TI 2015-115/VII
Tinbergen Institute Discussion Paper

Matching and Winning? The Impact of Upper and Middle Managers on Team Performance

Thomas Peeters[1]
Steven Salaga[2]
Matthew Juravich[3]

[1] *Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, the Netherlands;*
[2] *Texas A&M University, United States;*
[3] *University of Akron, United States.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Matching and Winning? The Impact of Upper and Middle Managers on Team Performance

THOMAS PEETERS, ERASMUS SCHOOL OF ECONOMICS, ERIM & TINBERGEN INSTITUTE

STEVEN SALAGA, TEXAS A&M UNIVERSITY

MATTHEW JURAVICH, UNIVERSITY OF AKRON

THIS VERSION: OCTOBER 2015

Corresponding author: Thomas Peeters. E-mail: peeters@ese.eur.nl. Postal address: Erasmus School of Economics room H12-22, Postbus 1738, 3000 DR Rotterdam, Netherlands. Other author e-mails: salaga@hlkn.tamu.edu (Steven Salaga), mjuravich@uakron.edu (Matthew Juravich). The authors would like to thank Alan Benson, Christophe Boone, Jan Bouckaert, Charles Brown, Tine Buyl, Josse Delfgaauw, Amanda Goodall, Brian Mills, Andrew Oswald, Stefan Szymanski, Jason Winfree, Nick Zubanov, and seminar/conference participants at Tilburg University, Erasmus School of Economics, 2015 Eastern Economic Association, SOLE/EALE, NASSM and EASE meetings for feedback and comments. Peeters gratefully acknowledges financial support from the Flanders Research Foundation during part of this research project.

ABSTRACT

In this paper we investigate the impact of upper and middle level managers on firm performance by simultaneously estimating manager and match qualities for management pairings in Major League Baseball (MLB). We document the economic significance of managers at both organizational levels and illustrate the importance of accounting for match quality in evaluating manager impact. Our results suggest assortative matching as managerial quality is positively correlated across organizational levels. Higher match quality in a pairing is further associated with higher individual manager qualities and longer joint employment spells. Mismatches in educational attainment are linked to lower match quality.

JEL-codes: M12, M54, L83

Keywords: match quality, management, team performance, Major League Baseball

1 Introduction

What is the impact of managers at different levels in the organizational hierarchy on firm policies and firm performance? This question has been at the heart of the management and personnel economics literature for decades. As a result, the evidence is mounting that both upper level managers, such as CEOs and CFOs, and middle level managers, such as supervisors and branch managers, have an economically relevant influence on firm practices and performance.[1] Furthermore, observable characteristics such as education, experience as a worker, reputation and personality traits have been linked to managerial performance at both the upper and middle level.[2] However, firm performance cannot be explained as the simple addition of inputs and managerial abilities across organizational levels. The quality of cooperation among managers at different levels and their interaction with the organizational culture (i.e., the "match" quality between managers and firms) is a factor that may either strengthen or weaken the contribution of management to production. Given that the empirical importance of match quality has already been shown in a variety of related settings, such as matching between CEOs and firms (Tervio, 2008), manager risk preference and firm compensation policies (Bandiera et al., 2015) and supervisors and workers (Lazear et al., 2015), it is clear that it cannot be neglected in the analysis of managerial impact along the firm hierarchy.[3]

In this paper we present a first attempt to gauge the importance of match quality between managers across hierarchical layers. To achieve this, we exploit a rich dataset on upper and middle managers in the North American professional baseball industry (i.e., Major League Baseball or MLB). As argued by Kahn

[1] In the management literature a seminal paper examining the impact of personal characteristics on firm policies is Hambrick and Mason (1984). See Hambrick (2007) for a review of the further development of this 'upper echelons' literature and Wooldridge et al. (2008) for a survey of the strategy literature on middle management. Bertrand and Schoar (2003) present the seminal econometric analysis of upper level manager effects on firm policy and performance. Recent analyses on game developers (Mollick, 2012), supervisors (Lazear et al, 2015), store managers (Siebert and Zubanov, 2010), operations managers (Hendricks et al, 2014) and car sales managers (Owen et al., 2015) document similar evidence for managers at lower organizational levels.

[2] For education, a positive correlation is usually found (Chevalier and Ellison, 1999, Goldfarb and Xiao, 2011), with the exception of Mair (2005). Among others, Goodall and Poterba (2014) and Goodall et al. (2011), document the positive impact of high quality worker experience on managerial leadership. See Kaplan et al. (2012) for a study on the impact of personality traits, such as empathy and directness, Falato et al. (2015) for an assessment of the importance of CEO reputation and Malmendier et al. (2011) for evidence on the impact of birth cohorts on managerial policy making.

