multiple sets of network matrices and exogenous regressors for each group, which produces enough variation to identify $\mathbf{\Lambda}_0$ and $\beta_{1,0}$.* ∎

B The Quasi-Maximum Likelihood Function

Suppressing t, the likelihood function for equation 5 (in the main text) is [45]

$$\ln L(\mathbf{\Lambda}, \beta_1, \mathbf{\Sigma}) = -\sum_{r=1}^R \frac{n_r - 1}{2} \ln(2\pi\sigma_r^2) + \ln |\bar{\mathbf{S}}(\mathbf{\Lambda})| - \sum_{r=1}^R \frac{\bar{\epsilon}_r(\theta_r)' \bar{\epsilon}_r(\theta_r)}{2\sigma_r^2}, \tag{17}$$

where $\mathbf{\Lambda} = (\mathbf{\Lambda}_1', ..., \mathbf{\Lambda}_R')'$ with $\mathbf{\Lambda}_r = (\lambda_{r1}, ..., \lambda_{rR})$, $\mathbf{\Sigma} = Diag(\sigma_1^2, ..., \sigma_R^2)$, $\bar{\mathbf{S}}(\mathbf{\Lambda}) = \mathbf{I}_N - \sum_{r=1}^R \sum_{k=1}^R \lambda_{rk} \bar{\mathbf{G}}_{rk}$, and $\bar{\epsilon}_r(\theta_r) = \bar{\mathbf{y}}_r - \sum_{k=1}^R \lambda_{rk} \bar{\mathbf{W}}_{rk} \bar{\mathbf{y}}_k - \bar{\mathbf{X}}_{1r} \beta_{1r}$ where $\theta_r = (\mathbf{\Lambda}_r, \beta_{1r})$ and $\bar{\epsilon}_r$ is a vector function of θ_r, and $\bar{\mathbf{W}}_{rk}$ and $\bar{\mathbf{X}}_{1r}$ are defined similarly. From Lemma 1 below, we show $\ln |\bar{\mathbf{S}}(\mathbf{\Lambda})| = -\ln f(\mathbf{\Lambda}) + \ln |\mathbf{S}(\mathbf{\Lambda})|$

[44] Note that $\mathbf{J}_Q \mathbf{G}_{rk} \mathbf{J}_Q = \mathbf{J}_Q \mathbf{G}_{rk}$ $\forall$ r,k which leads to $\mathbf{J}_Q \mathbf{S}_0^{-1} \mathbf{J}_Q = \mathbf{J}_Q \mathbf{S}_0^{-1}$

[45] The likelihood is conditional on the sufficient statistic, the mean of $\mathbf{y}_r$. Lee (2007a).

where $\mathbf{S}(\mathbf{\Lambda}) = \mathbf{I}_N - \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rk}$ and f is some scalar function of $\mathbf{\Lambda}$. For example, when $R = 2$, $f(\mathbf{\Lambda}) = (1-\lambda_{11})(1-\lambda_{22}) - \lambda_{12}\lambda_{21}$, and when $R = 3$, $f(\mathbf{\Lambda}) = (1-\lambda_{11})(1-\lambda_{22})(1-\lambda_{33}) - (1-\lambda_{11})\lambda_{23}\lambda_{32} - (1-\lambda_{22})\lambda_{13}\lambda_{31} - (1-\lambda_{33})\lambda_{21}\lambda_{12} - \lambda_{13}\lambda_{21}\lambda_{32}$. Using this result, we can evaluate the likelihood without $\mathbf{P}_r$ as

$$\ln L(\mathbf{\Lambda}, \beta_1, \mathbf{\Sigma}) = -\sum_{r=1}^{R}\frac{n_r - 1}{2}\ln(2\pi\sigma_r^2) - \ln f(\mathbf{\Lambda}) + \ln|\mathbf{S}(\mathbf{\Lambda})| - \sum_{r=1}^{R}\frac{\epsilon_r^{'}(\theta_r)Q_r\epsilon_r(\theta_r)}{2\sigma_r^2}, \tag{18}$$

where $\epsilon_r(\theta_r) = \mathbf{y}_r - \sum_{k=1}^{R}\lambda_{rk}\mathbf{W}_{rk}y_k - \mathbf{X}_{1r}\beta_{1r}$. There are two things to note here. First, we may need to further restrict the parameter space of $\mathbf{\Lambda}$ to guarantee that $f(\mathbf{\Lambda})$ is strictly positive. Second, it may be difficult to evaluate $|\mathbf{S}(\mathbf{\Lambda})|$. The Ord (1975) eigenvalue device may be used to compute the determinant. However, it may only work when the number of networks R is small, and all the network matrices are sparse. If the number of networks R is large, then GMM may be preferred to QML, as it avoids the computational difficulties of evaluating the determinant of $\mathbf{S}$.

To simplify estimation, we concentrate out β_1 and $\mathbf{\Sigma}$ in (18). The QML estimate of β_{1r} and σ_r^2, given $\mathbf{\Lambda}_r$ is $\hat{\beta}_{1r}(\mathbf{\Lambda}_r) = (\mathbf{X}_{1r}^{'}\mathbf{Q}_r\mathbf{X}_{1r})^{-1}\mathbf{X}_{1r}^{'}\mathbf{Q}_r\mu_r(\mathbf{\Lambda}_r)$ where $\mu_r(\mathbf{\Lambda}_r) = \mathbf{y}_r - \sum_{k=1}^{R}\lambda_{rk}\mathbf{W}_{rk}\mathbf{y}_k$ is a vector function, and

$$\hat{\sigma_r}^2(\mathbf{\Lambda}_r) = \frac{\epsilon_r^{'}(\theta_r)\mathbf{Q}_r\epsilon_r(\theta_r)}{n_r - 1} = \frac{\mu_r(\mathbf{\Lambda}_r)^{'}[\mathbf{Q}_r - \mathbf{Q}_r\mathbf{X}_{1r}(\mathbf{X}_{1r}^{'}\mathbf{Q}_r\mathbf{X}_{1r})^{-1}\mathbf{X}_{1r}^{'}\mathbf{Q}_r]\mu_r(\mathbf{\Lambda}_r)}{n_r - 1}, \tag{19}$$

is a scalar function. Then the concentrated log-likelihood function in $\mathbf{\Lambda}$ is

$$\ln L^c(\mathbf{\Lambda}) = -\sum_{r=1}^{R}\frac{n_r - 1}{2}[\ln(2\pi) + 1] - \ln f(\mathbf{\Lambda}) + \ln|\mathbf{S}(\mathbf{\Lambda})| - \sum_{r=1}^{R}\frac{n_r - 1}{2}\ln\hat{\sigma_r}^2(\mathbf{\Lambda}_r). \tag{20}$$

Then the QMLE, $\hat{\mathbf{\Lambda}}$, is the maximizer of the likelihood, and the QMLE of β_1 and $\mathbf{\Sigma}$ are $\hat{\beta}_{1r}(\hat{\mathbf{\Lambda}}_r)$ and $Diag(\hat{\sigma}_1^2(\hat{\mathbf{\Lambda}}_1), ..., \hat{\sigma}_R^2(\hat{\mathbf{\Lambda}}_R))$, respectively. The asymptotic distribution for these estimators can be derived from Lee et al. (2010, Appendix B) with appropriate modifications. When n_r is small, the simulation study in section 3 (of the main text) uncovers a sizable finite sample bias for the QML estimator of $\mathbf{\Lambda}$, which persists even as T increases. To remove the bias, we propose a residual bootstrap finite sample correction in section 3 (of the main text).

C Lemma 1

This generalizes a similar result in Lemma C.1 of Lee at al. (2010) to our setting with multiple networks and heterogeneous network effects.

Lemma 1. *Suppressing t, let the orthonormal matrix of $\mathbf{Q}_r$ be $[\mathbf{P}_r, \iota_{n_r}/\sqrt{n_r}]$. The columns in $\mathbf{P}_r$ are eigenvectors of $\mathbf{Q}_r$ corresponding to the eigenvalue one, such that $\mathbf{P}_r^{'}\iota_{n_r} = \mathbf{0}$, $\mathbf{P}_r^{'}\mathbf{P}_r = \mathbf{I}_{n_r-1}$ and $\mathbf{P}_r\mathbf{P}_r^{'} = \mathbf{Q}_r$. Denote $\mathbf{J}_P^{'} = Diag(\mathbf{P}_1^{'}, ..., \mathbf{P}_R^{'})$ and $\bar{\mathbf{G}}_{rk} = \mathbf{J}_P^{'}\mathbf{G}_{rk}\mathbf{J}_P$, then, $\ln|\bar{\mathbf{S}}(\mathbf{\Lambda})| = -\ln f(\mathbf{\Lambda}) + \ln|\mathbf{S}(\mathbf{\Lambda})|$ where $\bar{\mathbf{S}}(\mathbf{\Lambda}) = \mathbf{I}_{N-R} - \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\bar{\mathbf{G}}_{rk}$, $\mathbf{S}(\mathbf{\Lambda}) = \mathbf{I}_N - \sum_{r=1}^{R}\sum_{k=1}^{R}\lambda_{rk}\mathbf{G}_{rk}$ and $f(\mathbf{\Lambda})$ is some function of $\mathbf{\Lambda}$.*

Proof. *Here, we show that the Lemma holds for two networks. From this, we can easily see that the Lemma holds generally. Define $\mathbf{H} = \begin{bmatrix} [\mathbf{P}_1, \frac{\iota_{n_1}}{\sqrt{n_1}}] & \mathbf{0} \\ \mathbf{0} & [\mathbf{P}_2, \frac{\iota_{n_2}}{\sqrt{n_2}}] \end{bmatrix}$. Then, we can show that $|\mathbf{H}^{'}(\mathbf{I} - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk})\mathbf{H}| = |\mathbf{H}^{'}\mathbf{H}||\mathbf{I} - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk}| = |\mathbf{I}_N - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk}|$ as $|\mathbf{H}^{'}\mathbf{H}| = 1$. Next, we show*

$$\mathbf{H}^{'}(\mathbf{I}_N - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk})\mathbf{H} = \begin{bmatrix} [\mathbf{P}_1, \frac{\iota_{n_1}}{\sqrt{n_1}}]^{'} & \mathbf{0} \\ \mathbf{0} & [\mathbf{P}_2, \frac{\iota_{n_2}}{\sqrt{n_2}}]^{'} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n_1} - \lambda_{11}\mathbf{W}_{11} & -\lambda_{12}\mathbf{W}_{12} \\ -\lambda_{21}\mathbf{W}_{21} & \mathbf{I}_{n_2} - \lambda_{22}\mathbf{W}_{22} \end{bmatrix} \begin{bmatrix} [\mathbf{P}_1, \frac{\iota_1}{\sqrt{n_1}}] & \mathbf{0} \\ \mathbf{0} & [\mathbf{P}_2, \frac{\iota_2}{\sqrt{n_2}}] \end{bmatrix}. \tag{21}$$

