American Economic Journal: Applied Economics 2009, 1:4, 34–68
http://www.aeaweb.org/articles.php?doi=10.1257/app.1.4.34

Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments[†]

By Jonathan Guryan, Kory Kroft, and Matthew J. Notowidigdo*

This paper uses random assignment in professional golf tournaments to test for peer effects in the workplace. We find no evidence that playing partners' ability affects performance, contrary to recent evidence on peer effects in the workplace from laboratory experiments, grocery scanners, and soft fruit pickers. In our preferred specification, we can rule out peer effects larger than 0.043 *strokes for a one stroke increase in playing partners' ability. Our results complement existing studies on workplace peer effects and are useful in explaining how social effects vary across labor markets, across individuals, and with the form of incentives faced.* (*JEL* D83, J44, L83)

Is an employee's productivity influenced by the productivity of his or her nearby co-workers? The answer to this question is important for the optimal organization of labor in a workplace and for the optimal design of incentives.[1,2] Despite their importance, however, peer effects questions like this one are notoriously difficult to

* Guryan: University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637 and National Bureau of Economic Research (e-mail: jguryan@chicagobooth.edu); Kroft: UC-Berkeley Department of Economics, 195 Madison Avenue, Apartment D, Toronto, ON M5R 2S6 (e-mail: kroft@econ.berkeley.edu); Notowidigdo: Massachusetts Institute of Technology, Department of Economics, 50 Memorial Drive, E52-204F, Cambridge, MA 02142 (e-mail: noto@mit.edu). The authors thank Daron Acemoglu, David Autor, Oriana Bandiera, Marianne Bertrand, David Card, Kerwin Charles, Ken Chay, Stefano DellaVigna, Amy Finkelstein, Alex Mas, Sean May, Steve Pischke, Imran Rasul, Jesse Rothstein, and Emmanual Saez for helpful conversations regarding this paper. The authors also thank Phil Wengerd for outstanding research assistance. Guryan thanks the University of Chicago Booth School of Business and the Industrial Relations Section at Princeton University for funding support. Kroft thanks the Center for Labor Economics and The Institute of Business and Economic Research at Berkeley for funding support. Notowidigdo thanks the MIT Department of Economics for funding support. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research. This research was funded, in part, by the George J. Stigler Center for the Study of the Economy and the State at the University of Chicago Booth School of Business.

† To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at: http://www.aeaweb.org/articles.php?doi=10.1257/app.1.4.34.

[1] The relevance for the optimal design of incentives hinges on whether social effects are complements or substitutes for financial incentives.

[2] Complementarities between an employee's productivity and the productivity of his peers may arise for at least three reasons: (a) individuals may learn from their co-workers about how best to perform a given task, (b) workers may be motivated to exert effort when they see their co-workers working hard, performing well, or when they know their co-workers are watching, or (c) the nature of the production process may be such that the productivity of one worker mechanically influences the productivity of another worker directly (e.g., the assembly line). While the former two reasons are "behavioral," the latter is a mechanical effect that arises for purely structural or technological reasons. For clarity, we label the first two of these effects "peer effects" and the last effect a "production complementarity."

answer empirically. In this paper, we exploit the conditional random assignment of golfers to playing partners in professional golf tournaments to identify peer effects among co-workers in a high-skill professional labor market.[3]

A few other recent studies have found evidence that peer effects can be important in several work settings: among low-wage workers at a grocery store (Alexandre Mas and Enrico Moretti 2009), among soft fruit pickers (Oriana Bandiera, Iwan Barankay, and Imran Rasul 2009), and among workers performing a simple task in a laboratory setting (Armin Falk and Andrea Ichino 2006). However, the two field studies (Mas and Moretti 2009; and Bandiera, Barankay, and Rasul 2009) rely on observational variation in peers, while the Falk and Ichino (2006) study is of the behavior of high school students in a laboratory. Ours is the first field study of peer effects in the workplace to exploit random assignment of peers. It is also the first to study a high-skill professional labor market. As such, a comparison of results across markets speaks to heterogeneity in the importance of peer effects that is missed by looking at one particular type of labor market.

Random assignment is an important attribute to a peer effects research design (see e.g., Bruce Sacerdote 2001 and David J. Zimmerman 2003), but it is not a panacea. Shocks common to a randomly assigned peer group still make causal inference difficult. As we describe below, our design allows a convenient way to test and control for such common shocks (e.g., from common weather shocks that vary over the course of the day) by examining golfers who play nearby on the golf course and in close temporal proximity, but in separate groups with whom there are no social interactions.

The main result of this paper is that neither the ability nor the current performance of playing partners affect the performance of professional golfers. Output in golf is measured in strokes, the number of times a player hits the ball in his attempt to get it into the hole. In our preferred specification, we can rule out peer effects larger than 0.043 strokes for a 1 stroke increase in playing partners' ability, and our point estimate is actually negative. The results are robust to alternative peer effects specifications, and we rule out multiple forms of peer effect mechanisms including learning and motivation.

We also point out a bias that is inherent in typical tests for random assignment of peers and describe a simple correction. Because individuals cannot be their own peers, even random assignment generates a negative correlation in pre-determined characteristics of peers. Intuitively, the urn from which the peers of an individual are drawn does not include the individual. Thus, the population at risk to be peers with high-ability individuals is, on average, lower ability than the population at risk to be peers with low-ability individuals. As a result, the typical test for random assignment, a regression of *i*'s pre-determined characteristic on the mean characteristic of

[3] See Charles F. Manski (1993) and Robert Moffitt (2001) for descriptions of the problems associated with estimating peer effects. In general, it is difficult to disentangle whether an observed correlation is the effect of the group's behavior on an individual's behavior ("endogenous effects"), the effect of the group's characteristics on an individual's behavior ("contextual effects"), or the correlation between observed and unobserved determinants of the outcome ("correlated effects"). An additional difficulty is what Manski calls the "reflection problem." To understand the reflection problem, consider trying to estimate the effect of individual *i*'s behavior on individual *j*'s behavior. It is very difficult to tell which member of the pair is affecting the other's behavior, and whether the affected behavior by one member of the pair affects the other's in turn, and so on.

i's peers, produces a slightly negative coefficient even when peers are truly randomly assigned. This bias can cause researchers to infer that peers are randomly assigned when, in fact, there is positive matching. We present results from Monte Carlo simulations showing that this bias can be reasonably large, and that it is decreasing in the size of the population from which peers are selected. We then propose a simple solution to this problem—controlling for the average ability of the population at risk to be the individual's peers—and show that including this additional regressor produces test statistics that are well-behaved.

Our results compare the performance of professional golfers who are randomly grouped with playing partners of differing abilities. Within a playing group, players are proximate to one another and can therefore observe each others' shots and scores. This proximity creates the opportunity to learn from, and be motivated by, peers.

There are several learning opportunities during a round. For example, a player must judge the direction of the wind when hitting his approach shot to the putting green. Wind introduces uncertainty into shot and club selection. Thus, by observing the ball flight of others in the playing group, a player can reduce this uncertainty and increase his chance of hitting a successful shot. Another example is the putting green. Subtle slopes, moisture, and the type of grass all affect the direction and speed of a putt. The chance to learn how a skilled putter manages these conditions may confer an advantage to a peer golfer.[4]

Turning to motivation, there are several ways that motivation can affect performance. The chance to visualize a good shot may help a player to execute his own shot successfully. Similarly, seeing a competitor play well may directly motivate a player and help him to focus his mental attention on the task at hand. Still more, a player's self-confidence, and, in turn, his performance (Roland Benabou and Jean Tirole 2002), may be directly affected by the abilities or play of his peers (Leon Festinger 1954). Interestingly, there is a popular perception among PGA golfers that these psychological effects may facilitate play.[5]

In addition to a simple overall measure of skill, we present results based on multidimensional measures of ability that line up nicely with important potential underlying mechanisms of peer effects. We discuss why these data allow us to distinguish between learning effects and motivational effects, the latter of which are more psychological in nature. The results show no evidence of peer effects of either the learning or motivational kind.

[4] The prescribed order of play, described in Section IIA, clearly affects the learning opportunities available in a golf tournament. Better golfers tend to shoot first on the initial shot of a hole, but tend to follow in order on later shots. This variety may affect the extent to which learning occurs, though we suspect such a series of learning opportunities, in which the roles of actor and observer are continually exchanged, is typical of many workplaces. Furthermore, the variation induced by the order of play rules might be used to generate evidence of learning from peers. For example, if one could collect shot-by-shot data it would be interesting to see if behavior is affected by players who shoot immediately prior.

[5] For example, Darren Clarke stated that it is easier to play better when everyone in the group is playing better because "you see good shots go into the green all the time and that makes it a lot easier to do the same yourself" (www.pga.com/pgachampionship/2004/news_interviews_081304_clarke.html) (accessed August 27, 2009). Similarly, Billy Andrade was quoted saying "you kind of feed off each other and that's what we did" (i.pga.com/pga/images/events/2006/pgachampionship/pdf/20060818_andrade.pdf) (accessed August 27, 2009).

Golf tournaments are well designed to identify peer effects for two other reasons. First, our results are purged of relative incentive effects because the objective for each player in a golf tournament is to score the lowest, regardless of with whom that player is playing.[6] Pay is based on relative performance, but performance is compared to the entire field of entrants in a tournament, not relative to the players within a playing group. This contrasts with other settings, such as classrooms, where the performance of an individual is assessed relative to the individual's peers. For example, in a classroom, where grades are based on relative performance, we would expect to see students try hard to perform better than their peers. Such behavior is the result of response to incentives, not of learning from peers.[7]

And second, we are able to identify peer effects in a setting devoid of most production-technology complementarities, i.e., cross-productivity effects that stem from the production function rather than from social or behavioral effects. This is unlike, as one example, the grocery store setting considered in Mas and Moretti (2009). As they mention, in the supermarket checkout setting there is a shared resource, a bagger who helps some checkout workers place items in customers' bags. If the bagger spends more time helping the fastest checkout worker, this will negatively affect the productivity of others in his shift.

The remainder of this paper is organized as follows. Section I reviews the literature. Section II briefly describes the PGA Tour and our data. Section III describes the methodological point concerning bias in typical peer effect randomization tests and shows results verifying the random assignment mechanism. Section IV discusses our empirical approach. Section V discusses the validity of our empirical strategy and our results, and Section VI concludes.

I. Related Literature

It has long been recognized by psychologists that an individual's performance might be influenced by his peers. The first study to show evidence of such peer effects was done by Norman Triplett (1898), who noted that cyclists raced faster when they were pitted against one another, and slower when they raced only against a clock. While Triplett's study shows that the presence of others can *facilitate* performance, others found that the presence of others can *inhibit* performance. In particular, Floyd Henry Allport (1924) found that people in a group setting wrote more refutations of a logical argument, but that the quality of the work was lower than when they worked alone. Similarly, Joseph Pessin (1933) found that the presence

[6] This objective is accurate for almost every player, except, perhaps, the players who have a reasonable probability of earning the top few prizes. In those cases, the objective is to score lower than your opponents (where your opponents are the small universe of players who are competing with you for the top prizes). With this concern in mind, as a robustness check, we confirmed that dropping top-tier players does not change our main results. The other reason players might care about the performance of players in their group is that these are the scores of their competitors that they most easily observe. However, the scores of the current tournament leaders are typically posted around the course so that golfers have information about how they are playing relative to players outside of their group. This information makes the information about how their playing partners are performing significantly less valuable.

[7] Such relative performance effects might even show up in higher scores on statewide standardized tests, which are not graded on a relative scale, because the effort students put forth to earn a better grade than their classroom peer may lead to real learning.

of a spectator reduced individual performance on a memory task. Robert B. Zajonc (1965) resolved these paradoxical findings by pointing out that the task in these experimental setups varied in a way that confounded the results. In particular, he argued that for well-learned or innate tasks, the presence of others improves performance. For complex tasks, however, he argued that the presence of others worsens performance.

Guided by the intuition that peers may affect behavior and hence market outcomes, several economic studies of peer effects have recently emerged in a variety of domains. Examples include education (Bryan S. Graham 2008), crime (Edward L. Glaeser, Bruce Sacerdote, and Jose A. Scheinkman 1996), unemployment insurance take-up (Kroft 2008), welfare participation (Marianne Bertrand, Erzo F. P. Luttmer, and Sendhil Mullainathan 2000), and retirement planning (Esther Duflo and Emmanuel Saez 2003). The remainder of this section reviews three papers that are conceptually most similar to our research, in that they attempt to measure peer effects in the workplace or in a work-like task.