[3] The management literature has not adopted the term "match quality" to the same extent as it is used in personnel economics. Still, in our view similar (theoretical) concepts have been developed. Crocker and Eckardt (2014), for example, stress complementarities between worker and managerial human capital as a source of value creation in MLB, an idea closely related to worker-manager match quality. Other papers, such as Raes et al. (2011), specifically recognize the importance of the interaction between top and middle management in firm performance.

(2000), MLB data is particularly appealing in this context, as this industry has (a) an abundance of firm input (e.g., player wage data) and output measures (wins, results), (b) freely available personal characteristics for managers at all organizational levels and (c) plenty of manager dismissals, which ensures high turnover and hence facilitates the identification of manager effects. On top of this, MLB teams operate under a relatively uniform organizational hierarchy. Each organization has a clearly identified middle manager, called the coach or manager[4], who is responsible for the day-to-day management of the team, and an upper level executive, called the general manager (GM), who decides on longer term strategy issues. Given the structural homogeneity of organizations and the context-unique skills and expertise of managers, individuals are often rehired in similar roles across different organizations within the industry, resulting in numerous co-worker matches. Related, MLB managers perform largely similar tasks (e.g., planning, strategy, and recruitment) and take on comparable roles (e.g., communicator, motivator, etc.) as executives in many other industries, which broadens the applicability of our results.

Our first goal in this paper is to investigate the empirical importance of match and manager quality. To this end, we build an econometric model to explain team performance on the field (via wins and score difference) as a function of managerial talent (GMs and coaches), firm effects, match qualities, and other team inputs (players and stadiums). We contrast results from two procedures, one which incorporates and one which ignores match quality. We find that accounting for match quality significantly reduces the assessed heterogeneity in manager ability for both upper and middle managers. For example, the estimated standard deviation of added win percentage attributable to GMs decreases from 5.8% to 1.4%, whereas for coaches we find a reduction from 5.5% to 1.4%. At the very top of the managerial talent distribution, elite managers improve team performance by approximately 4% over their peer-group average.[5] These estimates place the impact of both managerial levels on par with widely accepted estimates of the impact star players generate for their team (Cameron, 2014). At the same time, a one standard deviation improvement in match quality is estimated to imply an additional 1% increase in win

[4] Unlike in other sports, in baseball, the head coach is called the "manager." To avoid any confusion between the manager and general manager, we refer to the "manager" as the "head coach" or "coach" from here out. However, when simultaneously referring to both management levels, we use the term "managers."

[5] Further support for our manager estimates comes from the fact that the top managers we identify are consistent with other research and popular opinion on baseball manager abilities. For example, in the case of GMs we find support for the notion that the 'moneyball' managers (see Lewis, 2003) have outperformed their peers, as also argued by Hakes and Sauer (2006) and Wolfe et al (2006). Our results in this respect diverge however from Goff (2013), which we suspect is due to large differences both in inputs and methodology.

probability. In other words, even though match quality is an empirically and economically significant phenomenon, managerial quality still produces a significant impact on performance.

In a second step, we analyze the relationships between manager characteristics, firm characteristics and match quality. We document a positive association between manager and firm quality as well as between managerial qualities across hierarchical organizational levels. As such, our results confirm the recent findings of Bandiera et al. (2015) and Spanos (2015), who respectively report evidence of assortative firm-worker and co-worker matching in the managerial labor market. Next, we illustrate that the value of match quality itself is positively correlated to the quality of both the upper and middle manager in the pairing. In other words, high quality managers tend to work with high quality co-workers and are also more likely to obtain higher match qualities with their co-workers. In terms of matching on observable characteristics, our results support the notion that a large mismatch in education leads to lower match quality between upper and middle managers. Even though work experience, management experience and age are associated with manager performance, we do not find any performance effect for matching across management levels on these characteristics.[6] A final result relates match and manager quality to spell duration. As expected, we find spells with low match quality to be shorter lived. Likewise, spells with lower estimated middle manager quality end sooner. This result, however, is not replicated at the upper level. This suggests not only that hiring and firing decisions at the middle level are decided directly at the upper level, but also that firm ownership is less likely to make a management change at the upper level based on low management quality.