Now, $\mathbf{P}_r^{'}\mathbf{W}_{rk}\iota_{n_k} = \mathbf{0}$ and $\iota_{n_r}^{'}\mathbf{W}_{rk}\iota_{n_k} = n_r$. Hence,

$$\mathbf{H}^{'}(\mathbf{I}_N - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk})\mathbf{H} = \begin{bmatrix} \mathbf{P}_1^{'}(\mathbf{I}_{n_1} - \lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & \mathbf{0} & -\lambda_{12}\mathbf{P}_1^{'}\mathbf{W}_{12}\mathbf{P}_2 & \mathbf{0} \\ \frac{\iota_{n_1}^{'}}{\sqrt{n_1}}(\mathbf{I}_{n_1} - \lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & 1 - \lambda_{11} & \frac{\iota_{n_1}^{'}}{\sqrt{n_1}}(-\lambda_{12}\mathbf{W}_{12})\mathbf{P}_2 & -\sqrt{\frac{n_1}{n_2}}\lambda_{12} \\ -\lambda_{21}\mathbf{P}_2^{'}\mathbf{W}_{21}\mathbf{P}_1 & \mathbf{0} & \mathbf{P}_2^{'}(\mathbf{I}_{n_2} - \lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 & \mathbf{0} \\ \frac{\iota_{n_2}^{'}}{\sqrt{n_2}}(-\lambda_{21}\mathbf{W}_{21})\mathbf{P}_1 & -\sqrt{\frac{n_2}{n_1}}\lambda_{21} & \frac{\iota_{n_2}^{'}}{\sqrt{n_2}}(\mathbf{I}_{n_2} - \lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 & 1 - \lambda_{22} \end{bmatrix} \tag{22}$$

Then, from Laplace's formula,

$$
\begin{aligned}
&|\mathbf{H}^{'}(\mathbf{I}_N - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk})\mathbf{H}| \\
&= (1-\lambda_{11})\begin{vmatrix} \mathbf{P}_1^{'}(\mathbf{I}_{n_1}-\lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & -\lambda_{12}\mathbf{P}_1^{'}\mathbf{W}_{12}\mathbf{P}_2 & \mathbf{0} \\ -\lambda_{21}\mathbf{P}_2^{'}\mathbf{W}_{21}\mathbf{P}_1 & \mathbf{P}_2^{'}(\mathbf{I}_{n_2}-\lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 & \mathbf{0} \\ \frac{\iota_{n_2}^{'}}{\sqrt{n_2}}(-\lambda_{21}\mathbf{W}_{21})\mathbf{P}_1 & \frac{\iota_{n_2}^{'}}{\sqrt{n_2}}(\mathbf{I}_{n_2}-\lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 & 1-\lambda_{22} \end{vmatrix} \\
&\quad + (-1)^{n_2}(-\sqrt{\frac{n_2}{n_1}}\lambda_{21})\begin{vmatrix} \mathbf{P}_1^{'}(\mathbf{I}_{n_1}-\lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & -\lambda_{12}\mathbf{P}_1^{'}\mathbf{W}_{12}\mathbf{P}_2 & \mathbf{0} \\ \frac{\iota_{n_1}^{'}}{\sqrt{n_1}}(\mathbf{I}_{n_1}-\lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & \frac{\iota_{n_1}^{'}}{\sqrt{n_1}}(-\lambda_{12}\mathbf{W}_{12})\mathbf{P}_2 & -\sqrt{\frac{n_1}{n_2}}\lambda_{12} \\ -\lambda_{21}\mathbf{P}_2^{'}\mathbf{W}_{21}\mathbf{P}_1 & \mathbf{P}_2^{'}(\mathbf{I}_{n_2}-\lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 & \mathbf{0} \end{vmatrix} \qquad (23) \\
&= (1-\lambda_{11})(1-\lambda_{22})\begin{vmatrix} \mathbf{P}_1^{'}(\mathbf{I}_{n_1}-\lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & -\lambda_{12}\mathbf{P}_1^{'}\mathbf{W}_{12}\mathbf{P}_2 \\ -\lambda_{21}\mathbf{P}_2^{'}\mathbf{W}_{21}\mathbf{P}_1 & \mathbf{P}_2^{'}(\mathbf{I}_{n_2}-\lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 \end{vmatrix} \\
&\quad - \lambda_{12}\lambda_{21}\begin{vmatrix} \mathbf{P}_1^{'}(\mathbf{I}_{n_1}-\lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & -\lambda_{12}\mathbf{P}_1^{'}\mathbf{W}_{12}\mathbf{P}_2 \\ -\lambda_{21}\mathbf{P}_2^{'}\mathbf{W}_{21}\mathbf{P}_1 & \mathbf{P}_2^{'}(\mathbf{I}_{n_2}-\lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 \end{vmatrix} \\
&= \Big((1-\lambda_{11})(1-\lambda_{22})-\lambda_{12}\lambda_{21}\Big)\underbrace{\begin{vmatrix} \mathbf{P}_1^{'}(\mathbf{I}_{n_1}-\lambda_{11}\mathbf{W}_{11})\mathbf{P}_1 & -\lambda_{12}\mathbf{P}_1^{'}\mathbf{W}_{12}\mathbf{P}_2 \\ -\lambda_{21}\mathbf{P}_2^{'}\mathbf{W}_{21}\mathbf{P}_1 & \mathbf{P}_2^{'}(\mathbf{I}_{n_2}-\lambda_{22}\mathbf{W}_{22})\mathbf{P}_2 \end{vmatrix}}_{\#}
\end{aligned}
$$

Now, the determinant # is equal to $|\mathbf{J}_P^{'}(\mathbf{I}_N - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk})\mathbf{J}_P| = |I_{N-2} - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\bar{\mathbf{G}}_{rk}|$, *implying that* $|\mathbf{I}_N - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\mathbf{G}_{rk}| = \Big((1-\lambda_{11})(1-\lambda_{22})-\lambda_{12}\lambda_{21}\Big)|\mathbf{I}_{N-2} - \sum_{r=1}^{2}\sum_{k=1}^{2}\lambda_{rk}\bar{\mathbf{G}}_{rk}|$. *Therefore, the Lemma holds for the* $R=2$ *case. From this, we can easily see that the Lemma holds for any number of networks. For example, when there are* $R=3$ *networks,* $f(\Lambda) = (1-\lambda_{11})(1-\lambda_{22})(1-\lambda_{33}) - (1-\lambda_{11})\lambda_{23}\lambda_{32} - (1-\lambda_{22})\lambda_{13}\lambda_{31} - (1-\lambda_{33})\lambda_{21}\lambda_{12} - \lambda_{13}\lambda_{21}\lambda_{32}$. ■

D Additional Simulated Results

To save space Table 2 (in the main text) contains selected results for the four DGPs defined in section 3 of the main text. In particular, Table 2 provides simulated results for estimation of network-level coefficients for only two sample sizes: $\{n_r, T\} = \{5, 500\}$ and $= \{5, 1000\}$, using two ways to estimate the selection probabilities: Multinomial Logit (MLN) and 5-fold cross-validated Random Forest (RF). Here, we provide additional results for intermediate sample size, $\{n_r, T\} = \{10, 500\}$, and an additional estimator of the selection probabilities: out-of-bag tuned RF. Table 4 provides the results for all three sample sizes for MNL estimation of the probabilities. Table 5 provides the results for both the 5-fold cross-validated and out-of-bag tuned RF (described below). Therefore, Table 2 (in the main text) provides a subset of the results in Tables 4 and 5 below. For comparison purposes, Table 6 provides the results for RF with no tuning.

We consider two ways of tuning the maximum number of splits in a tree. The first method is based on the out-of-bag (OOB) classification error. RF is a bagging method which grows trees on bootstrapped subsets of the observations. On average, around two-third of the observations are used to fit a tree, so the remaining observations can be used to calculate an out-of-sample classification error by excluding the trees grown on bootstraps that include observation i when computing the i^{th} observation's prediction. This is computationally easiest but may be subject to overfitting because the OOB error is similar to a leave-one-out cross-validation error when the number of trees grown is large (James et al., 2017). The second method uses five-fold cross-validation in which the data are divided into five subsets of similar size. This method is more computational intensive, but may be less subject to overfitting, as each possible subset is used as a test set in alternation and is validated against the complement of the subset. For comparison, we consider RF without tuning where we grow 1,000 trees with a maximum number of splits $= T - 1$.

In almost all cases, Tables 4 and 5 show that RF with five-fold cross-validation performs best in terms of both bias and variance of the estimates. When it is not tuned (Table 6), the bias of RF is as large as that of MNL, even in DGP 3 and 4 (Table 4). This result aligns well with the findings in Scornet et al. (2015) and Wager and Walther (2015) that the depth of the trees (i.e. the number of splits in a tree) should be controlled to eliminate bias from the inherent data-dependent procedures in growing trees. Wager and Athey (2018, Appendix B) explain that the bias may be due to the fact that RF tends to push outliers into corners of features space. Minimizing this sort of bias by a proper tuning of the number of splits may be important in our context as selection bias can only be corrected when outliers are correctly identified. The result may provide practical guidance on the selection of the maximum number of splits that satisfies the theoretical condition on depth of a decision tree for the consistency of random forest in Scornet et al. (2015) and Wager and Walther (2015).

E NBA Data and Additional Empirical Results

Descriptive statistics for the variables in the outcome and selection equations are in Tables 7 and 8. The first columns contain the abbreviated names of the 30 NBA teams, which are partitioned by division (e.g., *Atlantic Division* of the *East Conference*) and ranked within division by the mean of our outcome variable, $Wins$ (second column), which is measured as *Wins Produced* per minute. For example, in the 2015-16 season Cleveland (CLE) had an average *Wins Produced* per minute of 0.0066 (second column) with a standard deviation of 0.0264 (third column). Cleveland played in 1,883 sampled time periods (column heading $Periods$) with an average duration of 1.7611 minutes per period (column heading ADP =average duration per period). Over short time intervals player-time level *Wins Produced* can be highly variable and sometimes quite small (non-pivotal) due to many zeros in the box score. During this season, the Golden State Warriors (GSW) had the highest *Wins Produced* per minute (0.0081) and Philadelphia (PHI) had the lowest (0.0051). Not surprisingly, the Warriors won the most games that season (72) and Philadelphia the fewest (10). Therefore, if the goal is to win games, then our outcome measure seems appropriate for these data.