The first economic study of peer effects in a work-like setting is a laboratory experiment conducted by Falk and Ichino (2006). This experiment measures how an individual's productivity is influenced by the presence of another individual working on the same task (stuffing letters into envelopes). They find moderate and significant peer effects. A 10 percent increase in peers' output increases a given individual's effort by 1.4 percent. A criticism of this study is one that applies broadly to other studies in the lab. In particular, that it may have low external validity because of experimenter demand effects or because experimental subjects get paid minimal fees to participate and, as a result, their incentives may be weak (Steven D. Levitt and John A. List 2007).[8]

Two recent studies of peer effects in the workplace have examined data collected from the field. Mas and Moretti (2009) measure peer effects directly in the workplace using grocery scanner data. There is not explicit randomization, but the authors present evidence that the assignment of workers to shifts appears haphazard. Since grocery store managers do not measure individual output directly in a team production setting, one might expect to observe significant free riding and suboptimal effort. Instead, this study finds evidence of significant peer effects with magnitudes similar to those found by Falk and Ichino (2006). A 10 percent increase in the average permanent productivity of co-workers increases a given worker's effort by 1.7 percent. As Mas and Moretti (2009) discuss in their paper, grocery scanner is an occupation where compensation is not very responsive to changes in individual effort and output. They conjecture that "economic incentives alone may not be enough to explain what motivates [a] worker to exert effort in these jobs."

Bandiera, Barankay, and Rasul (2009) examine how the identity and skills of nearby workers affect the productivity of soft fruit pickers on a farm. Assignment of workers to rows of fruit is made by a combination of managers. Though assignment is not explicitly random, the authors present evidence to support the claim that it is orthogonal to worker productivity. The authors find that productivity responds to

[8] Another concern about laboratory experiments is that subjects are prevented from sorting into environments based on their social preferences (Edward Lazear, Ulrike Malmendier, and Roberto Weber 2006).

the presence of a friend working nearby, but do not show evidence that productivity responds to the skill-level of non-friend co-workers. High-skill workers slow down when working next to a less productive friend, and low-skill workers speed up when working next to a more productive friend. Workers are paid piece rates in this setting and are willing to forgo income to conform to a social norm along with friends. The authors also find that relatively low-skill workers respond to the cost of conforming. On high-yield days, fruit pickers who are less productive than their friends work harder.

In light of the fact that monetary incentives are weaker in the Mas and Moretti (2009) study than in Bandiera, Barankay, and Rasul (2009), it is interesting to note that Mas and Moretti (2009) find more general peer effects. A recent paper by Thomas Lemieux, W. Bentley MacLeod, and Daniel Parent (2009) points out that an increasing fraction of US jobs contain some type of performance pay based either on a commission, a bonus, or a piece rate.[9] Therefore, it is natural to ask whether peer effects also exist in settings with stronger incentives than either of the aforementioned peer effects studies.[10] In the setting that we consider—professional golf tournaments—compensation is determined purely by performance.[11] Pay is high and the pay structure is quite convex. For example, during the 2006 PGA season, the top prize money earner, Tiger Woods, earned almost $10 million, and 93 golfers earned in excess of $1 million during the season. We discuss the convexity of tournament payouts later in the paper.

The existing field studies in the literature on peer effects in the workplace have focused on low-skill jobs in particular industries. It is possible that there is heterogeneity in how susceptible individuals are to social effects at work. Motivated by this, we ask whether peer effects exist in workplaces made up of highly skilled professional workers.[12]

In terms of research design, our study is related to Sacerdote (2001) and Zimmerman (2003), who measure peer effects in higher education using the random assignment of dorm roommates at Dartmouth College and Williams College, respectively. Sacerdote finds that an increase in roommate's GPA by 1 point increases own freshman year GPA by 0.120. The coefficient on roommate's GPA, however, drops significantly—from 0.120 to 0.068—when dorm fixed effects are included, suggesting that common shocks might be driving some of the correlation in GPAs between roommates. As we discuss, we are able to test directly for the most likely source of common shocks to golfers.[13]

[9] Using the Panel Study of Income Dynamics (PSID), the authors find that the overall incidence of performance pay was a little more than 30 percent in the late 1970s but grew to over 40 percent by the late 1990s.

[10] In a prior paper, Bandiera, Barankay, and Rasul (2005) find an increase in effort in response to a switch in compensation regime from relative pay to piece-rate pay. Their evidence suggests that social preferences can be offset by appropriate monetary incentives.

[11] Golfers also receive a sizeable amount from professional endorsements. Presumably, endorsement earnings are related indirectly to performance. Tiger Woods, the golfer with the highest earnings from tournaments is also the golfer with the greatest endorsement income.

[12] In Section VF, we will also look at whether there are heterogeneous peer effects within this class of professional workers.

[13] Finally, other studies have used professional golf data to test economic theories. Jennifer Brown (2008) concludes that overall performance of top competitors is worse in tournaments in which Tiger Woods competes because golfers infer that the chance of earning the top prize declines. Our estimates net out this effect because we

II. Background and Data

A. *Institutional Details of the PGA Tour*

Relevant Rules.—[14] There is a prescribed set of rules to determine the order of play. The player with the best (i.e., lowest) score from the previous hole takes his initial shot first, followed by the player with the next best score from the previous hole. After that, the player who is farthest from the hole always shoots next.

Selection.—Golfers from all over the world participate in PGA tournaments that are held (almost) every week.[15] At the end of each season, the top 125 earners in PGA Tour events are made full-time members of the PGA Tour for the following year. Those not in the top 125, and anyone else who wants to become a full-time member of the PGA Tour, must go to "Qualifying School," where there are a limited number of spots available to the top finishers. For most PGA Tour tournaments, the players have the right, but not the obligation, to participate in the tournament. In practice, there is variation in the fraction of tournaments played by PGA Tour players. The median player plays in about 59 percent of a season's tournaments.[16] Some players avoid tournaments that do not have a large enough purse or high prestige, while other players might avoid tournaments that require a substantial amount of travel. Stephen Bronars and Gerald Oettinger (2008) look directly at several determinants of selection into golf tournaments, and they find a substantial effect of the level of the purse on entry decisions. Because membership on the tour for the following year is based on earnings, lower earning players have an incentive to enter tournaments that higher skill players choose not to enter. Conditional on the set of players who enter a tournament, playing partners are randomly assigned within categories (which we define and describe below). Unconditional on this fully interacted set of fixed effects, assignment is not random.

Assignment Rule.—Players are continuously assigned to one of four categories according to rules described in detail in Appendix A. Players in category 1 are typically tournament winners from the current or previous year or players in the top 25 on the earnings list from the previous year. These include players such as Tiger

condition on tournament-by-category. Ronald G. Ehrenberg and Michael L. Bognanno (1990) test whether professional golf tournaments elicit effort responses. They find that the level and structure of prizes in PGA tournaments influence players' performance. However, Jonathan M. Orszag (1994) shows their results might not be robust once weather shocks are accounted for.

[14] The Web Appendix gives a short introduction to the basic rules of golf.

[15] In 2007, there were 47 PGA tournaments played over 44 weeks. The three instances in which there are two tournaments during a week are the "major championships." Because these championships are only for qualifying golfers from around the world (and not exclusively for PGA Tour members), and because these tournaments are not sponsored by the PGA Tour, the PGA Tour also hosts tournaments during the major championships for the remaining PGA Tour members who did not qualify for the major tournaments. We drop all the major championships from our dataset because they do not use the same random assignment mechanism.

[16] Since most players play in most tournaments, when we construct our measure of ability we will usually have enough past tournament results to reliably estimate the ability of each player. The median player has 30 past tournament results with which to estimate his ability, 29 percent of the players have more than 40 past tournaments, and 14 percent of the players have fewer than 5 tournaments. We weight all regressions by the inverse of the sample variance of each player's estimated ability, except for the randomization tests which are unweighted.

Woods, Phil Mickelson, and Ernie Els. Players in category 1A include previous tournament winners who no longer qualify for category 1 and former major championship winners such as Nick Faldo and John Daly. Players in category 2 are typically those between 26 and 125 on the earnings list from the previous year, and players who have made at least 50 cuts during their career or who are currently ranked among the top 50 on the World Golf Rankings. Category 3 consists of all other entrants in the tournament. Within these categories, tournament directors randomly assign playing partners to groups of three golfers.[17] These groups play together for the first two (of a total of four) rounds of the tournament.

Tournament Details.—Tournaments are generally four rounds of 18 holes played over four days. Prizes are awarded based on the cumulative performance of players over all four rounds. At the end of the second round, there is a cut that eliminates approximately half of the tournament field based on cumulative performance. Most tournaments have 130–160 players, the top 70 (plus ties) of which remain to play the final two rounds. We evaluate performance from the first two rounds only, since in the third and fourth rounds players are assigned to playing partners based on performance of the previous rounds.[18]

Economic Incentives.—A player must survive the second-round cut to qualify to earn prize money.[19] The prize structure is extremely convex. First prize is generally 18 percent of the total purse. Furthermore, the full economic incentives are even stronger than this implies, since better performance generally attracts endorsement compensation. Figure 1, panel A shows the convexity in the prize structure of a typical tournament, and Figure 1, panel B shows the distribution of average earnings in our sample.

B. *Data*

Key Variables.—We collected information on tee times, groupings, results, earnings, course characteristics, and player statistics and characteristics from the PGA Tour Web site and various other Web sites.[20] Most of our data spans the 1999–2006 golf seasons. However, we only have tee times, groupings (i.e., the peer groups), and categories for the 2002, 2005, and 2006 seasons.[21] As we will discuss, to construct

[17] This is similar to the random assignment mechanism used to assign roommates at Dartmouth, as discussed in Sacerdote (2001).

[18] Francis J. Flynn and Emily T. Amanatullah (2008) examine a similar question using data on the third and fourth rounds of a particular major championship golf tournament. They find a positive correlation between golfers' scores and their partner's ability. In the rounds they study, however, peers are assigned based solely on the performance in the first two rounds of the current tournament. As a result, it is hard to distinguish whether this positive correlation is causal or the result of a mechanical relationship induced by the pairing mechanism (i.e., that players who are playing better are, by rule, more likely to be paired with better players).

[19] There is no entry fee for PGA Tour members to play in PGA tournaments. Nonmembers must pay a nominal $400 entry fee.

[20] We gathered data from www.pgatour.com, www.espn.com, www.usatoday.com, www.cnnsi.com, and www.yahoo.com.

[21] Tee times were collected each Thursday during the 2002 season since a historical list was not maintained either on Web sites or by the PGA Tour at that time. The 2005 and 2006 tee times were collected subsequently in an effort to increase sample size and power.

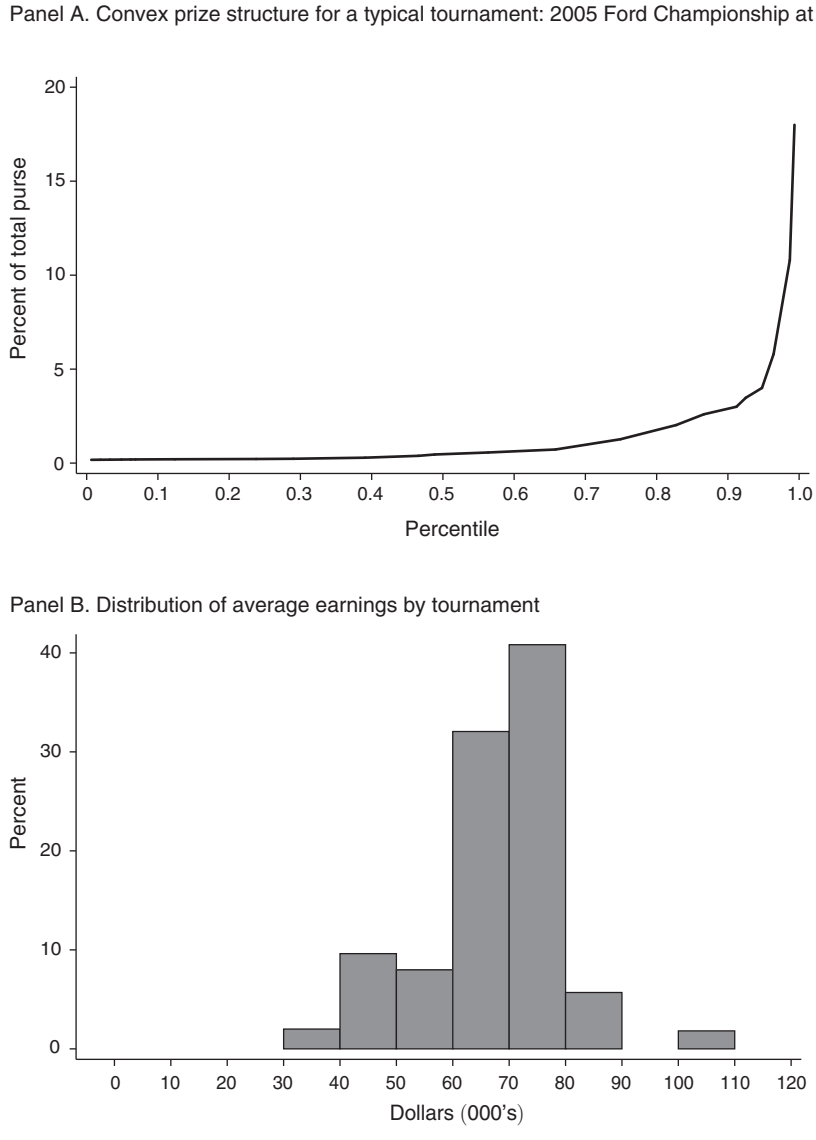

FIGURE 1. TOURNAMENT EARNINGS

Notes: Panel A shows the distribution of earnings within a typical tournament. The convex prize structure implies that the top few percentiles receive a majority of the purse. Panel B shows the distribution, in average earnings, across all tournaments in our sample.

a pre-determined ability measure, we use the 1999, 2000, and 2001 data for the players from the 2002 season, and the 2003 and 2004 data for the players from the 2005 and 2006 seasons. Our data allow us to observe the performance of the same individuals playing in many PGA tournaments over three seasons (2002, 2005, and 2006).