In the next section we explain the role of managers in baseball organizations and outline our performance model. Then we focus on the empirical implementation including the identification of manager and match qualities. The fourth section discusses our results and provides a host of robustness checks while the final section offers our conclusion.

[6] In line with earlier studies by Goodall et al. (2011) on coaches in the National Basketball Association, we find that former star players are more likely to be better coaches.

2 MODEL

2.1 THE ROLE OF TOP AND MIDDLE MANAGERS IN TEAM PRODUCTION

Sport organizations function as a result of the interpersonal relationships that exist between executive management, coaches and athletes. In MLB, organizational goals mirror those of non-sport organizations as teams pursue financial success as well as competitive success on the field-of-play. At the top level, an owner or ownership group is responsible for establishing a formal hierarchy of employees to serve as a chain of command for decision making, strategy development and implementation.

In a typical setting, a GM is hired to oversee and facilitate the hiring and firing of coaches and athletes based upon the extent to which individual abilities or performance align with the greater goals of the organization. GMs typically serve as upper-level managers with respect to payroll allocation and personnel decision making in their roles as executives. In MLB (as with other professional leagues) teams are engaged in a continuous cycle of financial and competitive performance improvement. For example, one strategic approach to improvement could be focused on making changes to player resources in hopes of enhancing the mix of abilities of the athletes comprising the team.

Another could be centered upon changing strategic direction via the replacement of a team's head coach and coaching staff (i.e., middle management). For example, if a team hires a coach who was a former player and thus, demonstrates unique expertise based on his playing experience, the GM should work with the coach to compile a roster of appropriately skilled players to support the strategic approach employed by the coach. As baseball middle managers (i.e., head coaches) are granted autonomy in organizing their human resources in such a manner to pursue competitive success in the most efficient manner possible, an optimum fit would be represented by a roster populated with a number of players who demonstrate attributes which complement the strategies identified by both middle and upper management. Thus, both GMs and coaches occupy key managerial roles which directly impact strategy formulation and implementation, financial investment and human resource management through their actions as employees. As such, we argue that these individuals occupy management roles in the same way that upper and middle management direct employees and impact organizational performance in non-sport industries.

We also note the parallel between professional sport and other industries regarding the role of management in hiring a support staff. In MLB, the GM is responsible for the selection and evaluation of personnel in the areas of baseball operations, player development, scouting and analytics. The total number of employees in these areas traditionally range from 20-60. Likewise, the head coach (i.e., middle manager) is responsible for hiring a staff of fellow coaches including a bench coach and hitting and pitching specialists. Head coaches traditionally supervise a staff of approximately ten coaches. Consequently, by modeling the impact of GMs and coaches on team performance, we are implicitly accounting for the quality of the support staff that the managers in both hierarchical layers hire.

2.2 A SIMPLE MODEL OF TEAM PRODUCTION IN BASEBALL

We now provide a simple model of team production, which we interpret as the production of on-field performance. We assume that in each game between teams i and j, clubs use four inputs, labor (L_{it}), capital (K_{it}), fixed firm-specific factors (ω_i) and management. We explicitly separate managerial from labor inputs, such that in our setting labor only refers to the workers (players) the team hires in season t. In terms of capital, the home stadium is by far the most important input MLB teams use. Fixed firm-specific factors include elements such as team history, fan loyalty and organizational culture, which may all influence the production of on-field performance from inputs. Note that this specification implies that we model firm productivity largely as a complement to managerial inputs, and not the channel through which management influences production. Management inputs encompass the managerial contribution to team production both at the middle level (μ_i), i.e. the head coach, and upper level (υ_i), i.e. the general manager or GM. As explained above, not every pairing between a team and two managers of fixed quality need be equally productive. We account for this by allowing the spell-specific match quality ($\sigma_{im_iu_i}$) between the team (i) and its current coach (m_i) and GM (u_i) to impact the production of on-field results.

Additionally, one team in each contest enjoys a "home advantage" (γ_i), because every competitive MLB game is played at one team's home stadium. The presence of home advantage has been documented across a wide range of sports, and appears remarkably stable over time, in particular in baseball (see Pollard and Pollard, 2005). Alternative explanations for this effect include the bias from fans towards the home team, less fatigue from traveling (Pollard and Pollard, 2005) and even referee bias (Garicano et al., 2005).