Rounding out Table 7 are the explanatory variables in the outcome equation. As previously stated, *Experience* is the cumulative minutes played by a player from the beginning of the game to the end of period $t-1$. The *Fatigue* variable is total minutes continuously played by the player at the end of period $t-1$. For example, Milwaukee (MIL) had values of experience and fatigue of 14.3271 and 5.5877 minutes, respectively. The Milwaukee cagers play many minutes (on average), and many of these minutes are continuous, compared to other teams in the league. Philadelphia (PHI) had the lowest values for the experience variable, while Atlanta (ATL) had the lowest fatigue variable. It is likely that these coaches substitute many players more frequently than the rest of the league. This is reflected in their relatively low values for "Average Duration per Period" (ADP) of 1.6193 minutes (PHI) and 1.6391 minutes (ATL). The RPI variable is "Ratings Percentage Index" and is a measure of the opposing team's power rating at the end of the previous season.[46] Higher ratings reflect tougher opposing teams, so Phoenix (PHX) faced the toughest schedule ($RPI = 0.5028$) during the 2015-16 season, based on power ratings at the end of the 2014-15 season.

Table 8 contains means and standard deviations for the variables in the selection equation. The second column contains dependent variable ($Nguard$), and we see that the Portland coach (POR) uses on average the most guards (2.5), while Chicago (CHI) uses the fewest (1.44). This is not surprising

[46]The RPI for each opposing team is obtained from ESPN: http://www.espn.com/nba/stats/rpi/year/2015.

given that Portland averaged the most guards on their roster ($NguradR = 4.98$) in any game and Chicago had the fewest ($NguradR = 3.14$). The Golden State Warriors (GSW) had the largest positive cumulative score differential *CumscoreDiff* =6.62, while Philadelphia (PHI) had the largest negative score differential (-6.03). Golden State tended to be leading in points over the season, while Philadelphia was often losing. This is also not surprising given their best and worst (respectively) records at the end of the season. Boston (BOS) had the largest number of cumulative fouls over a game ($CumFoul = 11.20$) and the San Antonio Spurs (SAS) had the fewest (8.54 fouls).

The reduced form of equation 15 (in the main text) provides insight onto the overall network effects in the production process. Suppressing subscripts, the matrix form of equation 15 is,

$$\mathbf{y} = \mathbb{M} \cdot \mathbf{y} + \mathbf{XB} + \mathbf{u},$$

where $\mathbf{y} = \left[\begin{smallmatrix} \mathbf{y}_{rtg} \\ \mathbf{y}_{ktg} \end{smallmatrix}\right]$, a 10×1 vector, and $\mathbb{M} = \left[\begin{smallmatrix} \lambda_{rr}\mathbf{W}_{rrtg} & \lambda_{rk}\mathbf{W}_{rktg} \\ \lambda_{kr}\mathbf{W}_{krtg} & \lambda_{kk}\mathbf{W}_{kktg} \end{smallmatrix}\right]$, a 10×10 network matrix. Solving for the reduced form and taking expectations conditional on $\mathbf{X}$ yields

$$E(\mathbf{y}|\mathbf{X}) = (\mathbf{I}_{10} - \mathbb{M})^{-1}\mathbf{XB}. \tag{24}$$

This implies that the reduced-form peer- and competitor-effects are polynomial functions of the structural peer- and competitor-effects of any head-to-head match-up. For example, the second-order network term, $\mathbb{M}^2$, is given by

$$\begin{aligned} \mathbb{M}^2 &= \left[\begin{smallmatrix} \lambda_{rr}\mathbf{W}_{rrtg} & \lambda_{rk}\mathbf{W}_{rktg} \\ \lambda_{kr}\mathbf{W}_{krtg} & \lambda_{kk}\mathbf{W}_{kktg} \end{smallmatrix}\right] \left[\begin{smallmatrix} \lambda_{rr}\mathbf{W}_{rrtg} & \lambda_{rk}\mathbf{W}_{rktg} \\ \lambda_{kr}\mathbf{W}_{krtg} & \lambda_{kk}\mathbf{W}_{kktg} \end{smallmatrix}\right] \\ &= \left[\begin{smallmatrix} \mathbf{m}_1 & \mathbf{m}_2 \\ \mathbf{m}_3 & \mathbf{m}_4 \end{smallmatrix}\right] \end{aligned}$$

where

$$\begin{aligned} \mathbf{m}_1 &= (\lambda_{rr}\mathbf{W}_{rrtg})^2 + \lambda_{rk}\lambda_{kr}\mathbf{W}_{rktg}\mathbf{W}_{krtg} \\ \mathbf{m}_2 &= \lambda_{rr}\lambda_{rk}\mathbf{W}_{rrtg}\mathbf{W}_{rktg} + \lambda_{kk}\lambda_{rk}\mathbf{W}_{kktg}\mathbf{W}_{rktg} \\ \mathbf{m}_3 &= \lambda_{rr}\lambda_{kr}\mathbf{W}_{rrtg}\mathbf{W}_{krtg} + \lambda_{kk}\lambda_{kr}\mathbf{W}_{kktg}\mathbf{W}_{krtg} \\ \mathbf{m}_4 &= (\lambda_{kk}\mathbf{W}_{kktg})^2 + \lambda_{rk}\lambda_{kr}\mathbf{W}_{rktg}\mathbf{W}_{krtg} \end{aligned}$$

The first component of $\mathbf{m}_1$ represents the second-order, pure peer-effect of team r on itself ($r \to r \to r$). The second component of $\mathbf{m}_1$ represents the competitor-effect that arises from the fact that r's performance affects k's performance but it feeds back to r through the competitor network ($r \to k \to r$). The first component of $\mathbf{m}_2$ represents a mixed effect that arise from the fact that k affects r's performance which, in turn, affects r again through its own peer network ($k \to r \to r$). The second component of $\mathbf{m}_2$ represents a mixed effect that arises from the enhanced performance due to the network effect of team k that is affecting team r's performance through the competition network ($k \to k \to r$). Terms $\mathbf{m}_3$ and $\mathbf{m}_4$ can be similarly understood, but from the perspective of team k.

We calculate the 10×10 reduced-form matrix $(\mathbf{I}_{10} - \mathbb{M})^{-1}$ in (24) for every network in each period t for every r vs. k match-up over the season. Since the elements of the upper, right-hand 5×5 submatrix of $(\mathbf{I}_{10} - \mathbb{M})^{-1}$ capture the competitor-effect from k to r for each of the 5 players on team r, we report the average of the row-sums of the submatrices for every period for every match-up of team r vs. k over the season. This average embodies a reduced-form "indirect competitor-effect" from k to r.[47] We also report the same for the lower left-hand 5×5 submatrix, which captures the competitor-effect from r to k. These two reduced-form average effects corresponded in sign and magnitude to the structural competitor-effects λ_{rk} and λ_{kr} within a division. Therefore, it appears that for these data the first-order structural competitor-effects are an excellent proxy for the reduced-form competitor-effects.[48]

Table 9 contains the reduced-form indirect competitor-effects. We do not include t-statics for the indirect competitor-effects, but the point of this table is to show the similarity between these results and those of Table 3 in the main text. For example, in Table 9 the average indirect competitor-effect of Boston (BOS) on Toronto (TOR) is -0.220, while the corresponding structural parameter in Table 3 is -0.229. In Table 9 the average indirect competitor-effect of New York (NYK) on Boston (BOC) is -0.508, while the corresponding structural parameter in Table 3 is -0.569. Results are similar for other match-ups, as well. Apparently, the reduced-form competitor-effects are largely determined by the first-order effect, which consists solely of the structural parameter, λ_{rk}, so we would not be making a grave error if we interpreted the structural competitor-effects in Table 3 as reduced-form indirect competitor-effects. We also experimented with a "direct" and "indirect own network effect" by calculating the average of the diagonal elements and the row-sums, but without the diagonal elements, respectively, of the 5×5 upper left-hand corner submatrix of the matrix $(\mathbf{I}_{10} - \mathbb{M})^{-1}$ over all time periods for the games played between teams r and k within a division. We found that this average was affected by the relative signs of λ_{rk} and λ_{kr}. If the signs were the same, then it increased the

[47] This is similar to the direct and indirect reduced-form spatial effects discussed in LeSage and Pace (2009) and Elhorst (2014). See Glass et al. (2016), Table 5 for an example.

[48] In fact, we find that any row sum of the upper, right-hand 5×5 submatrix approximately equals the competitor coefficient, λ_{rk}, which we believe is due to the topology of the competition matrix.

average effect, but if the signs were different, then this effect was smaller. The point is that without accounting for the competitor-effects (which the literature has often overlooked) the within network effects may not be accurately measured.

In Table 3 (of the main text) the sign of the estimated λ_{rk} is almost always the opposite of the sign of the estimated λ_{kr}, which makes sense from a competitive standpoint, but the 120 intra-divisional estimates are numerous and difficult to interpret on their own. Therefore, we also report the 60 products, $\lambda_{rk}\lambda_{kr}$, which appears in both m_1 and m_4 of the second-order network effect, of the reduced-form matrix $(\mathbf{I}_{10} - \mathbb{M})^{-1}$. This is only a partial structural effect, but it provides an easier way to simultaneously summarize the competitor-effects for *both teams* in any match-up between r and k than the individual competitor-effects alone. It is a single statistic that captures (in some sense) the intensity of the rivalry between r and k (and vice versa) after controlling for individual performance, within-team chemistry and coaching. These statistics should be particularly relevant for quantifying intensity of intra-divisional rivalry.[49] These are tabulated in Table 10. For example, consider the Toronto Raptors (TOR) in the first row. The signs and the relative magnitudes of the competitor-effect products suggest that the New York Knicks (NYK) are a much tougher opponent for Toronto (-0.141) than is Boston (-0.064). The "toughest" match-up in the Atlantic division in terms of competitor-effects is New York vs. Boston, where the product of the competitor-effects is largest in magnitude in the division (-0.319). Looking down the rows of the table, the toughest match-ups in each division in terms of competitor-effects are: CLE-MIL (-0.575) in the Central Division, CHA-ORL (-0.468) in the Southeast, UTA-DEN (-0.320) in the Northwest, PHX-LAL (-0.206) in the Pacific, and HOU-NOP (-0.199) in the Southwest.

Estimates of β_{1r}, estimates of β_{2r} and the coefficient on the selection bias based on Lee's approach, and estimates of β_{2r} based on Dahl's approach are in Table 11, 12 and 13, respectively. An interesting result in Table 12 is that the coefficient on the selection bias term in the final-step estimates is almost always insignificant, using Lee's approach. Therefore, either the selection model is poorly specified or NBA coaches had little strategic effect on player productivity after accounting for peer- and competitor-effects in the 2015-16 season. Only the LA Clippers (LAC in the MNL model) and the LA Lakers (LAL in the RF model) had significant coach selection effects at the 95% significance level.