We also collected a rich set of disaggregated player statistics for all years of the sample. These variables were collected in hopes of shedding light on the mechanism through which the peer effects operate. These measures of skill include the average number of putts per round, average driving distance (a measure of driving accuracy), and the average number of greens hit in regulation (the fraction of times a golfer gets the ball onto the green in at least two fewer shots than par). Below, we discuss how these measures might allow us to separately identify two specific forms of peer effects: learning and motivation.

Sample Selection.—As discussed in more detail in Appendix A, the scores from some tournaments must be dropped because random assignment rules are not followed. The four most prestigious tournaments (called "major championships") do not use the same conditional random assignment mechanism. For example, the US Open openly admits to creating "compelling" groups to stimulate television ratings. The vast majority of PGA tournaments, however, use the same random assignment mechanism, and we have confirmed this through personal communications with the PGA Tour, and through statistical tests reported below.

C. *A Measure of Ability*

To estimate peer effects, we require a measure of ability or skill for every player. We construct such a measure based on players' scores in prior years.[22] However, a simple average of scores from prior years understates differences in ability across players because better players tend to self-select into tournaments played on more difficult golf courses.

We address this problem using a simplified form of the official handicap correction used by the United States Golf Association (USGA), the major golf authority that oversees the official rules of the game. For the purpose of computing the handicap correction, the USGA estimates the difficulty of most golf courses in the United States. Using scores of golfers of different skill levels, the USGA assigns each course a *slope* and a *rating*, which are related to the estimated slope and intercept from a regression of score on ability. We normalize the slopes of the courses in our sample so that the average slope is one.[23] We then use the ratings and adjusted slopes of the course to regression-adjust each past score, indexed by n. Specifically, for each past score we compute

$$h_n = (score_n - rating_c)/adjusted\ slope_c,$$

where c indexes golf courses. For each golfer, in each year, we take the average of h_n for the scores from the previous two or three years to be our measure of ability.[24]

[22] Since our peer effects specifications only use the first two rounds of a tournament, we construct our ability measure using only the first two rounds from earlier tournaments.

[23] Specifically, we divide each slope by the average slope in our sample, which is 135.5.

[24] We use scores from the previous three years for 2002 data and scores from the previous two years for 2005 and 2006 data.

This ability measure is essentially an estimate of the number of strokes more than 72 (i.e., above par) that a golfer typically takes in a round, on an average course, that is used for professional golf tournaments. As is true for golf scores generally, higher values are worse. Thus, the measure of ability is positively correlated with a golfer's score.

This correction differs from the official USGA handicap correction in two ways. First, the official correction predicts scores on the average golf course for which the slope is 113, whereas we calibrate to the average course slope in our sample. This adjustment ensures that our measure of ability is in the same units as the dependent variable of our peer effects regressions. Second, the official handicap formula averages the 10 lowest of the last 20 adjusted scores. Because we are interested in predicted performance, we, instead, average over all past scores. We have experimented with several other estimates of ability, including the simple average of scores from the previous two years, the average score from the previous two years after adjusting for course fixed effects and a best linear predictor. Estimates based on these alternative measures of ability yield very similar results.

III. Bias in Typical Tests for Random Assignment of Peers

A. *Explaining the Problem*

Before presenting the empirical results, in this section, we describe an important methodological consideration. Given the importance of random assignment, papers that report estimates of peer effects typically present statistical evidence to buttress the case that assignment of peers is random, or as good as random. The typical test is an OLS regression of individual i's pre-determined characteristic x on the average x of i's peers, conditional on any variable on which randomization was conditioned. The argument is made that if assignment of peers is random, or if selection into peer groups is ignorable, then this regression should yield a coefficient of zero. In our case, this regression would be of the form

$$Ability_{ikt} = \pi_1 + \pi_2 \overline{Ability}_{-i,kt} + \delta_{tc} + \varepsilon_{ikt}, \tag{1}$$

where i indexes players, k indexes (peer) groups, t indexes tournaments,[25] c indexes categories, and δ_{tc} is a fully interacted set of tournament-by-category dummies, including main effects. This is, for example, the test for random assignment reported in Sacerdote (2001).

This test for the random assignment of individuals to groups is not generally well-behaved. The problem stems from the fact that an individual cannot be assigned to himself. In a sense, sampling of peers is done without replacement—the individual himself is removed from the "urn" from which his peers are chosen. As a result, the

[25] For the remainder of this paper, we define a "tournament" to be a tournament-by-year cell, since our dataset has several tournaments that are played in subsequent years. For example, the Ford Championship in 2002 has a separate dummy than the Ford Championship in 2005.

peers for high-ability individuals are chosen from a group with a slightly lower mean ability than the peers for low-ability individuals.

Consider an example in which four individuals are randomly assigned to groups of two. To make the example concrete, let the individuals have pre-determined abilities 1, 2, 3, and 4. If pairs are randomly selected, there are three possible sets of pairs. Individual 1 has an equal chance of being paired with either 2, 3, or 4. So, the ex ante average ability of his partner is 3. Individual 4 has an equal chance of being paired with either 1, 2, or 3, and thus the ex ante average ability of his partner is 2. This mechanical relationship between own ability and the mean ability of randomly assigned peers—which is a general problem in all peer effects studies—causes estimates of equation (1) to produce negative values of $\hat{\pi}_2$. Random assignment appears nonrandom, and positively matched peers can appear randomly matched.

This bias is decreasing in the size of the population from which peers are drawn, i.e., the size of the urn. As the urn increases in size, each individual contributes less to the average ability of the population from which peers are drawn, and the difference in average ability of potential peers for low and high ability individuals converges to zero. In settings where peers are drawn from large groups, ignoring this mechanical relationship is inconsequential. In our case, the average urn size is 60, and 25 percent of the time the urn size is less than 18.

We present Monte Carlo results in Figure 2 which confirm that estimates of (1) are negatively biased and that the bias is decreasing in the size of the urn. We report the results from two simulations. For the first simulation, we created 55 players with ability drawn from a normal distribution with mean 0 and standard deviation 1. We then created 100 tournaments and for each tournament randomly selected M players. Each of these M players were then assigned to groups of three. The size of the urn from which peers are drawn, M, was randomly chosen to be either 39, 42, 45, 48, or 51 with equal probability, for an average M of 45. We explain, below, the reason for the variation in M. Finally, we estimated an OLS regression of own ability on the average of partners' ability, controlling for tournament fixed effects, and the estimates of π_2 and p-values were saved. This procedure was repeated 10,000 times. For the second simulation, we increased the average size of M by creating 550 players, and allowed M to take on values of 444, 447, 450, 453, and 456 with equal probability, for an average M of 450.

The results from the first simulation are shown on the left side of Figure 2. As predicted, the typical OLS randomization test is not well-behaved. The test substantially overrejects, rejecting at the 5 percent level more than 33 percent of the time. Even though peers are randomly assigned, the estimated correlation between abilities of peers is, on average, -0.046. This negative relationship is exactly as one should expect, resulting from the fact that individuals cannot be their own peers.

The intuitive argument made above also implies that the size of the bias should be decreasing in M. Indeed, this is the case. The right side of Figure 2 shows results from the second simulation, where the urn size was increased by an order of magnitude. The typical randomization test is more well-behaved. The estimates of π_2 center around 0, the test rejects at the 5 percent level 7.5 percent of the time, and p-values are close to uniformly distributed between 0 and 1. In short, the Monte

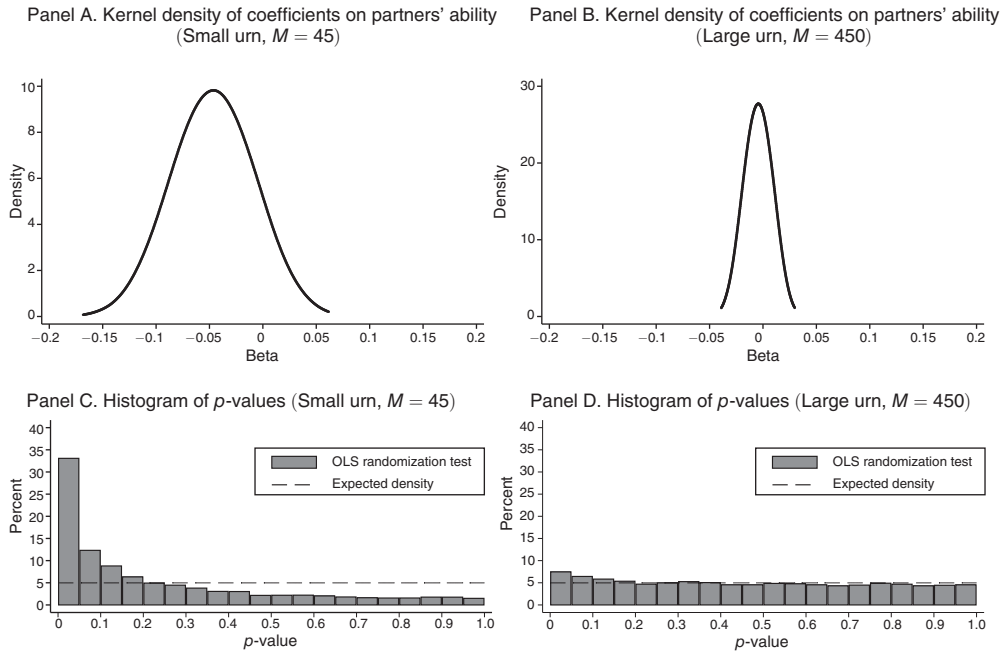

FIGURE 2. MONTE CARLO OF OLS RANDOMIZATION TESTS

Notes: This graph displays the results of a Monte Carlo study to show the bias of the randomization test described in the text. There are two sets of results. The "small urn" setting (where three-person pairings are randomly drawn from an urn of size 45), and the "large urn" setting (where three-person pairings are randomly drawn from an urn of size 450). In the large urn setting, the randomization test is fairly well-behaved, while in the small urn setting the distribution of betas is negatively biased, and *p*-values from a randomization test that tests whether beta is statistically significantly different from zero are very far from uniformly distributed. Monte Carlo results are based on 10,000 repetitions.

Carlo results show that the typical test for randomization is biased when the set of individuals from which peers are drawn is relatively small.

B. *A Proposed Solution*

To our knowledge, this point has not been made clearly in the literature.[26] We propose a simple correction to equation (1) that produces a well-behaved test of random assignment of peers, even with small urn sizes. Since the bias stems from the fact that each individual's peers are drawn from a population with a different mean ability, we simply control for that mean. Specifically, we add to equation (1)

[26] The closest discussion that we are aware of is by Michael A. Boozer and Stephen E. Cacciola (2001), who point out that the ability to detect peer effects in a linear-in-means regression of outcomes on mean outcomes is related to the size of the reference group, but they do not link this discussion to tests for random assignment nor to the fact that individuals cannot be assigned to themselves as peers.