Given the zero-sum nature of sports competition, it is common to model game results as a Tullock-type contest where each team's input use matters only in a relative sense, i.e. in relationship to its opponent's. To account for this, Peeters and Szymanski (2014) suggest to estimate the fraction of two team production functions, which for team i takes the following form in our application,

$$\tilde{y}_{it} = L_{it}^{\beta_l} K_{it}^{\beta_k} \exp(\gamma_i \omega_i \mu_i \upsilon_i \sigma_{im_i u_i}) \tag{1}$$

Observe that (1) is equivalent to a Cobb-Douglas type function where managerial and team effects enter multiplicatively. The impact of capital and labor are measured by the respective return parameters β_k and β_l. Rewriting this model in logs and adding an error term (ε_{ijt}), which can be interpreted to capture the chance factors inherent in any sports competition, yields a model with additive productivities for managers and teams,

$$y_{ijt} = \gamma_i - \gamma_j + \beta_l(l_{it} - l_{jt}) + \beta_k(k_{it} - k_{jt}) + \omega_i - \omega_j + \mu_i - \mu_j + \upsilon_i - \upsilon_j + \sigma_{im_i u_i} - \sigma_{jm_j u_j} + \varepsilon_{ijt}. \tag{2}$$

In our baseline specification we consider two indicators of on-field performance (y_{ijt}), namely the simple win/loss result and the run differential between the competing teams.[7]

3 EMPIRICAL IMPLEMENTATION

We implement the production model outlined above in a dataset of MLB games played in the 1989 through 2012 seasons, including both regular season and playoff games. Before explaining the way we proxy inputs and outputs empirically, we explain the different approaches implemented to estimate the managerial and match quality parameters in equation (2).

3.1 IDENTIFYING THE IMPACT OF MANAGERS AND MATCH QUALITY

In order to highlight the importance of match quality between managers across organizational layers, we compare two empirical strategies to identify the manager and match quality parameters in equation (2). The first approach is dubbed “Fixed Effects” in the remainder of this paper, as it relies on manager and firm dummies in the spirit of the seminal paper by Bertrand and Schoar (2003). The second method, referred to as the “Random Effects” approach, adapts the two step methodology developed in Jackson (2013) to estimate the productivity of teacher-school matches. We detail both approaches below.

3.1.1 Fixed Effects (FE) Approach

Following the seminal papers of Abowd et al. (1999) and Bertrand and Schoar (2003), empirical researchers have frequently introduced manager and firm fixed effects to identify the contribution of individual managers to firm policies and performance. As we are interested in the effect of individual

[7] Run differential accounts for the relative closeness of a contest between two competing teams.

heterogeneity at two separate levels of management, a natural extension of this methodology is to introduce three types of dummies to estimate the individual effect of each firm (ω_i), middle (μ_i) and upper-level manager (υ_i). In this case, it is not feasible to simultaneously identify a dummy for each coach-GM-team combination, i.e. to estimate match value ($\sigma_{im_iu_i}$). Match values can only be inferred from the average residual of a pairing.

The separate identification of upper management, middle management and firm fixed effects hinges on individuals moving across different management-firm pairings in the dataset. In our case, this implies that we require variation across two dimensions. First, each firm has to be paired with at least one manager who is also employed by another firm in the data. The "moving" managers allow identification of the time-invariant firm component, which can then be used in the identification of all non-moving managers. Second, each upper (middle) manager has to be paired with at least one middle (upper) manager who has worked with another upper (middle) manager. In other words, we can only include a dummy for a coach if he has worked with at least one GM who has worked with at least one other coach. Likewise, we must disregard every GM who has only worked with coaches which have only worked with him. As is the case for firms without moving managers, we simply lack the necessary variation to disentangle the impact of both levels of management for these observations.

Additionally, all manager and firm fixed effects are estimated with respect to a comparable reference i.e. those managers (and firm) for which there will be no dummy in the regressions. As argued by Abowd et al. (1999), firms which are connected in a network through moving managers automatically share a common reference; however this is not the case across all such networks. In our dataset, all firms are connected into the same network via moving coaches, whereas for GMs, there are two separate networks.[8] To arrive at comparable estimates, we repeat the fixed effect procedure using each possible pairing of firms across both GM networks as the reference and report the average effects across all estimates.[9]

In our application, a natural choice to serve as the reference category are the so called "caretaker" or interim managers, who did not succeed in securing long term management positions in MLB. Using these individuals as the reference has two distinct advantages. First, it removes the need to identify individual fixed effects from the limited amount of observations we are able to obtain for each of them. As such, we

[8] GM network 1 is Texas, Milwaukee and Cleveland, while network 2 is all other MLB clubs.