[49] We use the delta method to calculate standard errors of the competitor-effects products.

Table 4: Simulated Estimation of Network-Level Coefficients: MNL

		No Bias Term			MNL					
					Lee				Dahl	
$\{n_r, T\}$		$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	ρ	$\beta_2(1)$	$\beta_2(2)$
DGP1: Lee + Linear Utility Function										
$\{5, 500\}$	Bias	-1.40	-0.32	0.51	-0.01	-0.01	-0.00	0.01	-0.10	0.12
	SD	0.20	0.27	0.29	0.23	0.25	0.25	0.16	0.42	0.46
	Avg. Bootstrap SE	0.19	0.27	0.28	0.23	0.25	0.25	0.16	0.44	0.47
	RMSE	1.42	0.42	0.59	0.23	0.25	0.25	0.17	0.57	0.58
	95% Coverage Rate	0.00	0.78	0.54	0.95	0.95	0.95	0.97	0.99	0.99
$\{10, 500\}$	Bias	-1.40	-0.33	0.51	-0.03	-0.01	0.02	0.02	-0.08	0.11
	SD	0.19	0.27	0.26	0.22	0.24	0.24	0.16	0.43	0.43
	Avg. Bootstrap SE	0.19	0.27	0.27	0.22	0.24	0.24	0.16	0.43	0.45
	RMSE	1.41	0.43	0.57	0.23	0.24	0.24	0.17	0.57	0.55
	95% Coverage Rate	0.00	0.76	0.54	0.95	0.96	0.96	0.95	0.99	0.99
$\{5, 1000\}$	Bias	-1.40	-0.34	0.50	-0.01	0.00	0.00	0.00	-0.07	0.09
	SD	0.15	0.20	0.22	0.16	0.18	0.18	0.11	0.36	0.34
	Avg. Bootstrap SE	0.14	0.19	0.20	0.17	0.18	0.18	0.13	0.40	0.39
	RMSE	1.41	0.39	0.55	0.16	0.18	0.18	0.11	0.37	0.35
	95% Coverage Rate	0.00	0.59	0.26	0.95	0.94	0.94	0.96	0.97	0.98
DGP2: Dahl + Linear Utility Function										
$\{5, 500\}$	Bias	-1.48	-0.45	0.67	-0.41	-0.20	0.30	0.40	-0.03	0.12
	SD	0.20	0.28	0.31	0.23	0.28	0.28	0.13	0.42	0.34
	Avg. Bootstrap SE	0.20	0.28	0.29	0.23	0.27	0.27	0.14	0.50	0.47
	RMSE	1.49	0.53	0.74	0.47	0.35	0.41	0.42	0.42	0.36
	95% Coverage Rate	0.00	0.67	0.36	0.57	0.87	0.80	0.21	0.99	0.99
$\{10, 500\}$	Bias	-1.47	-0.47	0.67	-0.42	-0.21	0.30	0.40	-0.04	0.11
	SD	0.19	0.28	0.29	0.23	0.27	0.27	0.13	0.33	0.31
	Avg. Bootstrap SE	0.19	0.28	0.28	0.23	0.26	0.26	0.13	0.47	0.44
	RMSE	1.48	0.54	0.73	0.48	0.34	0.40	0.42	0.33	0.33
	95% Coverage Rate	0.00	0.59	0.34	0.54	0.85	0.80	0.15	0.99	0.99
$\{5, 1000\}$	Bias	-1.47	-0.46	0.67	-0.42	-0.20	0.30	0.40	0.00	0.07
	SD	0.15	0.22	0.21	0.16	0.19	0.18	0.09	0.25	0.24
	Avg. Bootstrap SE	0.14	0.20	0.21	0.17	0.19	0.19	0.11	0.33	0.31
	RMSE	1.48	0.51	0.71	0.45	0.28	0.35	0.41	0.25	0.25
	95% Coverage Rate	0.00	0.36	0.10	0.28	0.81	0.67	0.03	0.99	0.99
DGP3: Lee + Nonlinear Utility Function										
$\{5, 500\}$	Bias	-2.07	0.77	0.75	-0.16	0.17	0.18	0.06	0.16	0.17
	SD	0.17	0.30	0.30	0.27	0.27	0.27	0.19	0.29	0.29
	Avg. Bootstrap SE	0.17	0.30	0.30	0.27	0.27	0.27	0.19	0.29	0.30
	RMSE	2.08	0.82	0.81	0.31	0.32	0.33	0.19	0.33	0.34
	95% Coverage Rate	0.00	0.25	0.30	0.90	0.91	0.90	0.93	0.92	0.92
$\{10, 500\}$	Bias	-2.07	0.76	0.75	-0.16	0.20	0.18	0.05	0.19	0.18
	SD	0.19	0.30	0.29	0.27	0.27	0.25	0.19	0.28	0.28
	Avg. Bootstrap SE	0.17	0.29	0.29	0.27	0.26	0.26	0.18	0.29	0.29
	RMSE	2.08	0.81	0.80	0.32	0.33	0.31	0.19	0.34	0.33
	95% Coverage Rate	0.00	0.26	0.26	0.89	0.89	0.90	0.92	0.91	0.91
$\{5, 1000\}$	Bias	-2.07	0.76	0.76	-0.15	0.19	0.18	0.04	0.21	0.20
	SD	0.13	0.21	0.21	0.20	0.20	0.19	0.13	0.21	0.21
	Avg. Bootstrap SE	0.13	0.22	0.22	0.19	0.19	0.19	0.14	0.21	0.21
	RMSE	2.08	0.79	0.79	0.25	0.27	0.27	0.14	0.30	0.29
	95% Coverage Rate	0.00	0.06	0.06	0.87	0.84	0.85	0.93	0.84	0.85
DGP4: Dahl + Nonlinear Utility Function										
$\{5, 500\}$	Bias	-2.50	1.05	1.06	-0.83	0.57	0.56	0.27	0.49	0.49
	SD	0.16	0.31	0.31	0.24	0.28	0.28	0.16	0.30	0.30
	Avg. Bootstrap SE	0.16	0.30	0.30	0.25	0.28	0.28	0.16	0.31	0.31
	RMSE	2.50	1.10	1.10	0.87	0.63	0.62	0.31	0.57	0.58
	95% Coverage Rate	0.00	0.06	0.06	0.10	0.46	0.49	0.62	0.65	0.64
$\{10, 500\}$	Bias	-2.50	1.06	1.06	-0.83	0.57	0.56	0.27	0.49	0.48
	SD	0.14	0.29	0.28	0.24	0.26	0.27	0.15	0.29	0.29
	Avg. Bootstrap SE	0.15	0.29	0.29	0.24	0.27	0.27	0.16	0.30	0.30
	RMSE	2.50	1.10	1.10	0.87	0.62	0.62	0.31	0.56	0.57
	95% Coverage Rate	0.00	0.03	0.04	0.09	0.45	0.44	0.59	0.63	0.62
$\{5, 1000\}$	Bias	-2.50	1.06	1.07	-0.81	0.55	0.56	0.25	0.48	0.49
	SD	0.12	0.22	0.22	0.18	0.20	0.19	0.11	0.21	0.21
	Avg. Bootstrap SE	0.12	0.22	0.22	0.17	0.20	0.20	0.12	0.22	0.22
	RMSE	2.50	1.08	1.09	0.83	0.59	0.59	0.27	0.52	0.53
	95% Coverage Rate	0.00	0.01	0.00	0.01	0.20	0.18	0.45	0.39	0.36