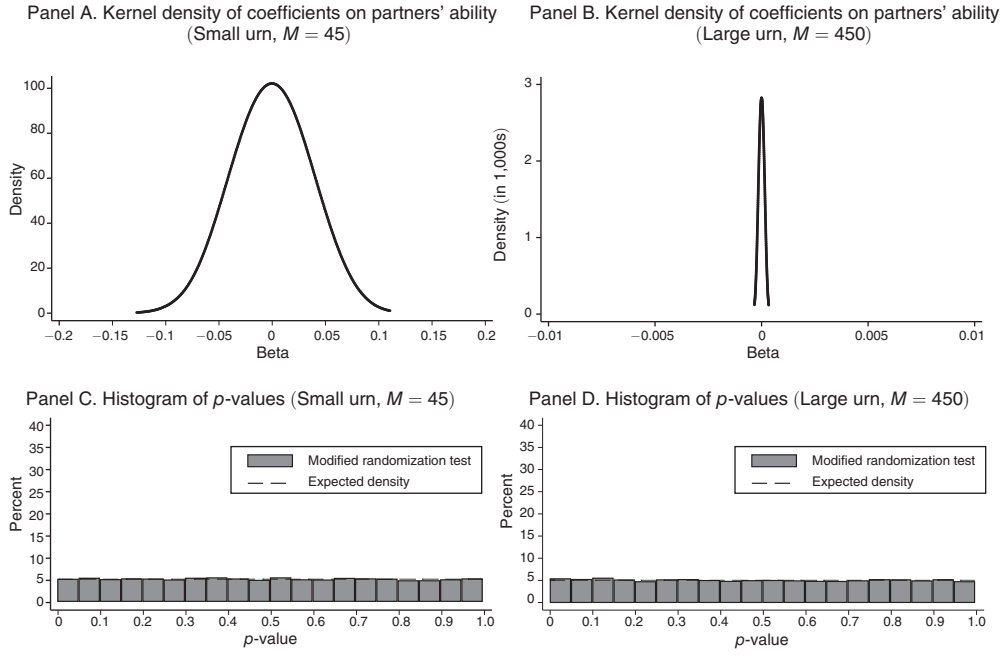

FIGURE 3. MONTE CARLO OF MODIFIED RANDOMIZATION TESTS

Notes: This graph displays the results of a Monte Carlo study to show that our proposed "correction" of the biased randomization test described in Figure 2 (and the accompanying text) produces a well-behaved randomization test regardless of the size of the urn. There are two sets of results: the "small urn" setting (where three-person pairings are randomly drawn from an urn of size 45) and the "large urn" setting (where three-person pairings are randomly drawn from an urn of size 450). In both settings, the randomization test is very well behaved (p-values from a randomization test that tests whether beta is statistically significantly different from zero are uniformly distributed). Monte Carlo results are based on 10,000 repetitions.

the mean ability of all individuals in the urn, excluding individual i. The modified estimating equation is

$$Ability_{ikt} = \pi_1 + \pi_2 \overline{Ability}_{-i,kt} + \delta_{tc} + \varphi \overline{Ability}_{-i,ct} + u_{ikt}, \tag{2}$$

where $\overline{Ability}_{-i,ct}$ is the mean ability of all players in the same category $\times$ tournament cell as player i, other than player i himself (i.e., all individuals that are eligible to be matched with individual i), and φ is a parameter to be estimated. It should be noted that it is necessary for there to be variation in the size of player i's urn to be able to separately identify π_2 and φ.[27]

Figure 3 shows the results from Monte Carlo simulations analogous to those reported above. The difference here is that instead of estimating the typical OLS regression, we include $\overline{Ability}_{-i,ct}$ as an additional regressor. As can be seen clearly

[27] If every urn has N players, then my ability $Ability_{ikt}$ is related to the mean ability in my urn $\overline{Ability}_{ct}$ and the "leave-me-out" mean $\overline{Ability}_{-i,ct}$ by the following identity: $Ability_{ikt} = N\,\overline{Ability}_{ct} - (N-1)\,\overline{Ability}_{-i,ct}$.

in Figure 3, the addition of this control makes the OLS test of randomization well-behaved regardless of whether average urn size is large or small. In both cases, the estimated correlation of peers' ability centers around 0, the test rejects at the 5 percent level approximately 5 percent of the time, and *p*-values are approximately uniformly distributed between 0 and 1. In results not reported here, we have also confirmed that the test can detect deviations from random assignment. Going forward, we include $\overline{Ability}_{-i,ct}$ as a control in tests for random assignment.

An alternative to the regression control approach that we propose is to compare the estimated π_2 to a distribution generated by randomly assigning golfers to counterfactual peer groups and estimating π_2.[28] In our case, we would repeatedly assign the golfers in our data to counterfactual groups of three according to the conditional random mechanism assumed by the null hypothesis. For each set of peer group assignments, we would estimate π_2 according to the typical OLS randomization test described by equation (1), repeating the process a large number of times. The $\hat{\pi}_2$ that we estimate from the real peer group assignments in our data could then be compared to the distribution of $\hat{\pi}_2$ generated from this process. As we describe in Section IIID, our estimate of $\hat{\pi}_2$ from such an exercise lies very close to the median of the randomly generated distribution of $\hat{\pi}_2$, yielding the same conclusion as our corrected randomization test.

C. *Summary Statistics*

Before turning to a test of randomization in our sample, we present descriptive statistics in Table 1. It is important to understand the units of our primary variables of interest, score, and ability. Score is a variable that represents the number of strokes the player took, and is the actual golf score the player achieved in a given tournament round. Ability, while in the same units as score, is typically expressed as deviation of score from par, which is 72 strokes. Throughout the results section, it is helpful to keep in mind that lower scores in golf indicate better performance (and, analogously, a lower ability measure indicates a higher ability player).

Figure 4 shows the distribution of players' handicaps in each tournament by category. Two things should be noted from the figure. First, there is a reasonable amount of variation in ability even among professional golfers. Across the four categories, the difference between the ninetieth percentile and tenth percentile in adjusted average score—our baseline measure of ability—is 1.97. Perhaps more importantly, given the stratification by category prior to random assignment, much of the variance in ability remains after separating by categories. As can be seen clearly in Figure 4, the average ability increases from category 1 to category 3. However, there is a great deal of overlap in the distributions. The 90–10 differences in measured ability in category 1, 1A, 2, and 3 are 1.93, 1.64, 1.86, and 2.84, respectively. These differences represent wide ranges in ability. They are differences in average scores *per round*, and tournaments are typically four rounds long. Using our data on earnings, a reduction in handicap of 1.97 translates into an increase in expected tournament earnings

[28] Such a procedure is closely related to exact inference (Ronald A. Fisher 1935) and randomization inference (Paul R. Rosenbaum 2002).

Table 1—Descriptive Statistics

Variable	#	Mean	SD	Percentiles 10th	25th	50th	75th	90th
Score	17,492	71.158	3.186	67	69	71	73	75
Average(score), partners	17,492	71.158	2.606	68	69.5	71	73	74.5
Ability (handicap)	17,492	−2.86	0.92	−3.81	−3.40	−2.93	−2.45	−1.84
Average(ability), partners	17,492	−2.86	0.78	−3.66	−3.32	−2.94	−2.53	−2.04
Driving distance (yards)	17,182	281.00	9.85	269.10	274.01	280.10	288.00	293.30
Putts per round	17,182	28.67	1.03	27.33	28.03	28.83	29.34	29.72
Greens per round	17,182	11.60	0.65	10.90	11.33	11.66	11.97	12.32
Years of experience	17,492	11.80	6.69	3	6	11	18	21
Tiger Woods in group	17,492	0.00	0.06	0	0	0	0	0
Category 1 players								
Ability	6,377	−3.14	0.77	−4.10	−3.62	−3.20	−2.65	−2.17
Average(ability), partners	6,377	−3.14	0.60	−3.86	−3.53	−3.15	−2.77	−2.41
Category 1A players								
Ability	6,818	−2.81	0.73	−3.67	−3.23	−2.87	−2.43	−2.03
Average(ability), partners	6,818	−2.82	0.58	−3.46	−3.18	−2.86	−2.50	−2.18
Category 2 players								
Ability	2,998	−2.86	0.90	−3.73	−3.36	−3.02	−2.51	−1.87
Average(ability), partners	2,998	−2.85	0.75	−3.59	−3.29	−2.95	−2.55	−2.08
Category 3 players								
Ability	1,299	−1.68	1.45	−3.13	−2.50	−1.65	−1.13	−0.29
Average(ability), partners	1,299	−1.71	1.25	−3.12	−2.50	−1.77	−1.13	−0.35

Note: See Appendix A for more information on player categories.

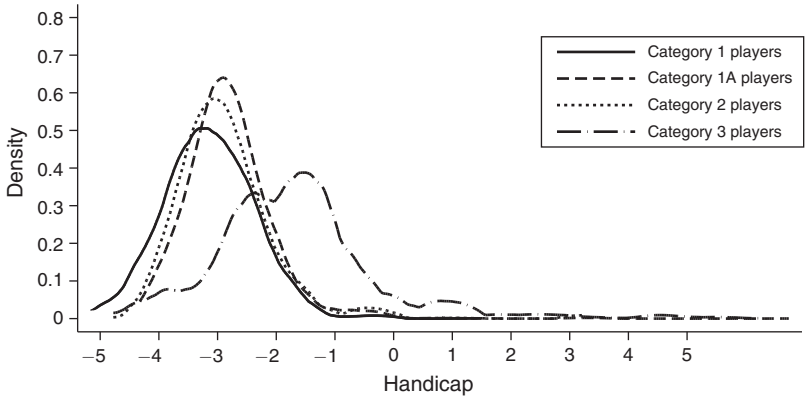

Figure 4. Kernel Density of Player Handicaps, by Category

Note: This graph displays the density of handicaps (our primary measure of player ability) for each player category.

Table 2—Testing for Random Assignment

	(1)	(2)	(3)	(4)
Average(ability), partners	−0.018	−0.088	0.158	0.089
	(0.015)	(0.022)	(0.021)	(0.020)
Leave-me-out average(ability), urn	−10.803			
	(1.629)			
Tournament fixed effects	Yes	Yes	Yes	No
Category fixed effects	Yes	Yes	No	Yes
Tournament × category fixed effects	Yes	Yes	No	No
Leave-me-out urn mean included	Yes	No	No	No
R^2	0.542	0.254	0.124	0.162
N	8,801	8,801	8,801	8,801

Notes: Results from estimating equation (2). Average(ability) is the average handicap of a player's playing partners, as described in the text. Column 1 includes the average of all of the other players in your tournament-by-category urn (not including yourself). This control is necessary to produce a well-behaved randomization test (see text for more details). Standard errors are in parentheses and are clustered by playing group. All regressions are unweighted (unlike the specifications in the other tables), since the weighting is not appropriate for the randomization test.

of 87 percent.[29] This suggests that differences in strokes on this order of magnitude should be quite salient to players.

D. *Verifying Random Assignment of Peers in Professional Golf Tournaments*

Using our various measures of ability, we test the claim that assignment to playing groups is random within a tournament-by-category cell.[30] In this section, we report results from estimating variations of equation (2), with various measures of pre-determined ability. As discussed earlier, the correct randomization test includes the full set of tournament-by-category fixed effects along with the control $\overline{Ability}_{-i,ct}$. Table 2 reports these results. In column 1, we present results from this "correct" randomization test. The coefficient on partners' average ability is −0.018 and insignificant. The small insignificant conditional correlation between own ability and partners' ability is consistent with players being randomly assigned to playing partners.[31]

In the remaining columns of Table 2, we present results from "incorrect" randomization tests to illustrate the importance of controlling for $\overline{Ability}_{-i,ct}$, and to show that tests of this form have the power to detect deviations from random assignment. In column 2, we present estimates of the randomization test that are most typical in the literature. It includes the full set of tournament-by-category fixed effects but

[29] A regression of log earnings on handicap with a full set of tournament fixed effects, category fixed effects, and their interactions gives a coefficient on handicap of −0.318 with a t-statistic of 13.11. Thus, a reduction in handicap of 1.97 will increase earnings by 0.626 log points, or 87 percent.

[30] This claim is based on the PGA Player Handbook and Tournament Regulations, and on numerous telephone conversations with PGA Tour officials.

[31] The coefficient on the bias correction term is −10.803 (standard error of 1.629). When the variation in urn size is small, we expect the magnitude of this coefficient to be roughly equal to $-(\overline{N}-1)$, where $\overline{N}$ is the average urn size. As the variation in urn size increases, the absolute value of this coefficient declines, as observations from smaller urns are given more weight in the regression. In our data, the average tournament by category cell has 28.9 golfers with a standard deviation of 17.3. We have replicated through Monte Carlo simulations a coefficient very close to ours using simulated groupings drawn from urns with the same average size and standard deviation.