[9] The correlations across fixed effects estimates with different reference firms is always above 0.95, so, empirically, this appears to be a fairly minor issue.

reduce the risk of incidental parameter bias. Second, given their inability to obtain a permanent position and the relatively short time period in which they have an ability to impact the firm, these managers are expected to have a modest influence on team production. They are therefore a natural lower bound to measure the impact of managers with established careers.

<Insert Table 1 here>

Table 1 shows the implications of this identification strategy for our sample. We start with a total of 147 separate coaches, 107 GMs and 31 teams which are combined into 345 spells. Slightly above 38% of coaches and 25% of GMs were employed by at least two teams. We consequently drop the three teams (Minnesota, Pittsburgh and Tampa Bay) which lack a moving manager at one of the two levels. Next, we find that 144 coaches (or 98%) and 103 GMs (or 96%) have been paired with a manager at the other level, which has at least one other co-worker in the data. We again exclude the coaches and GMs who fail to meet this requirement. This leaves us with 136 middle-level and 91 upper-level managers[10] in the final dataset. In terms of individual game observations, this sample selection procedure leads the data sample to shrink from an initial value of 110,212 down to 84,534.[11] As shown in Table 1, about 21% of all coaches and 13% of all GMs are sorted into the "caretaker" reference category.[12] However, these managers account for only 1,365 and 1,995 observations, respectively.

3.1.2 Random Effects (RE) Approach

As mentioned above, the fixed effects approach does not allow for the ability to uncover a separate estimate of the match quality term ($\sigma_{im_iu_i}$). Match quality therefore has to be absorbed by the error term (ε_{ijt}) in the regression of equation (2). This in turn forces mean match quality for a given manager or firm to equal 0 across all his pairings. In a hypothetical dataset, where each manager is observed in many different pairings, this assumption need not be problematic, as we might expect good and bad matches to cancel out over the long run. In a typical dataset such as ours however, each manager is observed in at most a handful of pairings. This approach therefore risks to incorrectly attribute a considerable part of the

[10] Of these, one GM observation is in fact a duo, i.e. Mike Flanagan and Jim Beattie of the Baltimore Orioles in the early 2000's.

[11] Note that each game is used twice, once from each team's perspective, such that this means a move from 55,106 to 42,267 independent games.

[12] In our baseline specification, we sort GMs with less than 260 observations (slightly less than two seasons) and coaches with less than 130 observations (slightly less than one season) for their entire career into the interim manager category. Note that a larger amount of data per estimated effect also helps to limit the amount of "shrinkage" in the RE estimates.

match value to manager and firm effects. This may severely bias the estimation results, overstating the importance of both manager and firm effects if these are positively correlated to match quality.[13]

To overcome this potential problem we adapt the two stage estimation technique developed by Jackson (2013). In the first step we introduce a composite spell fixed effect ($\varphi_{im_iu_i}$) to replace all manager, firm and match quality parameters in (2). The first stage equation therefore reduces to

$$y_{ijt} = \gamma_i - \gamma_j + \beta_l(l_{it} - l_{jt}) + \beta_k(k_{it} - k_{jt}) + \varphi_{im_iu_i} - \varphi_{jm_ju_j} + \tilde{\varepsilon}_{ijt}. \tag{3}$$

We sum both spell fixed effects ($\varphi_{im_iu_i}$ and $\varphi_{jm_ju_j}$) and the error term ($\tilde{\varepsilon}_{ijt}$) to create the second step dependent variable (η_{ijt}). In the second step we use maximum likelihood to estimate a mixed linear effects model decomposing this joint spell-error term into the coach, GM, team and match quality effects for team i and j. Therefore, we estimate

$$\eta_{ijt} = \omega_i - \omega_j + \mu_i - \mu_j + \upsilon_i - \upsilon_j + \sigma_{im_iu_i} - \sigma_{jm_ju_j} + \varepsilon_{ijt}. \tag{4}$$

As argued by Jackson (2013), this procedure uncovers the best linear unbiased predictors (BLUPs) of manager, firm and match quality. Intuitively, we take the total variation in manager, firm and match quality into account to form an estimate of the match qualities of managers with a limited number of spells. This replaces the assumption that each individual has a mean match value of zero, which is very likely unrealistic in our application.