Table 5: Simulated Estimation of Network-Level Coefficients: Random Forest

		RF + Out of Bag Tuning						RF + 5 Fold CV					
		Lee				Dahl		Lee				Dahl	
$\{n_r, T\}$		$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	ρ	$\beta_2(1)$	$\beta_2(2)$	$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	ρ	$\beta_2(1)$	$\beta_2(2)$
DGP1: Lee + Linear Utility Function													
$\{5, 500\}$	Bias	-0.12	-0.07	0.10	0.11	-0.14	0.20	-0.08	-0.06	0.07	0.04	-0.14	0.18
	SD	0.31	0.27	0.29	0.26	0.32	0.36	0.25	0.26	0.26	0.19	0.32	0.34
	Avg. Bootstrap SE	0.28	0.26	0.27	0.25	0.34	0.36	0.25	0.27	0.28	0.19	0.36	0.38
	RMSE	0.33	0.28	0.31	0.29	0.35	0.41	0.26	0.27	0.27	0.19	0.35	0.38
	95% Coverage Rate	0.89	0.93	0.90	0.89	0.94	0.91	0.93	0.95	0.95	0.94	0.96	0.96
$\{10, 500\}$	Bias	-0.11	-0.06	0.09	0.10	-0.16	0.20	-0.10	-0.06	0.09	0.05	-0.14	0.18
	SD	0.27	0.26	0.27	0.23	0.31	0.32	0.24	0.24	0.25	0.18	0.31	0.31
	Avg. Bootstrap SE	0.27	0.25	0.26	0.25	0.34	0.35	0.24	0.26	0.27	0.18	0.36	0.37
	RMSE	0.29	0.26	0.28	0.25	0.35	0.38	0.26	0.25	0.27	0.19	0.34	0.36
	95% Coverage Rate	0.90	0.94	0.92	0.91	0.95	0.93	0.93	0.95	0.95	0.93	0.96	0.96
$\{5, 1000\}$	Bias	-0.07	-0.07	0.07	0.08	-0.15	0.19	-0.06	-0.04	0.06	0.03	-0.14	0.17
	SD	0.23	0.21	0.21	0.23	0.25	0.25	0.18	0.19	0.19	0.13	0.24	0.25
	Avg. Bootstrap SE	0.20	0.19	0.19	0.19	0.25	0.26	0.18	0.19	0.20	0.14	0.26	0.28
	RMSE	0.24	0.22	0.22	0.24	0.29	0.31	0.19	0.20	0.20	0.14	0.28	0.30
	95% Coverage Rate	0.90	0.92	0.91	0.89	0.91	0.91	0.93	0.95	0.95	0.94	0.94	0.94
DGP2: Dahl + Linear Utility Function													
$\{5, 500\}$	Bias	-0.50	-0.25	0.36	0.49	-0.21	0.33	-0.47	-0.23	0.35	0.42	-0.17	0.30
	SD	0.27	0.28	0.30	0.20	0.33	0.36	0.24	0.29	0.28	0.15	0.32	0.33
	Avg. Bootstrap SE	0.27	0.28	0.28	0.21	0.36	0.37	0.24	0.28	0.29	0.16	0.37	0.38
	RMSE	0.57	0.37	0.47	0.53	0.39	0.48	0.53	0.37	0.45	0.45	0.37	0.44
	95% Coverage Rate	0.52	0.84	0.74	0.35	0.92	0.85	0.50	0.85	0.76	0.28	0.94	0.91
$\{10, 500\}$	Bias	-0.49	-0.28	0.36	0.49	-0.22	0.33	-0.47	-0.25	0.35	0.43	-0.18	0.30
	SD	0.26	0.26	0.28	0.19	0.32	0.34	0.24	0.27	0.27	0.15	0.32	0.33
	Avg. Bootstrap SE	0.26	0.27	0.27	0.20	0.35	0.36	0.24	0.27	0.28	0.15	0.36	0.37
	RMSE	0.56	0.38	0.45	0.53	0.39	0.47	0.53	0.37	0.44	0.46	0.36	0.44
	95% Coverage Rate	0.51	0.83	0.74	0.31	0.94	0.86	0.49	0.84	0.76	0.18	0.95	0.90
$\{5, 1000\}$	Bias	-0.45	-0.26	0.35	0.45	-0.20	0.31	-0.46	-0.23	0.34	0.42	-0.17	0.27
	SD	0.20	0.21	0.20	0.14	0.26	0.26	0.17	0.20	0.19	0.11	0.25	0.26
	Avg. Bootstrap SE	0.19	0.19	0.20	0.15	0.26	0.27	0.17	0.20	0.21	0.12	0.27	0.28
	RMSE	0.49	0.34	0.40	0.47	0.33	0.40	0.49	0.30	0.39	0.43	0.30	0.37
	95% Coverage Rate	0.34	0.72	0.59	0.18	0.87	0.79	0.24	0.78	0.63	0.05	0.92	0.86
DGP3: Lee + Nonlinear Utility Function													
$\{5, 500\}$	Bias	-0.18	0.05	0.04	0.10	-0.03	-0.03	-0.02	-0.05	-0.03	-0.03	-0.11	-0.09
	SD	0.43	0.32	0.33	0.34	0.35	0.36	0.31	0.29	0.29	0.23	0.32	0.32
	Avg. Bootstrap SE	0.44	0.33	0.33	0.36	0.38	0.38	0.30	0.29	0.29	0.21	0.35	0.35
	RMSE	0.47	0.32	0.33	0.35	0.35	0.36	0.31	0.30	0.29	0.23	0.34	0.33
	95% Coverage Rate	0.92	0.95	0.95	0.93	0.97	0.97	0.94	0.94	0.95	0.95	0.95	0.96
$\{10, 500\}$	Bias	-0.17	0.03	0.03	0.09	-0.03	-0.03	-0.03	-0.01	-0.03	-0.03	-0.06	-0.07
	SD	0.40	0.31	0.31	0.32	0.35	0.34	0.32	0.29	0.28	0.22	0.32	0.33
	Avg. Bootstrap SE	0.44	0.32	0.32	0.36	0.38	0.38	0.30	0.28	0.29	0.21	0.35	0.35
	RMSE	0.44	0.31	0.31	0.34	0.35	0.35	0.32	0.29	0.28	0.22	0.33	0.34
	95% Coverage Rate	0.92	0.95	0.96	0.93	0.96	0.97	0.93	0.95	0.95	0.94	0.96	0.95
$\{5, 1000\}$	Bias	-0.09	-0.01	-0.01	0.05	-0.06	-0.06	0.02	-0.05	-0.06	-0.04	-0.08	-0.09
	SD	0.29	0.23	0.23	0.23	0.25	0.27	0.21	0.21	0.21	0.15	0.24	0.24
	Avg. Bootstrap SE	0.31	0.24	0.24	0.25	0.27	0.27	0.21	0.20	0.20	0.15	0.24	0.24
	RMSE	0.31	0.23	0.23	0.23	0.26	0.27	0.21	0.22	0.22	0.16	0.26	0.26
	95% Coverage Rate	0.92	0.96	0.96	0.94	0.95	0.95	0.95	0.94	0.94	0.95	0.93	0.92
DGP4: Dahl + Nonlinear Utility Function													
$\{5, 500\}$	Bias	-0.81	0.40	0.41	0.27	0.19	0.21	-0.69	0.37	0.36	0.17	0.16	0.16
	SD	0.37	0.32	0.33	0.31	0.37	0.36	0.29	0.29	0.30	0.22	0.34	0.34
	Avg. Bootstrap SE	0.39	0.32	0.32	0.33	0.38	0.38	0.28	0.29	0.29	0.20	0.36	0.36
	RMSE	0.89	0.51	0.52	0.41	0.42	0.42	0.74	0.47	0.47	0.27	0.37	0.38
	95% Coverage Rate	0.47	0.76	0.73	0.82	0.92	0.92	0.31	0.76	0.76	0.83	0.95	0.94
$\{10, 500\}$	Bias	-0.81	0.42	0.42	0.26	0.22	0.22	-0.68	0.36	0.36	0.16	0.15	0.15
	SD	0.35	0.31	0.30	0.30	0.37	0.35	0.28	0.28	0.29	0.21	0.32	0.34
	Avg. Bootstrap SE	0.40	0.31	0.31	0.34	0.38	0.38	0.27	0.28	0.28	0.19	0.35	0.35
	RMSE	0.88	0.52	0.51	0.40	0.43	0.41	0.74	0.46	0.46	0.27	0.35	0.37
	95% Coverage Rate	0.46	0.71	0.74	0.84	0.91	0.92	0.29	0.75	0.76	0.84	0.95	0.94
$\{5, 1000\}$	Bias	-0.73	0.38	0.38	0.23	0.15	0.14	-0.64	0.34	0.35	0.16	0.10	0.11
	SD	0.26	0.24	0.23	0.21	0.28	0.27	0.19	0.22	0.21	0.14	0.25	0.24
	Avg. Bootstrap SE	0.28	0.23	0.23	0.23	0.27	0.27	0.19	0.21	0.21	0.14	0.25	0.25
	RMSE	0.77	0.45	0.44	0.31	0.31	0.30	0.67	0.40	0.41	0.21	0.27	0.26
	95% Coverage Rate	0.27	0.61	0.63	0.82	0.92	0.93	0.10	0.61	0.60	0.75	0.94	0.94

Table 6: Simulated Estimation of Network-Level Coefficients: RF (No Tuning)

		Lee				Dahl	
$\{n_r, T\}$		$\beta_2(0)$	$\beta_2(1)$	$\beta_2(2)$	ρ	$\beta_2(1)$	$\beta_2(2)$
DGP1: Lee + Linear Utility Function							
$\{5, 500\}$	Bias	-0.43	-0.04	0.12	-1.45	-0.18	0.27
	SD	0.24	0.27	0.27	0.40	0.27	0.28
	RMSE	0.49	0.27	0.30	1.50	0.32	0.39
$\{10, 500\}$	Bias	-0.42	-0.05	0.11	-1.45	-0.19	0.26
	SD	0.22	0.24	0.26	0.39	0.25	0.26
	RMSE	0.47	0.25	0.28	1.51	0.31	0.36
$\{5, 1000\}$	Bias	-0.40	-0.05	0.10	-1.52	-0.19	0.26
	SD	0.16	0.19	0.19	0.30	0.19	0.20
	RMSE	0.43	0.19	0.22	1.55	0.27	0.33
DGP2: Dahl + Linear Utility Function							
$\{5, 500\}$	Bias	-0.70	-0.23	0.39	-0.83	-0.31	0.47
	SD	0.23	0.27	0.27	0.40	0.28	0.28
	RMSE	0.74	0.36	0.48	0.92	0.42	0.55
$\{10, 500\}$	Bias	-0.71	-0.23	0.36	-0.81	-0.30	0.44
	SD	0.22	0.28	0.27	0.40	0.29	0.28
	RMSE	0.74	0.36	0.45	0.90	0.42	0.52
$\{5, 1000\}$	Bias	-0.68	-0.23	0.36	-0.89	-0.31	0.45
	SD	0.17	0.20	0.20	0.30	0.21	0.21
	RMSE	0.70	0.30	0.41	0.94	0.38	0.50
DGP3: Lee + Nonlinear Utility Function							
$\{5, 500\}$	Bias	-0.59	0.21	0.22	-1.51	0.23	0.24
	SD	0.26	0.29	0.30	0.44	0.29	0.30
	RMSE	0.65	0.35	0.37	1.57	0.37	0.39
$\{10, 500\}$	Bias	-0.60	0.20	0.21	-1.48	0.23	0.23
	SD	0.24	0.27	0.26	0.40	0.28	0.27
	RMSE	0.65	0.34	0.33	1.53	0.36	0.35
$\{5, 1000\}$	Bias	-0.57	0.19	0.20	-1.55	0.22	0.23
	SD	0.17	0.20	0.18	0.29	0.20	0.19
	RMSE	0.60	0.27	0.27	1.58	0.30	0.29
DGP4: Dahl + Nonlinear Utility Function							
$\{5, 500\}$	Bias	-1.13	0.54	0.55	-1.23	0.53	0.54
	SD	0.24	0.28	0.28	0.40	0.29	0.28
	RMSE	1.16	0.61	0.61	1.29	0.61	0.61
$\{10, 500\}$	Bias	-1.13	0.55	0.56	-1.25	0.53	0.54
	SD	0.22	0.27	0.27	0.37	0.28	0.28
	RMSE	1.15	0.61	0.62	1.31	0.60	0.61
$\{5, 1000\}$	Bias	-1.11	0.53	0.55	-1.30	0.51	0.53
	SD	0.16	0.19	0.20	0.27	0.20	0.21
	RMSE	1.12	0.56	0.58	1.33	0.55	0.57