TABLE 3—ROBUSTNESS TESTS OF RANDOM ASSIGNMENT

	Dependent variable, $X =$					
	Ability (1)	Driving distance (2)	Putts per round (3)	Greens per round (4)	Years of experience (5)	Length of name (6)
Panel A. Modified randomization test						
Average(X), playing partners	−0.018	0.002	−0.002	−0.004	−0.010	−0.013
	(0.015)	(0.015)	(0.012)	(0.014)	(0.012)	(0.014)
Leave-me-out average(X), urn	−10.803	−10.189	−12.529	−12.033	−15.793	−14.587
	(1.629)	(1.210)	(1.653)	(2.665)	(3.795)	(1.364)
Tournament × category fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Leave-me-out urn mean included	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.542	0.632	0.579	0.580	0.593	0.515
N	8,801	8,646	8,646	8,646	8,646	8,646
Panel B. Biased randomization test						
Average(X), playing partners	−0.088	−0.081	−0.071	−0.080	−0.050	−0.063
	(0.022)	(0.021)	(0.022)	(0.023)	(0.020)	(0.020)
Tournament × category fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Leave-me-out urn mean included	No	No	No	No	No	No
R^2	0.254	0.382	0.163	0.251	0.167	0.035
N	8,801	8,646	8,646	8,646	8,646	8,646

Notes: Results from estimating equation (2). Average(ability) is the average handicap of a player's playing partners, as described in the text. All specifications in panel A include the average of all of the other players in your tournament-by-category urn (not including yourself). This control is necessary to produce a well-behaved randomization test. Standard errors are in parentheses and are clustered by playing group. All regressions are unweighted (unlike the specifications in the other tables), since the weighting is not appropriate for the randomization test. Length of the name is the total number of characters (including spaces) in the player's name.

excludes the bias correction described in Section IVA, $\overline{Ability}_{-i,ct}$. The correlation between own and partners' ability is negative and significant (−0.088 with a standard error of 0.022). Ignoring the bias discussed in Section IVA would lead to the erroneous conclusion that peers were negatively assortatively matched. The test results would erroneously be interpreted as evidence of nonrandom assignment. In columns 3 and 4, we present estimates of equation (2) that drop the controls for category and tournament fixed effects, respectively, and omit the bias correction. Failing to account for the conditional nature of random assignment generates inference of positively matched peer groups. The positive and significant estimates from these specifications show that the test has sufficient power to detect deviations from random assignment in a setting where we know assignment is not random.

We also test for random assignment using various disagreggated measures of ability (e.g., driving distance, putts per round, greens in regulation per round, years of experience, and length of a player's name which is measured by the number of characters in the name). We report estimates of equation (2), replacing average adjusted score with these alternative measures of ability in columns 2–6 of Table 3. Panel A reports the correct test (i.e., those that include the $\overline{Ability}_{-i,ct}$ control, calculated for the respective measure of ability), while panel B reports results from the typical test excluding the correction term. In specifications with the correction control, correlations between all other measures of pre-determined ability are small in magnitude and statistically insignificant. Just as with the measure of overall ability, all

specifications that exclude the correction term yield estimated correlations that are negative and significant.[32]

IV. Empirical Framework

Having established that peers are randomly assigned, we now turn to the estimation of peer effects. We estimate peer effects using a simple linear model in which own score depends on own ability and playing partners' ability. The key identifying assumption, which was tested in the previous section, is that, conditional on tournament and category, players are randomly assigned to groups. Our baseline specification is

$$Score_{iktr} = \alpha_1 + \beta_1 Ability_i + \gamma_1 \overline{Ability}_{-i,kt} + \delta_{tc} + e_{iktr}, \tag{3}$$

where i indexes players, k indexes groups, t indexes tournaments, r indexes each of the first two rounds of a tournament, c indexes categories, δ_{tc} is a full set of tournament-by-category fixed effects to be estimated, α_1, β_1, and γ_1, are parameters, and e_{iktr} is an error term.[33] The parameter γ_1 measures the effect of the average ability of playing partners on own score, and is our primary measure of peer effects. Its magnitude is generally evaluated in relation to the magnitude of β_1, which is the effect of own ability on own score. Since playing partners are randomly assigned, the coefficients in equation (3) can be estimated consistently using OLS.[34]

Even with the handicap correction described above, a remaining potential problem with estimating equation (3) is that $Ability_i$ might be measured with error. This should be a concern for all studies of peer effects that estimate exogenous effects. A nice feature of the specification above is that it contains a simple correction for measurement error. If each golfer had a single peer and each golfer's individual measure of ability contained the same amount of measurement error, then the estimates of $\hat{\beta}_1$ and $\hat{\gamma}_1$ would be equally attenuated. In this case, the ratio $\hat{\gamma}_1/\hat{\beta}_1$ would give us a measurement-error-corrected estimate of the reduced-form exogenous peer effect. In addition to reporting this ratio, we also report measurement-error-corrected estimates following Wayne A. Fuller and Michael A. Hidiroglou (1978) and David Card and Lemieux (1996). This estimator corrects for attenuation bias of a known (and estimated) form and is described in more detail in Appendix B. The advantage of this estimator is that it allows for the degree of measurement error to vary by player and leverages the structure imposed on the measurement error from the fact that the regressor of interest is an average of two error-ridden measures of ability.

[32] In addition to testing for randomization using our modified randomization test, we have also tested for random assignment by repeatedly drawing new sets of groupings and computing the correlation between own ability and partners' ability for each drawing. The location of -0.088 (see column 1 of panel B in Table 3) in the empirical distribution of correlations provides a valid test of random assignment. The median correlation based on 10,000 iterations is -0.089 with an empirical 95 percent confidence interval of $(-0.049, -0.129)$. We thus find no evidence of nonrandom assignment ($p = 0.839$). We have conducted this randomized inference on all of the variables in Table 3, and we find no evidence against random assignment for any of the variables.

[33] The results reported throughout the paper include a fixed effect for the second round and are essentially unchanged if we omit it or include a full set of tournament-by-round fixed effects.

[34] The peer effects estimates that we report throughout are essentially unaffected, though slightly more precise, if the bias correction term from the randomization test is included.

A key advantage to estimating the reduced-form specification in (3) is that the average ability of playing partners is a pre-determined characteristic. Thus, our estimate of γ_1 is unlikely to be biased due to the presence of common unobserved shocks.[35] An alternative commonly estimated specification replaces peers' ability with peers' score. This outcome-on-outcome specification estimates a combination of endogenous and contextual effects, but intuitively examines how performance relates to the contemporaneous performance of peers, rather than just to peers' pre-determined skills. Even with random assignment, one cannot rule out that a positive relationship between own score and peers' score is driven by shocks commonly experienced by individuals within a peer group. We nevertheless run regressions of the following form to get an upper bound on the magnitude of peer effects:

$$Score_{iktr} = \alpha_2 + \beta_2 Ability_i + \gamma_2 \overline{Score}_{-i,ktr} + \delta_{tc} + \nu_{iktr}, \tag{4}$$

where $\overline{Score}_{-i,ktr}$ is the average score in the current round of player i's playing partners. Because common shocks are expected to cause positive correlation in outcomes, the estimate of γ_2 should be viewed as an upper bound on the extent of peer effects. Since we can observe playing groups playing at the same time nearby on the course, we are able to gauge the magnitude of the bias created by common shocks. We report estimates that take advantage of this feature of the research design in the following section.

V. Results

A. *Visual Evidence*

To get a sense of the importance of peer effects, we first plot regression-adjusted scores against playing partners' handicap. To do this, we regress each player's score on tournament-by-category fixed effects and a dummy variable for the second round. Then we take the residuals from this regression, compute means by each decile of the partners' ability distribution, and graph the average residual against each decile bin. Figure 5 reports this graph for the full sample, which shows zero correlation between own score and the ability of randomly assigned playing partners. Those who were randomly assigned to partners with higher average scores scored no differently than those who were assigned to partners with low average scores. There also does not appear to be evidence of nonlinear peer effects. We take this to be a first piece of evidence that peer effects among professional golfers are economically insignificant. To place a confidence interval around this estimate, we estimate the linear regression model in equation (3).

[35] Note, however, that these common shocks are likely to affect the standard errors. Hence, in the peer effect regressions, we cluster at the group level. The estimated standard errors are virtually unchanged if we cluster by tournament-by-category, and are actually smaller if we cluster by tournament.

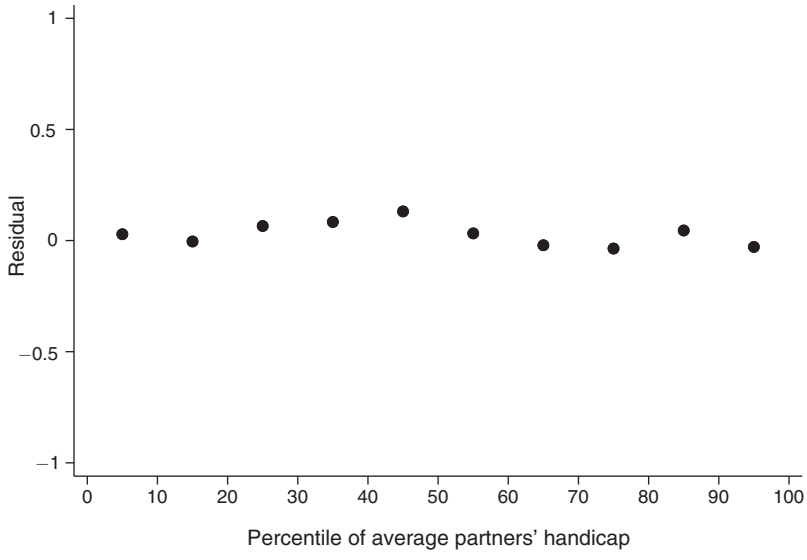

FIGURE 5. ESTIMATES OF PEER EFFECTS BY DECILE OF PARTNERS' ABILITY

Notes: This figure displays the average of a residual by decile of playing partners' handicap. The residual is from a regression of own score on own handicap and a full set of player category by tournament fixed effects. If these points suggested a positive slope, we would interpret that as evidence of positive peer effects.

B. *Regression Estimates of the Effect of Playing Partners' Ability on Own Score*

The results of estimating equation (3) are shown in the first column of Table 4. Since our measure of ability is an average of varying numbers of prior adjusted scores, we weight all regressions by the inverse of the sample variance of each player's estimated $Ability_i$. Shown in column 1, the coefficient on own ability is strongly statistically significant and large in magnitude, as expected. A one-stroke increase in a player's average score in past rounds is associated with an increase in that player's score of 0.672 strokes. That this coefficient is not equal to one suggests there is some measurement error in our measure of ability, but as a conditional reliability ratio this is reasonably large in magnitude. If we think of $\hat{\beta}_1$ this way, then, as we described above, $\hat{\gamma}_1/\hat{\beta}_1$ is a measurement-error corrected estimate of the effect of partners' ability on own score.

The estimate of γ_1, the effect of playing partners' ability on own score, is not statistically significant, and the point estimate is actually negative. The insignificant point estimate suggests that improving the average ability of one's playing partners by one stroke actually increases (i.e., worsens) one's score by 0.035 strokes. Our estimates make it possible to rule out positive peer effects larger than 0.043 strokes for an increase in average ability of one stroke. One stroke is 28 percent greater than one standard deviation in partners' average ability (0.78). If we divide the upper bound of the 95 percent confidence interval by the estimate of $\hat{\beta}_1$ to correct for

TABLE 4—THE EFFECT OF PEERS' ABILITY ON OWN SCORE

	(1)	(2)	(3)
Own ability	0.672	0.949	
	(0.039)	(0.057)	
Average(ability), partners	−0.035	−0.036	−0.032
	(0.040)	(0.063)	(0.040)
Tournament × category fixed effects	Yes	Yes	Yes
Measurement error correction	No	Yes	No
Player fixed effects	No	No	Yes
N	17,492	17,492	17,492

Notes: Results in column 1 are from baseline specifications as specified in equation (3). Column 2 reports a measurement-error-corrected estimate using the estimator described in Appendix B. Column 3 reports results using player fixed effects instead of own ability. The dependent variable is the golf score for the round. The ability variable is measured using the player's handicap. Standard errors are in parentheses and are clustered by playing group. All regressions include round fixed effects and weight each observation by the inverse of the sampling variance of estimated ability of each player.

measurement error, we can still rule out that a one-stroke increase in partners' average ability increases own score by more than 0.065 strokes. The results from our baseline specification suggest that there are not significant peer effects overall.

We address measurement error more formally in column 2, which reports results using the measurement-error-corrected estimator described in more detail in Appendix B. The coefficient on own ability increases from 0.672 to 0.949 and it is no longer statistically significantly different from 1, suggesting that much of the measurement error has been eliminated. The coefficient on partners' average ability remains essentially unchanged (from −0.035 to −0.036), and the standard error on the peer effect coefficient increases (from 0.040 to 0.063), which results in a slightly larger upper bound of the 95 percent confidence interval of 0.087. Interestingly, the point estimates suggest measurement error affects the coefficient on own ability more than the coefficient on average ability of playing, which is consistent with peers' ability being an average of two values.

Column 3 verifies that the results are insensitive to controlling for player fixed effects instead of player ability. Going forward, we report results from the specification that includes own ability rather than player fixed effects for ease of interpretation.