Note that this "random effects" (RE) procedure changes the way we attribute variation in the data across different factors, but does not introduce a different identification strategy. Hence, the identification of individual effects still depends on observing moving managers across the two dimensions described above. We therefore apply both procedures to the same data sample. This also renders estimates to be more comparable across both approaches, which facilitates the interpretation of our results. Table 1 shows that out of the 341 original spells, 291 (slightly below 85%) figure in the final first stage estimation. When analyzing match quality in the following sections, we restrict attention to those spells for which we have a sufficient number of observations, both on the GM, coach and the pairing they share.[14] As depicted in Table 1, this leads to an additional 81 spells being disregarded in this analysis.

[13] See Jackson (2013) for a more elaborate explanation of these issues.

[14] Here we use 65 game observations or slightly less than half a season as the cut-off value. Experimentation has shown that results are not sensitive to this choice.

3.2 MEASURING OUTPUTS AND INPUTS

Table 2 shows summary statistics for our final sample. We retain 84,534 observations for which we have all relevant information. It is important to note that each individual game features twice in the data, once from each team's perspective. Throughout the paper we cluster standard errors at the game level to correct for this issue.

We gauge team production using two output measures, game results and run differential. The game result variable is a simple binary measure where we code a loss as 0 and a win as 1. As our dataset is symmetric, the mean and standard deviation of this variable are both equal to 0.5. By the same logic, the run difference variable is completely symmetric and its average is 0. The standard deviation of run differential is slightly above 4, with a maximum difference of 27 runs. For both outcome variables we estimate equation (2) and (3) using linear regression models.

<Insert Table 2 here>

In the economic literature on sports leagues, it is common to assess the labor inputs teams employ by looking at player salaries, rather than measures such as hours worked or number of full time employees (FTEs). On the one hand, researchers often lack accounting or survey data to infer the number of FTEs, whereas payroll data is readily available. On the other hand, leagues impose limits on roster sizes[15] (the amount of players teams are allowed to hire), and likewise the number of players involved in any particular game is fixed by sporting rules. The positive effect of a team's total wage bill on sporting performance is well documented in the literature (e.g., Szymanski, 2003). Moreover, the MLB player market is not regulated by a salary cap/floor arrangement, which gives rise to much larger discrepancies in payroll across teams than in the National Football League, for example.[16] In addition to the total payroll of the team, the intra-team distribution of salaries may play a role in team production. For example, Bloom (1999) and Depken (2000) find that teams with more skewed salary distributions perform worse ceteris paribus than teams with more homogeneous pay structures. These effects may be due to a number of factors, including motivational issues or equity concerns.

[15] Leagues usually limit either the number of players a team may have on its payroll at any point during the season or the number of different players eligible to compete over the entire season.

[16] MLB has a luxury tax, which imposes an additional tax of 22.5% on salary payments above a certain threshold. The threshold for this tax is quite high however, such that only the NY Yankees and Boston Red Sox regularly pay a contribution. Over the last ten years only three other teams have ever paid anything under the tax, and even when affected, the tax represents a very minor share of overall payroll costs.

In Table 2 we report summary statistics for the total payroll of each team expressed in 2009 US$. Even in real terms, total salaries rose over the observation period, with the large market teams (primarily the New York Yankees and Boston Red Sox) spending around $200m per season in recent years. Average total team payroll over the sample period was only a third of this amount at approximately $70m yearly. The average standard deviation of team payroll is approximately $3m. Given that the active roster size for MLB teams is 25 players, this number is roughly comparable in size to the average individual player salary over our sample period.

In terms of capital, a team's main production input is its stadium. To quantify capital use we account for both the capacity and age of the team's stadium. This approach avoids the need to obtain the book or market value of each baseball park, which in almost all cases is unavailable. Furthermore, the limited alternative uses of a baseball park may lead its market value to lie significantly below the production value it has to the team. A baseball park is clearly a long-term investment and clubs have very limited options to adapt their stadium to short-run productivity and demand shocks. Table 2 displays summary statistics for both stadium measures in the dataset. The average stadium has a capacity slightly less than 48,000 and is roughly 28 years old. However, both measures show significant variation with the oldest stadium in use for over a century.

An additional concern when estimating production functions is that firms may choose their inputs conditional on unobservable factors, such as productivity, leading to omitted variable bias. This is particularly true for adjustable inputs, such as labor, which are therefore treated as potentially endogenous. A further concern is that results on the field of play could have a feedback effect on total payroll, for example through performance-related bonuses. To check the robustness of our results with respect to these potential biases we estimate equation (3) with and without instrumenting the wage variables and compare the results. As instruments we employ lagged payroll values, a set of market characteristics for the metropolitan statistical area (MSA) where the team is located, a time trend and a dummy for teams in the American League. Table 2 also presents summary statistics for each of these variables.