Table 7: NBA 2015-16 Descriptive Statistics: Outcome Function

	$Wins$ (wins/min)		$Experience$ (mins)		$Fatigue$ (mins)		RPI					
Team	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Games	Periods	Obs.	ADP
East Conference												
Atlantic Division												
TOR	0.0068	0.0254	13.9906	8.9852	4.8002	4.2214	0.4975	0.0402	82	1,802	9,010	1.8506
BOS	0.0061	0.0266	13.5679	8.7688	4.5319	4.1359	0.4970	0.0414	82	1,894	9,470	1.6840
NYK	0.0058	0.0251	13.2521	8.9107	4.4921	4.3276	0.5001	0.0409	82	1,847	9,235	1.8075
BKN	0.0053	0.0252	13.5829	8.9497	4.5742	4.2478	0.5008	0.0399	82	1,838	9,190	1.7807
PHI	0.0051	0.0265	13.1285	8.5196	3.6186	3.5643	0.5030	0.0347	82	2,061	10,305	1.6193
Central Division												
CLE	0.0066	0.0264	13.7307	9.0387	4.5803	4.2359	0.4948	0.0402	82	1,883	9,415	1.7611
IND	0.0061	0.0255	13.9000	8.9143	4.6639	4.1757	0.5003	0.0401	82	1,839	9,195	1.8162
DET	0.0059	0.0257	14.6659	9.4099	4.8601	4.5260	0.4981	0.0419	82	1,754	8,770	1.9050
CHI	0.0059	0.0249	13.4434	8.8153	4.6841	4.4133	0.4996	0.0418	82	1,793	8,965	1.8500
MIL	0.0055	0.0242	14.3271	9.5646	5.5877	5.3570	0.5006	0.0394	82	1,786	8,930	1.8766
Southeast Division												
MIA	0.0065	0.0253	14.4091	8.8739	4.7797	4.2856	0.4982	0.0397	82	1,847	9,235	1.8138
ATL	0.0063	0.0264	13.1450	8.6175	3.3534	3.3129	0.4996	0.0384	82	2,021	10,105	1.6391
CHA	0.0068	0.0254	13.5900	8.9516	4.9244	4.5149	0.4972	0.0406	82	1,741	8,705	1.9243
WAS	0.0061	0.0257	13.7136	9.0765	4.7863	4.4908	0.4996	0.0414	82	1,914	9,570	1.7293
ORL	0.0060	0.0249	13.3663	9.1810	5.1182	4.9035	0.4984	0.0415	82	1,815	9,075	1.8481
West Conference												
Northwest Division												
OKC	0.0073	0.0265	13.5447	8.9972	4.5932	4.3073	0.4994	0.0413	82	1,857	9,285	1.7843
POR	0.0066	0.0261	13.4148	8.7288	4.6371	4.1711	0.5018	0.0416	82	1,839	9,195	1.8099
UTA	0.0059	0.0256	13.5649	8.9888	3.8109	3.7618	0.5001	0.0446	82	2,035	10,175	1.6233
DEN	0.0058	0.0257	13.4496	8.7122	5.0957	4.6100	0.5029	0.0438	82	1,843	9,215	1.8071
MIN	0.0058	0.0253	13.1619	8.9522	4.8360	4.3706	0.5012	0.0406	82	1,891	9,455	1.7825
Pacific Division												
GSW	0.0081	0.0268	13.1567	9.1074	4.2704	4.0395	0.4968	0.0372	82	1,892	9,460	1.7745
LAC	0.0065	0.0263	13.0637	8.7264	4.6467	4.2978	0.4997	0.0422	82	1,862	9,310	1.7728
SAC	0.0064	0.0261	14.1710	9.0768	4.8797	4.5854	0.5018	0.0431	82	1,957	9,785	1.7061
PHX	0.0053	0.0261	13.9407	9.4149	5.2119	5.0957	0.5028	0.0416	82	1,821	9,105	1.8327
LAL	0.0056	0.0256	12.7579	8.3083	4.9857	4.4327	0.5042	0.0403	82	1,752	8,760	1.9016
Southwest Division												
SAS	0.0072	0.0253	12.2602	8.2964	3.8246	3.6986	0.4980	0.0398	82	1,878	9,390	1.7312
DAL	0.0058	0.0257	13.4748	8.8809	4.0182	3.9327	0.5012	0.0434	82	1,968	9,840	1.7049
MEM	0.0056	0.0252	13.5685	8.8102	4.8904	4.4279	0.4992	0.0439	82	1,855	9,275	1.7359
HOU	0.0063	0.0269	13.8216	9.4650	4.4653	4.7234	0.5012	0.0419	82	1,964	9,820	1.7123
NOP	0.0060	0.0252	14.2266	9.1684	5.0658	4.5300	0.5015	0.0408	82	1,875	9,375	1.7772

Teams ranked on team chemistry ($\lambda_r r$) from the homogeneous competitor model in Table 3. SD: Standard Deviation; ADP: Average Duration per Period; RPI: Ratings Percentage Index is a power measure of the opposing team.

Table 8: NBA 2015-16 Descriptive Statistics: Selection Equation

	$Nguard_{rt}$		$CumScoreDiff_{rt}$		$CurScoreDiff_{rt}$		$CumFoul_{rt}$		$CurFoul_{rt}$		$CurTime_{rt}$		$Duration_{rt}$		$NguardR_{rt}$		$NguardOPP_{rt}$	
Team	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
East Conference																		
Atlantic Division																		
TOR	2.17	0.54	1.92	8.99	0.17	3.18	9.87	6.24	0.69	0.87	25.11	12.44	1.70	1.48	3.79	0.75	1.95	0.61
BOS	1.89	0.59	1.59	10.07	0.06	3.02	11.20	6.79	0.71	0.95	25.64	12.49	1.55	1.42	3.59	0.83	1.93	0.62
NYK	2.18	0.42	-0.37	10.60	-0.13	3.05	9.75	5.96	0.67	0.87	25.15	12.47	1.69	1.46	4.43	0.54	1.86	0.58
BKN	1.87	0.66	-3.85	9.96	-0.32	3.04	9.30	5.48	0.61	0.80	25.58	12.35	1.64	1.52	4.04	0.71	1.86	0.58
PHI	1.97	0.58	-6.03	10.54	-0.30	3.02	10.76	6.61	0.64	0.80	24.72	12.64	1.42	1.20	4.06	0.69	1.96	0.61
Central Division																		
CLE	2.15	0.55	3.31	11.19	0.23	3.01	9.64	5.98	0.67	0.86	24.84	12.30	1.61	1.43	4.39	0.85	1.80	0.58
IND	1.90	0.40	1.09	9.26	0.11	3.01	9.88	6.20	0.67	0.88	25.54	12.55	1.66	1.44	3.81	0.83	1.96	0.64
DET	1.68	0.47	0.41	10.80	0.02	3.17	9.39	5.82	0.67	0.85	25.63	12.50	1.76	1.59	3.06	0.42	1.89	0.56
CHI	1.44	0.52	-0.78	9.28	-0.10	3.09	9.08	5.78	0.65	0.83	25.24	12.45	1.66	1.43	3.14	0.56	1.89	0.59
MIL	1.74	0.62	-1.96	9.56	-0.19	3.08	10.22	6.30	0.71	0.88	24.82	12.59	1.69	1.45	3.56	0.64	1.81	0.56
Southeast Division																		
MIA	1.83	0.46	1.23	9.92	0.09	3.02	9.14	5.58	0.62	0.84	24.99	12.56	1.64	1.41	3.39	0.72	1.91	0.61
ATL	1.89	0.52	1.23	9.91	0.09	2.95	9.38	5.99	0.57	0.80	25.50	12.67	1.47	1.30	4.28	0.87	1.85	0.58
CHA	2.10	0.61	1.17	10.88	0.15	3.21	9.03	5.89	0.66	0.87	25.04	12.44	1.81	1.53	4.37	1.26	1.92	0.60
WAS	2.32	0.64	0.19	10.54	0.03	3.06	10.32	6.28	0.68	0.87	25.11	12.49	1.57	1.44	4.07	0.73	1.97	0.62
ORL	1.60	0.54	0.01	10.27	-0.04	3.08	10.38	6.30	0.74	0.93	25.00	12.54	1.71	1.45	3.25	0.77	1.93	0.59
West Conference																		
Northwest Division																		
OKC	2.36	0.56	4.60	10.05	0.29	3.08	10.32	6.14	0.68	0.85	25.36	12.37	1.60	1.46	5.04	0.71	1.88	0.59
POR	2.50	0.55	0.80	11.29	0.02	3.09	10.84	6.69	0.73	0.93	25.22	12.47	1.65	1.48	4.98	1.06	1.91	0.59
UTA	1.76	0.53	1.39	11.04	0.12	2.81	9.77	6.27	0.59	0.78	25.15	12.40	1.47	1.28	3.53	0.53	1.90	0.57
DEN	1.83	0.52	-2.06	9.31	-0.10	3.09	10.57	6.44	0.68	0.90	25.27	12.55	1.61	1.44	3.37	0.59	1.95	0.60
MIN	1.58	0.51	-1.62	9.37	-0.15	3.05	10.14	6.04	0.67	0.89	24.56	12.51	1.61	1.38	3.17	0.51	1.89	0.55
Pacific Division																		
GSW	2.19	0.46	6.62	10.55	0.41	3.20	10.34	6.32	0.72	0.89	24.89	12.54	1.65	1.41	4.53	0.60	1.99	0.62
LAC	2.36	0.54	2.22	10.00	0.14	2.99	10.67	6.68	0.69	0.91	25.00	12.43	1.57	1.43	4.59	0.81	1.96	0.57
SAC	1.73	0.55	-1.90	10.37	-0.12	2.97	10.12	6.02	0.63	0.85	24.87	12.35	1.56	1.34	3.15	0.64	1.99	0.60
PHX	2.03	0.50	-4.06	10.85	-0.18	3.17	11.59	6.97	0.77	0.93	25.09	12.44	1.62	1.39	4.02	0.74	1.90	0.58
LAL	1.91	0.30	-6.61	10.43	-0.36	3.34	10.13	6.30	0.73	0.94	24.81	12.53	1.77	1.56	3.38	0.56	1.86	0.58
Southwest Division																		
SAS	1.51	0.55	5.56	10.28	0.35	2.90	8.54	5.43	0.56	0.79	25.10	12.37	1.57	1.40	3.31	0.64	1.92	0.57
DAL	1.98	0.57	-0.78	9.82	-0.07	3.01	9.60	6.05	0.60	0.78	25.15	12.62	1.51	1.30	3.72	0.66	2.02	0.63
MEM	1.48	0.52	-0.77	10.22	-0.08	2.94	10.68	6.58	0.73	0.92	24.97	12.45	1.61	1.43	2.93	0.74	1.95	0.58
HOU	1.92	0.45	-0.61	10.19	-0.01	3.16	10.74	6.75	0.70	0.90	24.52	12.52	1.52	1.27	3.49	0.50	2.00	0.65
NOP	1.90	0.55	-1.99	9.75	-0.15	3.14	10.55	6.21	0.69	0.86	25.07	12.48	1.60	1.33	3.49	0.77	1.94	0.61

SD: Standard Deviation. $CumScoreDiff_{rt}$ ($CurScoreDiff_{rt}$) is the cumulative (current) score difference between the two teams at the end of period $t-1$ (in period $t-1$); $CumFoul_{rt}$ ($CurFoul_{rt}$) is the the cumulative (current) number of fouls committed at the end of period $t-1$ (in period $t-1$): $CurTime_{rt} \in [0, 48]$ is the game time at the start of period t in minutes; $Duration_{rt}$ is the duration of period $t-1$; $NguardR_{rt}$, $NguardP_{rt}$, and $NguardOPP_{rt}$ are the number of guards on the roster of the game, of the period $t-1$ and of the opposing team of the period $t-1$, respectively.