C. *The Effect of Different Dimensions of Ability: Does the Overall Effect Hide Evidence of Learning or Motivation?*

As described earlier, we collected data on various dimensions of player skill. We hypothesize that players might learn about wind conditions or optimal strategies from more accurate players (i.e., those who take the fewest shots to get the ball on the green), or from better putters. In contrast, we assume that players cannot learn how to hit longer drives by playing alongside longer hitters, and that any effect of

Table 5—Peer Effects with Alternate Measures of Player Ability

	(1)	(2)	(3)	(4)	(5)
Own ability	0.672				
	(0.039)				
Average(ability), playing partners	−0.035				
	(0.040)				
Driving distance		−0.022			−0.009
		(0.004)			(0.004)
Average(driving), playing partners		0.003			0.003
		(0.004)			(0.004)
Putts			0.137		0.172
			(0.030)		(0.030)
Average(putts), playing partners			−0.038		−0.045
			(0.039)		(0.039)
Greens per round				−0.687	−0.682
				(0.049)	(0.051)
Average(greens), playing partners				−0.021	−0.023
				(0.058)	(0.060)
Tournament × category fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.154	0.141	0.140	0.149	0.152
N	17,492	17,182	17,182	17,182	17,182

Notes: Results from baseline specifications as specified in equation (3). The dependent variable is the golf score for the round. The ability variable is measured using the player's handicap. Standard errors are in parentheses and are clustered by playing group. All regressions include round fixed effects and weight each observation by the inverse of the sample variance of estimated ability of each player.

playing alongside a longer driver must operate through increased motivation.[36] If these assumptions are correct, specifications comparable to (3) but replacing partners' average ability with partners' average driving distance, putts per round, or greens reached in regulation can separately identify motivation and learning effects. An effect of the accuracy measures (putts per round and greens reached in regulation) would be interpreted as evidence of learning from peers, while an effect of partners' driving distance would be interpreted as evidence of motivation by peers.[37]

The results are presented in columns 2–5 of Table 5 (column 1 reproduces baseline results with average ability). In column 2, we present results using average driving distance. While the coefficient on own driving distance is negative and strongly statistically significant (longer drives enable a player to achieve a lower score), the point estimate on partners' driving distance is small and statistically insignificant. The results for putts per round, shown in column 3, are similar. Own

[36] Just as in the measures of overall ability, there is also a good deal of variation across players in these specific dimensions of past performance. The ninetieth percentile golfer hits his initial drive 24 yards farther than the golfer at the tenth percentile in average driving distance. This is 8.6 percent of the mean drive length. For putting, the differences are similar. The ninetieth percentile putter hits 2.4 fewer putts per round than the tenth percentile putter (8.3 percent of the mean). And for accuracy, the 90–10 difference is 12.2 percent of the mean for greens hit in regulation.

[37] One might argue that the former is a production complementarity in the typology set out at the outset of the paper, but the latter is clearly a purely social peer effect.

putting skill has a large and strongly significant effect on own score, but golfers do not appear to shoot lower scores when they play with better putters. The results for shot accuracy, shown in column 4, similarly show strong effects of own accuracy, but no effect of partners' accuracy on a golfer's performance.[38] Finally, we present a specification that jointly estimates the effects of all three measures of ability in column 5. A golfer's putting accuracy and number of greens hit in regulation have the most significant effects on his score,[39] but, as in the separately estimated specifications, no dimension of his partners' ability appears to have any effect on score.

Another form learning might take is that the effect of playing alongside a better golfer would manifest as time passes. There does not appear to be any evidence of such a pattern in our data. In results not presented here, we find no differential effect of playing partners' ability in the second nine holes as compared with the first nine holes of the round, no differential effect on the second day played with the same partners, and no carryover effect of partners' ability one, two, or three tournaments (i.e., weeks) later.

D. *Using Alternative Measures of Peer Ability*

Having seen no evidence that the average ability of peers affects individual performance, we ask whether the linear-in-means specification obscures peer effects in a different way. It is possible that it is not the mean ability of co-workers that matters, but rather the minimum or maximum ability of co-workers. Possibly playing with bad players matters, but playing with good players does not. Or, maybe playing alongside one very good player or one very bad player affects performance. In each of these cases, the mean ability of peers would not measure the relevant peer environment accurately. Motivated by these possibilities, in Table 6, we present estimates of specification (3), where $\overline{Ability}_{-i,kt}$ is replaced with alternative measures of peers' ability. We report the baseline specification in column 1 for comparison. In column 2, we replace the average ability with the maximum ability of the player's peers. The point estimate is slightly smaller, but virtually unchanged. In column 3, we show that the estimated effect of the minimum of peers' ability is, again, negative and insignificant, and virtually the same as the average ability effect.

In columns 6–9, we investigate whether there appears to be a nonlinear effect of partners' ability. To do this, we include indicators for whether individual i was assigned to a player in the top decile, top quartile, bottom quartile, or bottom decile of the ability distribution in his category. None of the four estimates are statistically significant, though suggestively the point estimates for the top-quantile specifications are positive while those for the bottom-quantile specifications are

[38] We also collected data on driving accuracy, specifically the fraction of fairways the golfer hits on tee shots. We do not include specifications using this variable because it does not strongly predict own score. This result is consistent with the work of Donald L. Alexander and William Kern (2005), who find large effects of putting accuracy on earnings, smaller effects of driving distance on earnings, and very small effects of driving accuracy on earnings. When we do estimate peer effects using driving accuracy, however, we find no effect of partners' driving accuracy on own score.

[39] Two old golfing cliches seem to be consistent with the data: "Drive for show, putt for dough" and "Hit fairways and greens."

Table 6—Peer Effects with Alternative Measures of Peer Ability

Measure of peer ability:	Average ability (1)	Maximum ability (2)	Minimum ability (3)	Tiger Woods is partner (4)	Average ability × (average ability − own ability) (5)
Own ability	0.672	0.673	0.672	0.674	0.672
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Peer ability	−0.035	−0.023	−0.031	−0.348	−0.016
	(0.040)	(0.032)	(0.040)	(0.462)	(0.014)
R^2	0.154	0.154	0.154	0.154	0.154
N	17,492	17,492	17,492	17,492	17,492

Measure of peer ability:	**1**{any partner in top 10%} (6)	**1**{any partner in top 25%} (7)	**1**{any partner in bottom 25%} (8)	**1**{any partner in bottom 10%} (9)
Own ability	0.673	0.672	0.673	0.673
	(0.039)	(0.039)	(0.039)	(0.039)
Peer ability	0.027	0.049	−0.041	−0.178
	(0.070)	(0.059)	(0.069)	(0.141)
R^2	0.154	0.154	0.154	0.154
N	17,492	17,492	17,492	17,492

Notes: Column 1 is reproduced from Table 3. Other columns present results from modifying baseline specifications as specified in equation (3) to support heterogeneous peer effects. The dependent variable is the golf score for the round. The ability variable is measured using the player's handicap. Standard errors are in parentheses and are clustered by playing group. All regressions weight each observation by the inverse of the sample variance of estimated ability of each player. All regressions include tournament-by-category fixed effects and round fixed effects. In column 9, average ability of playing partners is also included in regression. The estimated coefficient for this variable is −0.014 (0.048).

negative. Recall that lower scores are better, so this pattern would suggest that players play worse when they are matched with much better players. We also ask whether playing with Tiger Woods, the best player of his generation, affects performance. The point estimate in column 4 suggests that being partnered with Tiger Woods reduces golfers' scores, but the standard errors are large enough that we cannot rule out a zero effect. In our sample, there are only 70 golfer-days paired with Tiger Woods.

The specifications thus far have assumed it is the absolute level of peers' ability that affects performance. An alternative hypothesis is that relative ability also matters. To investigate this possibility, we present a specification that allows the effect of peers' ability to vary with the difference between peers' and own ability. The results are reported in column 5 of Table 6, and they suggest that the effect of peers' ability does not vary with relative ability. Similar specifications based on the other measures of ability also yield small and statistically insignificant results.

Finally, we consider whether there exist peer effects more generally, beyond the skill-based peer effects for which we have tested thus far. In particular, we create a set of J partner dummy variables, equal to one if player i is partnered with player j, where j ranges from one to J. We then estimate a model in which each player's score depends on this set of playing partner fixed effects. The F-test of the joint

significance of these playing partner fixed effects indicates whether performance systematically varies with the identity of one's playing partner. This test is more general than those presented thus far because it allows playing partners' effects to be based on unobservable characteristics. For example, the *F*-test would reject if a group of mediocre golfers improved the scores of their partners, for example, because they were pleasant people. The test also would detect peer effects if there were performance-enhancing and performance-reducing partners. In our data, however, the *F*-test that the coefficients on the full set of playing partner dummies are jointly zero fails to reject. Consistent with our previous results, we do not find any evidence that there are heterogeneous peer effects.[40]

E. *Endogenous Effect Regressions and Common Shocks*

Table 7 reports results of equation (4), the specification that replaces partners' ability with partners' score as the regressor of interest. As described earlier, the coefficient estimate on playing partners' score overstates the true peer effect if there are unobserved common shocks affecting all players uniformly in the group. Nevertheless, this regression is informative as an upper bound on γ_2. The first column in Table 7 shows, contrary to the results above, that the peer effect is positive and statistically significant—an increase in the average score of one's playing partners is associated with an increase of own score by 0.055 strokes. An important point, however, is that without accounting for the upward bias in this estimate due to common shocks, the coefficient is still small in magnitude. Mas and Moretti's (2009) elasticities, evaluated at the mean of our dependent variable (own score), would predict that an increase in average partners' score of one stroke would raise own score by 0.170 strokes, more than two times what we estimate. We do not correct for measurement error here, since each player's score, and the average score of peers, is measured without error.[41]

To look at the importance of common shocks more systematically, we try several additional controls. We hypothesize that the most likely sources of common shocks are variation in weather and crowd size. Because these shocks also simultaneously affect nearby playing groups, we construct controls for common shocks that are based on comparing groups with similar starting times.

In column 2, we interact time-of-day (early morning, mid-morning, afternoon) fixed effects with the full set of tournament fixed effects. This should capture weather shocks and other changes in course conditions that affect all groups that

[40] To get an unbiased test statistic, we simulate an empirical distribution of *F*-statistics by randomly reassigning pairings within a tournament-by-category cell. Our *p*-value, based on the location of the actual *F*-statistic within the simulated distribution, is 0.242 (based on 1,000 simulations). We also try clustering the standard errors on playing group and repeat this bootstrap procedure and find a *p*-value of 0.409. Both pieces of evidence lead us to conclude that there are no playing-partner-specific peer effects.

[41] One might argue that score is a noisy measure of performance. Even though score is the measure by which players are judged, a golfer may play well but have a high score because of a few unlucky bounces. In this case, the measurement error correction used in the previous section will likely produce estimates that are too large because $Ability_i$ is an even noisier measure of the contemporaneous performance of player i than $Score_{-i,ktr}$ is of his partners' contemporaneous performance.

TABLE 7—THE EFFECT OF PEERS' SCORE ON OWN SCORE

	(1)	(2)	(3)	(4)	(5)
Own ability	0.677	0.676	0.673	0.668	0.666
	(0.039)	(0.039)	(0.039)	(0.039)	(0.040)
Average(score), partners	0.055	0.028	0.019	0.011	0.003
	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)
Tournament × category fixed effects	Yes	Yes	Yes	Yes	Yes
Tournament × time-of-day fixed effects	No	Yes	No	No	No
Time cubic per tournament	No	No	Yes	No	No
Time quartic per tournament	No	No	No	Yes	No
Time quintic per tournament	No	No	No	No	Yes
R^2	0.155	0.170	0.175	0.179	0.184
N	17,492	17,492	17,492	17,492	17,492

Notes: Results from alternate specifications using average partners' score instead of average partners' ability as the primary independent variable of interest. The dependent variable is the golf score for the round. The ability variable is measured using the player's handicap. Standard errors are in parentheses and are clustered by playing group. All regressions include round fixed effects and weight each observation by the inverse of the sample variance of estimated ability of each player.

play at the same part of the day (e.g., a common complaint on the PGA Tour is that afternoon groups experience more "spike marks" on the green, which make it more difficult to putt effectively). The point estimate on partners' score drops by about 50 percent, and is marginally significant. To capture the fact that weather and other common shocks vary more smoothly than the dummy specification assumes, in column 3, we introduce a cubic in start-time which is allowed to vary by tournament. The coefficient on partners' score is further reduced to 0.019 and is no longer significant at conventional levels. In additional specifications, we include higher-order polynomials in start time, which allow the effect of weather, and other common shocks that change over the course of a day, to vary more and more flexibly. Moving across the columns, as the order of the polynomial increases from a cubic to a quartic to a quintic, the estimated effect of partners' score decreases. With the control for a quintic in start time, the endogenous peer effect coefficient is 0.003, and is insignificantly different from zero. Interestingly, adding controls for start time does not affect the estimate of own ability on own score. It appears that the correlations between own score and partners' score are driven primarily by common shocks.

As with any peer effects regression of own outcome on peer's outcome, it is difficult to interpret the regressions shown in Table 7. They should certainly be regarded as upper bounds for peer effects since any remaining common shocks that are not controlled for will bias the estimates upwards. Furthermore, regressions of outcomes on peers' outcomes suffer from the "reflection problem" described by Manski (1993). In short, the fact that more and more extensive controls for common shocks reduce the estimate of γ_2 but do not appreciably affect the estimate of β_2, along with the fact that the estimates of γ_1 are consistently zero, lead us to conclude that peer effects are negligible among professional golfers.