<Insert Table 3 around here>

In our empirical analysis, we further ask whether personal characteristics (and matching based on these characteristics) are correlated to a manager's performance and the quality of his matches. We therefore gather data on education, age and working experience both as a player and manager. Table 3 reveals that upper managers are on average more highly educated than middle managers, with more GMs having attended college (86% vs. 71%) and graduated from college (79% vs. 39%). For GMs, we also include a dummy to indicate whether their alma mater was a top 100 institution in the 2014/15 QS rankings, and as such an institution with a strong academic reputation.[17] In contrast, coaches clearly have more playing experience as compared to GMs. Almost all coaches have played professionally (97% vs. 42% of the GMs) and coaches are far more likely to have played at the highest professional level (78% MLB players against 19%). Around a third of all coaches were selected to appear in the MLB All Star Game, signaling they were truly elite players. Only 5.2% of GMs reached this level of playing proficiency during their career. In terms of managerial experience, coaches trail GMs by about 4.5 years on average. Despite this, the mean GM is slightly younger than his coaching counterpart. Finally, around 15% of all coaches have a minority background (Hispanic or African American in almost all cases), while this number stands at only 7.8% for GMs.

4 Empirical Analysis and Results

4.1 Baseline model

4.1.1 Results 1st stage regressions

Table 4 presents a series of estimation results beginning with the standard baseline model including no fixed effects and no wage instruments. In the next four columns, team, coach and GM dummies are sequentially added until all are contained in a single model. The depicted F-test results from these four models are a first indication that each fixed effect layer is significant in impacting team production. The only exception is the non-significant team effects in the fully specified model. This suggests that regardless of specification, both top and middle level management fixed effects matter in the production of team success. Lastly, the sixth and final column only includes fixed effects to capture the unique coach/GM/team spell associated with each management combination. In other words, this is the first stage estimation of equation (3).

<Insert Table 4 around here>

In line with previous studies, estimates on team payroll are positive and significant. This suggests that franchises which spend more on labor experience higher levels of on-field success. Coefficients on the

[17] As a robustness check we varied the threshold, instead including institutions in the top 200 of the QS rankings. This did not alter any of the results described below.

standard deviation of total team payroll are negative and significant. This infers a positive association between winning and reduced dispersion in within-team player salaries. This finding supports previous results for MLB found in Bloom (1999) and Depken (2000). In agreement with existing literature, we find a strong home advantage with the home team variable significant at the 1% level in all models. Stadium capacity is negative and generally significant, signifying that winning is negatively affected by venue size. At first glance this result appears odd. However, the nature of venue construction changed substantially over the examination period as the massive parks built in the 1960s and 1970s were gradually replaced by or renovated into more intimate venues beginning in the mid-1990s. As such, a smaller venue is usually indicative of higher/more recent investments in the team's stadium, rather than the other way around. The point estimates on stadium age are all small and suggest no significant impact on firm performance.

We abstain from an overly detailed discussion of the estimated parameters in the run difference model displayed in appendix Table 1A. In general these results are in concert with the findings in the game result estimations in Table 4. Home advantage and total team payroll are positively associated with run differential. Alternatively, a high standard deviation of total team payroll is negatively associated with run differential. Stadium capacity is generally negative and significant while the impact of stadium age is weak and inconsistent across models.

4.1.2 Robustness checks

In Table 5 we assess the robustness of these results by estimating a host of alternative specifications and then compare the spell effects from the baseline model against the alternative models. The first column compares the non-IV and IV approaches for both the winning percentage and run differential models. In Check 1, we enter the time-varying coach and GM characteristics (age and experience) directly into the baseline model. Check 2 drops playoff games from the sample to eliminate the potential of overestimating the impact of managers that win in the postseason. Check 3 adds interaction terms between stadium characteristics and the home indicator as it is possible that stadium effects have differing impacts at home versus on the road. In Check 4, we drop the intra-team standard deviation of total team payroll, as one might argue that GMs have a degree of control over this input. Lastly, Check 5 directly adds a set of additional inputs related to the market of the franchise – including population, the number of teams in market, and the club's tenure in the market.

<Insert Table 5 around here>

The results in Table 5 indicate that regardless of the exact specification, the spell effects are stable across models. All correlation coefficients are significant at the 1% level, with all but Check 5 displaying values larger than 0.95. This indicates that our estimates are not overly dependent on the exact choice of specification adopted for the baseline model.