Table 9: Reduced-Form Indirect Competitor Effects within Division

East Conference

Atlantic Division

	TOR	BOS	NYK	BKN	PHI
TOR	-	-0.220	0.507	-0.167	-0.040
BOS	0.268	-	-0.508	0.451	-0.460
NYK	-0.231	0.384	-	-0.312	0.193
BKN	-0.007	-0.415	0.344	-	-0.290
PHI	0.060	0.215	-0.183	0.388	-

Central Division

	CLE	IND	DET	CHI	MIL
CLE	-	0.251	-0.408	-0.331	0.541
IND	-0.319	-	0.288	-0.238	0.129
DET	0.424	-0.453	-	-0.364	-0.542
CHI	0.459	0.197	0.180	-	-0.483
MIL	-0.470	-0.233	0.452	0.356	-

Southeast Division

	MIA	ATL	CHA	WAS	ORL
MIA	-	-0.016	0.098	-0.190	-0.088
ATL	-0.096	-	0.325	0.470	-0.273
CHA	-0.061	-0.563	-	-0.034	0.587
WAS	0.144	-0.316	0.059	-	-0.180
ORL	0.419	0.327	-0.398	0.262	-

West Conference

Northwest Division

	OKC	POR	UTA	DEN	MIN
OKC	-	0.068	0.348	-0.423	0.123
POR	0.035	-	0.243	-0.234	-0.349
UTA	-0.280	-0.182	-	0.446	0.018
DEN	0.148	0.134	-0.409	-	-0.181
MIN	-0.140	0.326	0.180	0.072	-

Pacific Division

	GSW	LAC	SAC	PHX	LAL
GSW	-	0.025	-0.195	-0.494	-0.268
LAC	0.036	-	0.156	0.064	-0.214
SAC	-0.006	-0.210	-	-0.414	0.387
PHX	0.292	-0.109	0.422	-	0.489
LAL	0.297	0.208	-0.271	-0.305	-

Southwest Division

	SAS	DAL	MEM	HOU	NOP
SAS	-	0.288	0.067	-0.134	0.370
DAL	-0.370	-	0.115	0.386	-0.362
MEM	-0.038	-0.055	-	0.031	0.070
HOU	0.266	-0.304	0.160	-	0.374
NOP	-0.254	0.282	0.016	-0.375	-

Average Indirect effect from competitors is the average of row sum of the right-hand corner 5×5 submatrix of the 10×10 network matrix over all time periods for a given team and opposing team.

Table 10: Product Competitor Effect Estimates, $\lambda_{rk}\lambda_{kr}$, within Division

East Conference

	Atlantic Division					t-statistics				
	TOR	BOS	NYK	BKN	PHI	TOR	BOS	NYK	BKN	PHI
TOR		-0.064	-0.141	0.001	-0.002		-0.773	-1.196	0.051	-0.165
BOS			-0.319	-0.298	-0.128			-2.516	-2.939	-1.344
NYK				-0.121	-0.038				-1.187	-0.681
BKN					-0.147					-1.717
PHI										

	Central Division					t-statistics				
	CLE	IND	DET	CHI	MIL	CLE	IND	DET	CHI	MIL
CLE		-0.094	-0.239	-0.204	-0.575		-0.777	-1.862	-2.115	-3.386
IND			-0.174	-0.052	-0.030			-1.327	-0.496	-0.621
DET				-0.067	-0.483				-0.628	-2.876
CHI					-0.239					-1.449
MIL										

	Southeast Division					t-statistics				
	MIA	ATL	CHA	WAS	ORL	MIA	ATL	CHA	WAS	ORL
MIA		0.002	-0.006	-0.030	-0.040		0.101	-0.235	-0.658	-0.465
ATL			-0.302	-0.223	-0.105			-1.983	-2.186	-1.011
CHA				-0.002	-0.468				-0.126	-2.872
WAS					-0.049					-0.567
ORL										

West Conference

	Northwest Division					t-statistics				
	OKC	POR	UTA	DEN	MIN	OKC	POR	UTA	DEN	MIN
OKC		0.002	-0.116	-0.072	-0.017		0.384	-1.196	-0.906	-0.378
POR			-0.045	-0.033	-0.138			-0.817	-0.535	-1.115
UTA				-0.320	0.003				-2.585	0.121
DEN					-0.013					-0.228
MIN										

	Pacific Division					t-statistics				
	GSW	LAC	SAC	PHX	LAL	GSW	LAC	SAC	PHX	LAL
GSW		0.001	0.001	-0.204	-0.096		0.225	0.036	-1.387	-0.675
LAC			-0.034	-0.007	-0.050			-0.427	-0.316	-0.438
SAC				-0.254	-0.125				-2.144	-1.338
PHX					-0.206					-1.037
LAL										

	Southwest Division					t-statistics				
	SAS	DAL	MEM	HOU	NOP	SAS	DAL	MEM	HOU	NOP
SAS		-0.140	-0.002	-0.036	-0.118		-1.572	-0.098	-0.655	-1.100
DAL			-0.006	-0.155	-0.141			-0.208	-1.784	-1.552
MEM				0.004	0.001				0.160	0.140
HOU					-0.199					-1.830
NOP										

The t-statistics are computed using the Delta method.

Table 11: Outcome Function Estimates

	Parameter Estimate			t-statistics		
Team	Exper.	Fatigue	σ_r^2	Exper.	Fatigue	σ_r^2
TOR	0.0000	0.0002	0.0007	0.4322	2.1627	60.0333
BOS	0.0002	0.0000	0.0007	1.9461	0.1305	61.5467
NYK	0.0001	0.0001	0.0007	1.0801	1.5110	60.7783
BKN	0.0001	-0.0001	0.0007	0.7985	-1.3967	60.6300
PHI	0.0001	0.0001	0.0007	0.7050	1.3797	64.2028
CLE	-0.0002	0.0004	0.0007	-2.1448	3.9403	61.3677
IND	0.0001	0.0001	0.0007	0.7459	0.8906	60.6465
DET	-0.0001	0.0002	0.0007	-0.8755	2.6751	59.2284
CHI	0.0001	0.0001	0.0006	0.7086	1.1942	59.8832
MIL	0.0001	0.0000	0.0006	1.4290	-0.3539	59.7662
MIA	0.0000	0.0001	0.0006	0.0699	1.2007	60.7783
ATL	0.0003	-0.0001	0.0007	2.9096	-0.7113	63.5767
CHA	0.0000	0.0001	0.0007	0.2671	1.3312	59.0085
WAS	0.0000	0.0000	0.0007	-0.1675	0.4690	61.8708
ORL	0.0001	0.0000	0.0007	0.9473	-0.4786	60.2495
OKC	-0.0001	0.0002	0.0007	-0.9113	2.2144	60.9426
POR	0.0001	0.0001	0.0007	0.5428	1.0831	60.6465
UTA	0.0001	0.0003	0.0007	1.0949	2.9773	63.7966
DEN	0.0001	0.0001	0.0007	1.4732	1.1264	60.7124
MIN	0.0000	0.0001	0.0007	0.4925	1.4218	61.4980
GSW	0.0001	0.0000	0.0007	0.9787	0.4454	61.5142
LAC	-0.0002	0.0003	0.0007	-2.0434	2.8251	61.0246
SAC	0.0000	0.0001	0.0007	0.0606	0.8664	62.5620
PHX	-0.0001	0.0000	0.0007	-1.0618	0.0162	60.3490
LAL	0.0000	0.0001	0.0007	0.3291	1.5397	59.1946
SAS	-0.0001	0.0002	0.0007	-0.6380	2.3176	61.2862
DAL	0.0001	-0.0001	0.0007	1.6625	-0.8558	62.7375
MEM	0.0000	0.0001	0.0007	-0.4087	0.9881	60.9098
HOU	0.0000	0.0002	0.0007	0.5125	2.2838	62.6738
NOP	0.0001	0.0001	0.0006	1.2368	1.7580	61.2372

Exper. is Experience.
Asymptotic standard errors and t-statistics.

Table 12: Outcome Function Estimates (Network-Level): Lee (1983) Parametric Approach