F. *Do Peer Effects Vary with Player Skill or Player Experience?*

One possible explanation for why we find such different results than previous studies is that there may exist heterogeneity in the susceptibility of workers to social influences by co-workers. Professional golfers are elite professionals subject to a selection process, and perhaps the most successful professional golfers are those who are able to avoid these social responses. If heterogeneity in ability across occupations explains the differences between our results and those in Mas and Moretti (2009), it may also be the case that there is heterogeneity among golfers in the susceptibility to social influences. In Table 8, we present estimates of equation (3) that allow the effect of partners' ability to vary by the reference player's skill. This interaction tells us, for example, whether high- or low-skilled players respond more to playing with high-skilled players.[42] The results are shown in column 2. The positive coefficient on this interaction term implies that lower-skill players respond more to their co-workers' ability than do better players. The coefficient is statistically significant at conventional levels and is consistent with the idea that more skilled workers are less responsive to peer effects. Together with the small point estimate for the average-skill golfer, this interaction indicates that there are some high-skill players who appear to experience small negative peer effects and some low-skill players who appear to experience small positive peer effects.

Columns 3 and 4 report estimates of equation (3) that allow the effect of partners' ability to vary by the reference player's experience. We measure experience as the number of years since the player's first full year on the PGA Tour.[43] The positive and statistically significant coefficient on the interaction of experience and partners' ability (in columns 3 and 4) implies that more experienced players respond more to their co-workers' ability. This is inconsistent with the idea that experience mitigates peer effects. However, more experienced players appear to get higher (i.e., worse) scores, suggesting that comparisons based on experience are complicated by selection. Thus, it is difficult to discern whether the positive experience interaction indicates that players learn to benefit from their peers with experience, or that lower ability players (as proxied by high experience on the PGA Tour) are more prone to influence from their peers.[44]

VI. Concluding Comments

We use the random assignment of playing partners in professional golf tournaments to test for peer effects in the workplace. Contrary to recent evidence on supermarket checkout workers and soft fruit pickers, we find no evidence that the ability

[42] It is worth pointing out that high and low skill are relative terms. All professional golfers are extremely high skilled relative to the population.

[43] Results using other measures of experience (number of years on PGA Tour, number of years since "turned pro") are very similar.

[44] In unreported regressions, we have also tested whether peer effects vary with the strength of financial incentives. We did not find any evidence that larger financial incentives reduce peer effects, although the specifications produced large standard errors. We also note that even in the tournaments with the smallest purses and the least convex prize structure, the financial incentives are still very strong.

Table 8—Peer Effects by Skill and Experience

	(1)	(2)	(3)	(4)
Own ability	0.672	0.668	0.660	0.656
	(0.039)	(0.039)	(0.039)	(0.039)
Average(ability), partners	−0.035	−0.033	−0.038	−0.036
	(0.040)	(0.040)	(0.040)	(0.040)
Average(ability), partners × own ability		0.077		0.081
		(0.032)		(0.033)
Years of experience			0.019	0.019
			(0.004)	(0.004)
Average(ability), partners × years of experience			0.015	0.015
			(0.005)	(0.005)
Tournament × category fixed effects	Yes	Yes	Yes	Yes
R^2	0.154	0.154	0.156	0.156
N	17,492	17,492	17,492	17,492

Notes: Column 1 is reproduced from Table 3. Other columns present results from modifying baseline specifications as specified in equation (3) to support heterogeneous peer effects. The dependent variable is the golf score for the round. The ability variable is measured using the player's handicap. Standard errors are in parentheses and are clustered by playing group. All regressions include round fixed effects and weight each observation by the inverse of the sample variance of estimated ability of each player.

or current performance of playing partners affects the performance of professional golfers. With a large panel dataset, we observe players repeatedly, and the random assignment of players to groups makes it straightforward to estimate the causal effect of playing partners' ability on own performance. The design of professional golf tournaments also allows for direct examination of the role of common shocks, which typically make identification of endogenous peer effects difficult. In our preferred specification, we are able to reject positive peer effects of more than 0.043 strokes for a one stroke increase in playing partners' ability. We are also able to rule out that the peer effect is larger than 6.5 percent of the effect of own ability.

Interestingly, we find a small positive effect of partners' score on own score, but we interpret this as mostly due to common shocks (and we present evidence that controlling for these common shocks reduces this correlation). The raw correlation in scores, though, might help explain why many PGA players perceive peer effects to be important. Our results suggest that players might be misinterpreting common shocks as peer effects.

Our estimates contrast with a number of recent studies of peer effects in the workplace. Mas and Moretti (2009) find large peer effects in a low-wage labor market where workers are not paid piece rates and do not have strong financial incentives to exert more effort. Bandiera, Barankay, and Rasul (2009) find peer effects specific to workers' friends in a low-skilled job where workers are paid piece rates. Experimental studies (e.g., Falk and Ichino 2006) find evidence of peer effects in work-like tasks. We conclude by speculating several nonexclusive explanations for our contrasting findings and conclude that there is much to be learned from the difference between our results and those of other recent studies.

First, ours is the only study of which we are aware that estimates peer effects in a workplace where peers are randomly assigned. As we discuss above, random assignment is not a panacea for estimating peer effects, but it is extremely helpful in overcoming many of the difficult issues associated with identifying such models.

Second, the PGA Tour is a unique labor market that is characterized by extremely large financial incentives for performance. In such a situation, it may be the case that the incentives for high effort are already so high that the marginal effect of social considerations is minimal, or zero. In similar labor markets where there are high-powered incentives for better performance (e.g., surgeons, floor traders at an investment bank, lawyers in private practice, tenure track professors), the social effects of peers may not be as important as implied by existing studies. Consistent with this view is Mas and Moretti's (2009) conclusion that the peer effects they observe are mediated by co-worker monitoring. As incentives become stronger and monitoring output becomes easier, monitoring effort becomes less necessary.

Perhaps just as interesting is the implication that social incentives may be a substitute for financial incentives. This would suggest that when creating strong financial incentives is difficult (such as when monitoring costs are high, or measuring individual output is difficult), firms should optimally organize workers to take advantage of social incentives.

Third and, we speculate, most importantly, the sample of workers under study has been subject to extreme selection. Many people play golf, but only the very best are professional golfers. Even among professionals, PGA Tour players are among the elite. It is quite possible that an important selection criterion is the ability to avoid the influences of playing partners. Relatedly, successful, skilled workers may have chosen over the course of their life to invest in human capital whose productivity is not dependent on social spillovers, whether positive or negative, in order to avoid risks out of their control. The results described at the end of the previous section are suggestively consistent with this view. Even among the highly selected group of professional golfers, the least skilled are the only ones whose productivity respond positively to the composition of their peers. We view this as an interesting finding because it suggests that there is a great deal of heterogeneity across individuals in their susceptibility to social influences in the workplace. It is an open question whether professional golfers are rare exceptions or representative of a larger class of high-skill professional workers.

This conclusion also implies that workers may sort across firms according to the potential importance of peer effects. In settings where positive peer effects—such as the learning story described above—are potentially important, we should expect to see workers who respond relatively well to these learning opportunities. Ignoring such market-induced sorting can lead to misleading generalizations about the importance of social effects at work.

Though our results are different than those found in recent studies of peer effects in the workplace, we view our results as complementary. There is much to learn from the differences in findings. Primarily, our results suggest that there is heterogeneity in the importance of peer effects, both across individuals and across settings. Sorting of workers across occupations according to how they are affected by social pressures and other social spillovers, whether positive or negative, is likely

an important feature of labor markets. By focusing solely on occupations with low skill requirements, the existing studies miss this rich heterogeneity. Perhaps just as importantly, we show that peer effects are not important in a setting with strong financial incentives. Similarly, Peter Arcidiacono and Sean Nicholson (2005) find no evidence of peer effects among medical school students in their choice of specialty or their medical board exam scores, both of which carry large financial returns. These findings suggest that social effects may be substitutes for incentive pay, and are consistent, at least in spirit, with the work of List (2006) and Levitt and List (2007) who argue that the expression of social preferences is likely to vary according to whether behavior is observed in a market setting, the strength of incentives in that setting, and the selection of subjects that researchers observe. We hope our findings will spur other researchers to further explore this heterogeneity in peer effects in the workplace.

Appendix A: Player Categories and Random Assignment

Based on their past performance over their career, players are placed in one of four categories: 1, 1A, 2, and 3. The categories are assigned using the following rules:

- Category 1 players are tournament winners, and the top 25 money winners from the previous year, and PGA Tour life members (e.g., Tiger Woods, Phil Mickelson, and Ernie Els).
- Category 1A players include former champions of the four majors and The Players Championship (e.g., former British Open champions John Daly and Nick Faldo), as well as tournament winners who no longer qualifiy for Category 1 and who played in five or more PGA Tour events in the prior year.
- Category 2 players include those in the top 125, players with 50 or more career cuts made, and players in the top 50 of the World Golf Rankings.
- Category 3 includes all others, such as local qualifiers. These golfers get the first and last tee times of each session. When Annika Sorenstam became the first woman in 58 years to play in a PGA Tour event, she was a Category 3 player and was assigned the earliest and latest tee times in her two rounds.

Players might be paired with players from a different category if the number of players in a given category is not a multiple of three. In that case, Category 1A players are paired with Category 1 players and Category 3 players are paired with Category 2 players.

Categories are assigned at the beginning of the season, and, for the most part, the category assignments are static. However, if a player wins a tournament or enters the top 25 money list, then that player can be "promoted" to Category 1 status during the season. Likewise, if a player drops out of the top 50 World Golf Ranking and does not satisfy the other definitions of a Category 2 player, then that player is 'demoted' to Category 3 status during the season. We are never able to directly observe players changing categories in our data, however we infer category changes based on observed assignments. In the 2002 season, we have two "snapshots" of category status directly from the PGA Tour (at the beginning of the season and halfway

through the season). In the 2005–2006 seasons, we were not able to get the category status of the players on the tour, but we used the definitions above to assign players as best we could at the beginning of the season, and then we used a probabilistic matching algorithm to assign the remaining players. We tested the matching algorithm on the 2002 season (where we had the categories given to us directly from the PGA Tour), and we verified that we got more than 99 percent of category assignments correct.

The algorithm works as follows. We start with a list of players for whom we are sure we know the category status throughout the season (this is most obvious for elite players, former major champions, and former Nationwide Tour players). Then we look at every playing partner of those players during the season and assign the playing partners to the same category. For players who get matched to different categories, we flag them and manually decide which category they belong in.

Using these categories, we test for random assignment by tournament, and we drop the following tournaments which fail the test for random assignment: The Masters, US Open, British Open, PGA Championship, Walt Disney Championship, Tour Championship, Players Championship, Mercedes Championship, and all of the World Golf Championship events.

Appendix B: Measurement-Error-Corrected Estimator

This section describes the estimator of Fuller and Hidiroglou (1978) which corrects for attenuation bias of a known (estimated) form. The baseline model from the main text is reproduced below (with the fixed effects written out since they will be estimated as parameters below):

$$\begin{aligned} Score_{iktr} = \; & \alpha_1 + \beta_1 \, Ability_i + \gamma_1 \, \overline{Ability}_{-i,kt} + \cdots \\ & + \delta_{1,1} \, d_{1,1} + \delta_{1,1A} \, d_{1,1A} + \delta_{1,2} \, d_{1,2} + \delta_{1,3} \, d_{1,3} \\ & + \delta_{2,1} \, d_{2,1} + \cdots + \delta_{T,3} \, d_{T,3} + e_{iktr}, \end{aligned}$$

where $d_{t,k}$ is a dummy variable that is 1 in tournament t and category k, and 0 otherwise, and $\delta_{t,k}$ is the coefficient on the dummy variable (the fixed effect for the given tournament-category).

$Ability_i$ and $\overline{Ability}_{-i,kt}$ are measured with error because individual ability is constructed by sampling past golf scores (where the number of past scores varies by individual). Assume that past golf scores are independently and identically distributed unbiased measures of permanent ability, and that for each player N_i scores have been used to estimate ability and that these past scores have sample variance $\hat{s}_i^2$ (so that the sample variance of estimated ability is $\hat{s}_i^2/N_i$). Furthermore, assume that noise in past scores is uncorrelated across players, so that the variance in the estimate of my playing partners' average ability is proportional to the sum of the

variance of each playing partner's estimated ability. Under these assumptions, the measurement-error-corrected estimator for the parameters is the following:

$$\hat{\boldsymbol{\beta}}_{\mathbf{ME-CORRECTED}} = \left(\frac{\mathbf{X}'\mathbf{X}}{N} - \hat{\boldsymbol{\Sigma}}\right)^{-1}\left(\frac{\mathbf{X}'y}{N}\right),$$

where

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= [\beta_1 \ \gamma_1 \ \delta_{1,1} \ \delta_{1,1A} \ \delta_{1,2} \ \cdots \ \delta_{T,2} \ \delta_{T,3}] \\
\mathbf{X} &= [ABILITY \ \overline{ABILITY} \ d_{1,1} \ d_{1,1A} \ d_{1,2} \cdots d_{T,2} \ d_{T,3}] \\
y &= SCORE \\
N &= \text{row}(\mathbf{X}) \\
\hat{\boldsymbol{\Sigma}} &= \frac{1}{N}\sum_{iktr=1}^{N} \hat{\mathbf{V}}_{iktr}
\end{aligned}$$

$$\hat{\mathbf{V}}_{iktr} = \begin{bmatrix}
\frac{\hat{s}_i^2}{N_i} & 0 & 0 & \cdots & 0 \\
0 & \left(\sum_{j\neq i}\frac{\hat{s}_j^2}{2^2 N_j}\right) & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}.$$

We compute the standard errors using the following formula for the variance-covariance matrix:

$$\begin{aligned}
\hat{\mathbf{V}} &= \frac{1}{N}\left(\frac{\mathbf{X}'\mathbf{X}}{N} - \hat{\boldsymbol{\Sigma}}\right)^{-1} \hat{W}\left(\frac{\mathbf{X}'\mathbf{X}}{N} - \hat{\boldsymbol{\Sigma}}\right)^{-1} \\
\hat{W} &= \frac{1}{N}\sum_{iktr=1}^{N}\left[\left(x'\hat{\varepsilon} - \hat{\mathbf{V}}_{iktr}\hat{\boldsymbol{\beta}}\right)\left(x'\hat{\varepsilon} - \hat{\mathbf{V}}_{\iota ktr}\hat{\boldsymbol{\beta}}\right)'\right] \\
\hat{e} &= y - x\hat{\boldsymbol{\beta}}.
\end{aligned}$$

We use a weighted version of this estimator using the inverse of the sample variance of each player's estimated ability as the weights, mirroring the same weights used in the baseline specification. See Fuller (1987) for a discussion on choosing appropriate weights, which will generally be related to the variances of the measurement error in the model.

REFERENCES

Alexander, Donald L., and William Kern. 2005. "Drive for Show and Putt for Dough? An Analysis of the Earnings of PGA Tour Golfers." *Journal of Sports Economics*, 6(1): 46–60.

Allport, Floyd Henry. 1924. *Social Psychology*. Boston: Houghton Mifflin.

Arcidiacono, Peter, and Sean Nicholson. 2005. "Peer Effects in Medical School." *Journal of Public Economics*, 89(2–3): 327–50.

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

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2009. "Social Incentives in the Workplace." http://ftp.iza.org/dp4190.pdf.

Benabou, Roland, and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics*, 117(3): 871–915.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. "Network Effects and Network Cultures." *Quarterly Journal of Economics*, 115(3): 1019–55.

Boozer, Michael A., and Stephen E. Cacciola. 2001. "Inside the 'Black Box' of Project Star: Estimation of Peer Effects using Experimental Data." Yale Economic Growth Center Discussion Paper 832.

Bronars, Stephen, and Gerald Oettinger. 2008. "Giving 110 percent and Going for Broke: The Effort and Risk-Taking Responses of Professional Golfers to Tournament Incentives." Unpublished.

Brown, Jennifer. 2008. "Quitters Never Win: The Adverse Incentive Effects of Competing with Superstars." http://www.kellogg.northwestern.edu/~/media/Files/Faculty/Research/ArticlesBookChaptersWorkingPapers/Brown_Quitters_Never_Win.ashx.

Card, David, and Thomas Lemieux. 1996. "Wage Dispersion, Returns to Skill, and Black-White Wage Differentials." *Journal of Econometrics*, 74(2): 319–61.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment." *Quarterly Journal of Economics*, 118(3): 815–42.

Ehrenberg, Ronald G., and Michael L. Bognanno. 1990. "Do Tournaments Have Incentive Effects?" *Journal of Political Economy*, 98(6): 1307–24.

Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics*, 24(1): 39–57.

Festinger, Leon. 1954. "A Theory of Social Comparison Processes." *Human Relations*, 7(2): 117–140.

Fisher, Ronald A. 1935. *The Design of Experiments*. Oliver & Boyd: Edinburgh.

Flynn, Francis J., and Emily T. Amanatullah. 2008. "Psyched Up or Psyched Out? The Influence of Co-actor Status on Individual Performance." Unpublished.

Fuller, Wayne A. 1987. *Measurement Error Models*. John Wiley & Sons: New York.

Fuller, Wayne A., and Michael A. Hidiroglou. 1978. "Regression Estimation after Correcting for Attenuation." *Journal of the American Statistical Association*, 73(361): 99–104.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics*, 111(2): 507–48.

Graham, Bryan S. 2008. "Identifying Social Interactions through Excess Variance Restrictions." *Econometrica*, 76(3): 643–60.

Kroft, Kory. 2008. "Takeup, Social Multipliers and Optimal Social Insurance." *Journal of Public Economics*, 92(3–4): 722–37.

Lazear, Edward, Ulrike Malmendier, and Roberto Weber. 2006. "Sorting in Experiments with Application to Social Preferences." National Bureau of Economic Research Working Paper 12041.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent. 2009. "Performance Pay and Wage Inequality." *Quarterly Journal of Economics*, 124(1): 1–49.

Levitt, Steven D., and John A. List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Tell Us About the Real World?" *Journal of Economic Perspectives*, 21(2): 153–74.

List, John A. 2006. "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions." *Journal of Political Economy*, 114(1): 1–37.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60(3): 531–42.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112–45.

Moffitt, Robert. 2001. "Policy Interventions, Low-Level Equilibria, and Social Interactions." In *Social Dynamics*, ed. Steven N. Durlauf and H. Peyton Young, 45–82. Cambridge: MIT Press.

Orszag, Jonathan M. 1994. "A New Look at Incentive Effects in Golf Tournaments." *Economic Letters*, 46(1): 77–88.

Pessin, Joseph. 1933. "The Comparative Effects of Social and Mechanical Stimulation on Memorizing." *American Journal of Psychology*, 45(2): 263–70.

Rosenbaum, Paul R. 2002. "Attributing Effects to Treatment in Matched Observational Studies." *Journal of the American Statistical Association*, 97(457): 183–92.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116(2): 681–704.

Triplett, Norman. 1898. "The Dynamogenic Factors in Pacemaking and Competition." *American Journal of Psychology*, 9(4): 507–33.

Zajonc, Robert B. 1965. "Social Facilitation." *Science*, 149(3681): 269–74.

Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics*, 85(1): 9–23.

This article has been cited by:

1. Ángel Martín-Román, Alfonso Moral, Marta Martínez-Matute. 2015. Peer effects in judicial decisions: Evidence from Spanish labour courts. *International Review of Law and Economics* **42**, 20-37. [CrossRef]
2. Samuel Smithers. 2015. Goals, motivation and gender. *Economics Letters* . [CrossRef]
3. Theresa Beltramo, Garrick Blalock, David I. Levine, Andrew M. Simons. 2015. Does Peer Use Influence Adoption of Efficient Cookstoves? Evidence From a Randomized Controlled Trial in Uganda. *Journal of Health Communication* **20**, 55-66. [CrossRef]
4. Olga Yakusheva, Jason Fletcher. 2015. Learning from Teen Childbearing Experiences of Close Friends: Evidence using Miscarriages as a Natural Experiment. *Review of Economics and Statistics* **97**, 29-43. [CrossRef]
5. Fangwen Lu, Michael L. Anderson. 2015. Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups. *Journal of Labor Economics* **33**, 91-122. [CrossRef]
6. Ian M. Schmutte. 2015. Job Referral Networks and the Determination of Earnings in Local Labor Markets. *Journal of Labor Economics* **33**, 1-32. [CrossRef]
7. K. R. Ahern, R. Duchin, T. Shumway. 2014. Peer Effects in Risk Aversion and Trust. *Review of Financial Studies* **27**, 3213-3240. [CrossRef]
8. Ezra Golberstein, Daniel Eisenberg, Marilyn F. Downs. 2014. SPILLOVER EFFECTS IN HEALTH SERVICE USE: EVIDENCE FROM MENTAL HEALTH CARE USING FIRST-YEAR COLLEGE HOUSING ASSIGNMENTS. *Health Economics* n/a-n/a. [CrossRef]
9. Joshua D. Angrist. 2014. The perils of peer effects. *Labour Economics* **30**, 98-108. [CrossRef]
10. Jamie Emerson, Brian Hill. 2014. Gender Differences in Competition: Running Performance in 1,500 Meter Tournaments. *Eastern Economic Journal* . [CrossRef]
11. Serkan Ozbeklik, Janet Kiholm Smith. 2014. Risk taking in competition: Evidence from match play golf tournaments. *Journal of Corporate Finance* . [CrossRef]
12. Amanda L. Griffith, Kevin N. Rask. 2014. Peer Effects in Higher Education: A look at heterogeneous impacts. *Economics of Education Review* . [CrossRef]
13. Daniel Eisenberg, Ezra Golberstein, Janis L. Whitlock. 2013. Peer Effects on Risky Behaviors: New Evidence from College Roommate Assignments. *Journal of Health Economics* . [CrossRef]
14. Daniel Eisenberg, Ezra Golberstein, Janis L. Whitlock, Marilyn F. Downs. 2013. SOCIAL CONTAGION OF MENTAL HEALTH: EVIDENCE FROM COLLEGE ROOMMATES. *Health Economics* **22**:10.1002/hec.v22.8, 965-986. [CrossRef]
15. Muzhe Yang, Rui Huang. 2013. Asymmetric Association between Exposure to Obesity and Weight Gain among Adolescents. *Eastern Economic Journal* . [CrossRef]
16. JOSEPH PRICE, LARS LEFGREN, HENRY TAPPEN. 2013. INTERRACIAL WORKPLACE COOPERATION: EVIDENCE FROM THE NBA. *Economic Inquiry* **51**:10.1111/ecin.2013.51.issue-1, 1026-1034. [CrossRef]
17. Ryuichi Tanaka, Kazutoshi Ishino. 2012. Testing the incentive effects in tournaments with a superstar. *Journal of the Japanese and International Economies* **26**, 393-404. [CrossRef]
18. Scott E. Carrell, Mark Hoekstra. 2012. Family Business or Social Problem? The Cost of Unreported Domestic Violence. *Journal of Policy Analysis and Management* n/a-n/a. [CrossRef]
19. Francis J. Flynn, Emily T. Amanatullah. 2012. Psyched Up or Psyched Out? The Influence of Coactor Status on Individual Performance. *Organization Science* **23**, 402-415. [CrossRef]

20. Andries De Grip, Jan Sauermann. 2012. The Effects of Training on Own and Co-Worker Productivity: Evidence from a Field Experiment. *The Economic Journal* no-no. [CrossRef]
21. R. Chetty, J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, D. Yagan. 2011. How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics* **126**, 1593-1660. [CrossRef]
22. Jennifer Brown. 2011. Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy* **119**, 982-1013. [CrossRef]
23. Mir M. Ali, Aliaksandr Amialchuk, Francesco Renna. 2011. Social Network and Weight Misperception among Adolescents. *Southern Economic Journal* **77**, 827-842. [CrossRef]
24. Jorg Franke. 2011. The incentive effects of levelling the playing field - an empirical analysis of amateur golf tournaments. *Applied Economics* 1-8. [CrossRef]
25. Robert Bifulco,, Jason M. Fletcher,, Stephen L. Ross. 2011. The Effect of Classmate Characteristics on Post-Secondary Outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy* **3**:1, 25-53. [Abstract] [View PDF article] [PDF with links]
26. Devin G. Pope,, Maurice E. Schweitzer. 2011. Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes. *American Economic Review* **101**:1, 129-157. [Abstract] [View PDF article] [PDF with links]
27. Lawrence E. Blume, William A. Brock, Steven N. Durlauf, Yannis M. IoannidesIdentification of Social Interactions 853-964. [CrossRef]
28. Eric D. Gould, Todd R. Kaplan. 2010. Learning unethical practices from a co-worker: The peer effect of Jose Canseco. *Labour Economics* . [CrossRef]
29. Giacomo De Giorgi,, Michele Pellizzari,, Silvia Redaelli. 2010. Identification of Social Interactions through Partially Overlapping Peer Groups. *American Economic Journal: Applied Economics* **2**:2, 241-275. [Abstract] [View PDF article] [PDF with links]
30. Carrell Scott E., Hoekstra Mark L.. 2010. Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics* **2**:1, 211-228. [Abstract] [View PDF article] [PDF with links]
31. C. Kirabo Jackson,, Elias Bruegmann. 2009. Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics* **1**:4, 85-108. [Abstract] [View PDF article] [PDF with links]