	Random Forest								Multinomial Logit							
	Parameter Estimates				t-statistics				Parameter Estimates				t-statistics			
Team	Home	Half	RPI	SB	Home	Half	RPI	SB	Home	Half	RPI	SB	Home	Half	RPI	SB
TOR	0.0000	-0.0002	-0.0016	0.0002	0.0938	-0.4976	-0.1709	0.3155	0.0001	-0.0003	-0.0016	0.0003	0.1063	-0.5431	-0.1700	0.6110
BOS	0.0003	-0.0025	-0.0176	-0.0005	0.5481	-4.0623	-1.7804	-0.6415	0.0003	-0.0025	-0.0175	-0.0005	0.5355	-4.0573	-1.7703	-0.9455
NYK	-0.0003	-0.0021	-0.0020	0.0001	-0.4591	-3.7866	-0.1946	0.0905	-0.0003	-0.0021	-0.0020	0.0000	-0.4628	-3.7699	-0.1935	-0.0192
BKN	0.0009	-0.0005	-0.0134	-0.0014	1.4976	-0.8796	-1.3412	-1.7153	0.0010	-0.0005	-0.0134	-0.0009	1.5140	-0.8834	-1.3369	-1.3758
PHI	0.0001	-0.0004	-0.0122	0.0003	0.1985	-0.7865	-0.8640	0.3908	0.0001	-0.0004	-0.0123	0.0004	0.1957	-0.7942	-0.8717	0.6255
CLE	0.0002	0.0022	-0.0153	0.0010	0.3476	3.6505	-1.5563	1.2754	0.0002	0.0022	-0.0150	0.0005	0.3453	3.5973	-1.5264	0.8906
IND	0.0004	-0.0011	-0.0049	-0.0012	0.6670	-1.9592	-0.5085	-0.5567	0.0004	-0.0011	-0.0049	-0.0006	0.6323	-1.9770	-0.5014	-1.1669
DET	0.0011	0.0011	-0.0226	0.0001	2.1510	1.7195	-2.3251	0.0921	0.0011	0.0011	-0.0226	0.0002	2.1476	1.7657	-2.3252	0.2946
CHI	0.0003	-0.0007	-0.0370	0.0012	0.5459	-1.1011	-2.8896	1.1375	0.0003	-0.0006	-0.0370	0.0005	0.4537	-1.0449	-2.8882	0.7572
MIL	0.0011	-0.0011	0.0043	0.0000	1.7120	-1.7102	0.4242	-0.0314	0.0011	-0.0011	0.0043	-0.0001	1.7147	-1.7076	0.4257	-0.1289
MIA	0.0005	-0.0002	-0.0214	0.0004	0.8208	-0.3501	-2.0987	0.1721	0.0005	-0.0002	-0.0215	0.0002	0.8314	-0.3462	-2.1085	0.3623
ATL	0.0001	-0.0027	-0.0361	-0.0018	0.0924	-4.5454	-3.8706	-1.1457	0.0000	-0.0027	-0.0361	-0.0002	0.0733	-4.5916	-3.8689	-0.4334
CHA	0.0012	-0.0005	-0.0075	-0.0009	2.0826	-0.8465	-0.9330	-0.4459	0.0012	-0.0005	-0.0074	-0.0004	2.0879	-0.8611	-0.9326	-0.8246
WAS	0.0007	0.0001	-0.0213	-0.0009	1.1842	0.1053	-1.8594	-1.1590	0.0007	0.0001	-0.0211	-0.0008	1.2088	0.1241	-1.8509	-1.6182
ORL	0.0001	-0.0016	-0.0011	-0.0002	0.2044	-2.7118	-0.0998	-0.2885	0.0001	-0.0016	-0.0012	-0.0001	0.2051	-2.7117	-0.1017	-0.1635
OKC	0.0002	0.0006	-0.0120	-0.0005	0.3707	0.9853	-1.1747	-0.5338	0.0002	0.0006	-0.0120	-0.0005	0.3600	0.9748	-1.1687	-0.8969
POR	0.0007	-0.0008	0.0018	-0.0002	1.4196	-1.4378	0.1697	-0.1983	0.0007	-0.0008	0.0018	0.0000	1.4205	-1.4205	0.1693	0.0456
UTA	0.0009	-0.0009	-0.0052	-0.0003	1.6589	-1.6497	-0.6611	-0.3410	0.0009	-0.0009	-0.0052	-0.0002	1.6639	-1.6474	-0.6594	-0.3878
DEN	0.0005	-0.0024	-0.0036	0.0012	0.8774	-4.2105	-0.3757	1.4980	0.0005	-0.0023	-0.0038	0.0007	0.8521	-4.1666	-0.3882	1.3846
MIN	0.0000	-0.0011	0.0069	-0.0002	-0.0496	-1.9122	0.6609	-0.2508	0.0000	-0.0011	0.0070	0.0001	-0.0541	-1.8840	0.6710	0.1472
GSW	0.0008	-0.0015	-0.0030	-0.0007	1.3449	-2.5837	-0.2933	-0.9381	0.0008	-0.0015	-0.0031	-0.0009	1.3142	-2.5912	-0.3013	-1.3852
LAC	0.0002	0.0023	-0.0064	0.0020	0.3681	3.8951	-0.5227	2.3730	0.0002	0.0023	-0.0061	0.0017	0.3641	3.9584	-0.5023	2.9547
SAC	0.0004	-0.0004	-0.0129	-0.0004	0.8405	-0.7953	-1.1531	-0.5442	0.0004	-0.0004	-0.0130	-0.0003	0.8383	-0.8174	-1.1546	-0.5838
PHX	0.0007	0.0019	-0.0034	0.0008	1.0747	3.3409	-0.2891	0.8144	0.0007	0.0019	-0.0034	0.0006	1.0716	3.3599	-0.2900	0.9725
LAL	0.0007	0.0002	-0.0204	-0.0027	1.0083	0.3828	-2.1520	-2.3189	0.0006	0.0002	-0.0204	-0.0016	0.9651	0.3488	-2.1594	-1.7254
SAS	0.0006	0.0009	-0.0342	-0.0013	1.0194	1.6524	-3.9568	-0.6696	0.0006	0.0008	-0.0335	-0.0009	1.0701	1.4917	-3.8256	-1.4721
DAL	0.0007	-0.0015	-0.0096	-0.0005	1.3417	-2.9593	-1.0208	-0.2442	0.0007	-0.0016	-0.0097	0.0001	1.3210	-3.0266	-1.0263	0.2342
MEM	0.0008	-0.0004	-0.0176	-0.0008	1.4729	-0.8341	-1.7251	-0.9201	0.0008	-0.0004	-0.0175	-0.0005	1.4781	-0.8326	-1.7222	-0.8356
HOU	0.0010	0.0004	-0.0130	0.0005	1.5626	0.5905	-1.2372	0.2069	0.0010	0.0004	-0.0130	0.0002	1.5523	0.5978	-1.2362	0.3927
NOP	0.0011	-0.0004	-0.0222	-0.0002	2.0907	-0.8260	-2.4852	-0.2451	0.0011	-0.0004	-0.0223	-0.0003	2.1008	-0.8231	-2.4999	-0.5976

SB: Selection Bias Coefficient = σ_{12}

Bootstrap standard errors and t-statistics.

Table 13: Outcome Function Estimates (Network-Level): Dahl (2002) Semi-Parametric Approach

	Random Forest						Multinomial Logit					
	Parameter Estimates			t-statistics			Parameter Estimates			t-statistics		
Team	Home	Half	RPI	Home	Half	RPI	Home	Half	RPI	Home	Half	RPI
TOR	0.0001	-0.0003	-0.0016	0.0952	-0.4881	-0.1691	0.0001	-0.0003	-0.0016	0.1144	-0.4697	-0.1651
BOS	0.0003	-0.0025	-0.0176	0.5516	-4.0062	-1.7817	0.0003	-0.0025	-0.0175	0.5388	-3.9773	-1.7589
NYK	-0.0003	-0.0021	-0.0020	-0.4599	-3.5814	-0.1941	-0.0003	-0.0020	-0.0019	-0.4724	-2.5579	-0.1902
BKN	0.0009	-0.0005	-0.0134	1.4782	-0.8570	-1.3371	0.0009	-0.0005	-0.0134	1.4738	-0.8823	-1.3204
PHI	0.0002	-0.0004	-0.0114	0.2669	-0.6814	-0.8005	0.0002	-0.0004	-0.0118	0.2674	-0.6757	-0.8264
CLE	0.0002	0.0023	-0.0143	0.3559	3.6716	-1.4480	0.0002	0.0020	-0.0155	0.3178	2.8157	-1.5774
IND	0.0003	-0.0012	-0.0050	0.6166	-2.1052	-0.5114	0.0005	-0.0011	-0.0047	0.8011	-1.7782	-0.4850
DET	0.0011	0.0011	-0.0228	2.0616	1.7453	-2.3415	0.0011	0.0009	-0.0224	2.0278	1.4673	-2.3043
CHI	0.0003	-0.0007	-0.0371	0.5362	-1.1239	-2.8910	0.0004	-0.0007	-0.0359	0.7090	-1.0655	-2.7865
MIL	0.0011	-0.0011	0.0044	1.7108	-1.7047	0.4268	0.0011	-0.0011	0.0044	1.6884	-1.6991	0.4319
MIA	0.0005	-0.0002	-0.0214	0.8189	-0.3470	-2.0953	0.0005	-0.0002	-0.0214	0.7845	-0.2801	-2.1017
ATL	0.0001	-0.0026	-0.0355	0.0981	-4.3300	-3.7971	0.0000	-0.0027	-0.0361	0.0819	-4.5506	-3.8924
CHA	0.0012	-0.0005	-0.0075	2.0737	-0.8479	-0.9379	0.0012	-0.0005	-0.0074	2.0724	-0.7985	-0.9334
WAS	0.0007	0.0001	-0.0213	1.1771	0.1255	-1.8571	0.0007	0.0003	-0.0197	1.2687	0.4984	-1.7179
ORL	0.0001	-0.0017	-0.0009	0.1996	-2.7319	-0.0823	0.0001	-0.0016	-0.0011	0.1987	-2.6877	-0.0923
OKC	0.0003	0.0006	-0.0122	0.4932	0.8934	-1.1982	0.0002	0.0006	-0.0119	0.3711	0.9487	-1.1716
POR	0.0007	-0.0009	0.0027	1.3850	-1.5655	0.2512	0.0007	-0.0008	0.0018	1.4068	-1.4104	0.1693
UTA	0.0009	-0.0009	-0.0053	1.6466	-1.6512	-0.6665	0.0009	-0.0009	-0.0052	1.6601	-1.6156	-0.6601
DEN	0.0005	-0.0024	-0.0036	0.8727	-4.0382	-0.3739	0.0005	-0.0023	-0.0036	0.8582	-4.1444	-0.3715
MIN	0.0000	-0.0011	0.0069	-0.0493	-1.8248	0.6582	0.0000	-0.0011	0.0069	-0.0530	-1.8358	0.6690
GSW	0.0008	-0.0015	-0.0031	1.3367	-2.4139	-0.2950	0.0008	-0.0015	-0.0031	1.3179	-2.3322	-0.2942
LAC	0.0002	0.0023	-0.0065	0.3680	3.6931	-0.5288	0.0002	0.0023	-0.0062	0.3838	3.7818	-0.5078
SAC	0.0004	-0.0005	-0.0129	0.8699	-0.8803	-1.1506	0.0003	-0.0009	-0.0135	0.6069	-1.5301	-1.1918
PHX	0.0007	0.0019	-0.0034	1.0887	3.3329	-0.2909	0.0006	0.0016	-0.0013	0.9074	2.5161	-0.1132
LAL	0.0007	0.0002	-0.0203	0.9952	0.3652	-2.1370	0.0006	0.0002	-0.0205	0.9509	0.3316	-2.1368
SAS	0.0006	0.0009	-0.0342	1.0150	1.6469	-3.9480	0.0006	0.0009	-0.0347	1.0274	1.5818	-3.9721
DAL	0.0007	-0.0015	-0.0096	1.3436	-2.9338	-1.0233	0.0007	-0.0015	-0.0097	1.3232	-2.8085	-1.0123
MEM	0.0008	-0.0004	-0.0175	1.4684	-0.8341	-1.7207	0.0008	-0.0004	-0.0175	1.4825	-0.8136	-1.7272
HOU	0.0010	0.0004	-0.0130	1.5623	0.5864	-1.2369	0.0010	0.0004	-0.0130	1.5495	0.5763	-1.2314
NOP	0.0011	-0.0003	-0.0226	2.1017	-0.6382	-2.5118	0.0011	-0.0004	-0.0223	2.1106	-0.7952	-2.4880

Bootstrap standard errors and t-statistics.