4.2 WHAT IS THE IMPACT OF FIRMS, MIDDLE MANAGERS, UPPER MANAGERS AND MATCH QUALITY ON PRODUCTION?

4.2.1 Are spell effects important in evaluating manager impact?

<Insert Table 6 around here>

Table 6 displays the impact of firms, upper management, middle management, and match quality on team production in both the FE and RE methodologies. It is evident that the fixed effects procedure leads to substantially larger estimates of heterogeneity in managerial ability. Moving towards the two stage procedure, even when match effects are not included (see middle panel), the estimated standard deviation is reduced by a factor of around three for each level (5%-6% against 1.5%-2%). Mirroring the findings of Jackson (2013), there is an additional reduction in the standard deviation of the RE estimates once match effects are introduced (right panel). The differences between procedures for the impact on run differential are completely similar to those for winning percentage. In our view, the RE approach leads to added winning percentage estimates that are far more realistic, and furthermore similar in size to the estimated effects of individual players (Cameron, 2014). Note as well that the match quality estimates[18] are more dispersed for the fixed effects approach (2% vs. 1%), but imply a significant impact of match quality on production under both approaches.

<Insert Table 7 around here>

In Table 7 we report correlation coefficients for individual manager and team effects across empirical approaches. Individual coach and GM estimates are correlated at the 1% level in each combination, while team effects are only correlated at significant levels across the RE models. The two RE approaches generate manager estimates that are correlated at 0.96 or greater. This suggests that the difference between approaches is primarily a scaling issue, and does not overly impact on the productivity ordering of managers. In other words, all managers are estimated to have lower impact regardless of their place in the managerial quality distribution.

An underlying assumption for the identification of manager and match effects is that sorting of managers is idiosyncratic to match quality. Lazear et al. (2015) suggest that this assumption may be tested

[18] These are based on average residuals in the FE approach.

by calculating the population average of the mean estimated match effects within each individual. If these means average to zero, the underlying assumption is not violated, whereas significantly positive (negative) means could indicate sorting based on unobserved match quality. Performing this test on our baseline results we find average mean match quality estimates of 0.00028 for GMs, 0.000067 for coaches and 0.00075 for teams. With standard errors of, respectively, 0.00062, 0.00067 and 0.00064, these averages are not significantly different from zero. Hence, we find no violation of this assumption in our application.

4.2.2 Which individual managers generate largest impacts?

We now place our attention on the estimates for individual managers, i.e. their individual random or fixed effects. Table 8 illustrates the top contributors to both team winning percentage and run differential as compared to the mean manager in the industry. It is again clear that the RE and FE approaches produce substantially different estimates of managerial impact. The more conservative RE estimates illustrate that even the most productive coaches have a relatively modest impact on club outcomes in comparison to their peers. In practice, there is relatively little difference among the vast majority of managers in the industry. At the same time Table 8 also reveals that there is a small, elite group of both coaches and GMs who generate large impacts.

<Insert Table 8 around here>

In the middle management ranks, Lou Piniella clearly stands out, generating the largest impacts on both winning percentage and run differential, with 4.09% and 0.36 runs respectively. Mike Scioscia, Jack McKeon and Joe Torre are the only other coaches with estimated impacts in the 3.6% to 4.0% range, whereas no coach outside the top 10 crosses the 3.1% mark. Outside this elite group, firms may find it hard to hire coaches who can deliver appreciable performance gains over an incumbent manager. This may explain why many coach firings are perceived to generate little positive long-term effects.

Among GMs, Brian Cashman is estimated as the most productive manager in both outcomes with an RE estimate of his impact on winning percentage at 5.23%. This implies an additional four wins per season by employing Cashman as opposed to the league average GM. Since an All-Star quality player contributes approximately five wins per season in comparison to a replacement level player, Cashman's worth to a franchise is roughly on par with the value of a star player. However, a conservative estimate of the open market cost of a player of this quality is $5 million per win or $25 million per season. Given that Cashman is one of the highest paid GMs and is reported to earn 'only' $3 million per year, franchises appear to be able to generate a substantial surplus by hiring an elite GM. Along with Cashman, Theo Epstein, Billy Beane and Pat Gillick are also estimated to generate more than 3% in additional wins. Regular followers of MLB will notice that all but Gillick were at the forefront of the sport's adoption of analytical techniques in decision making. In that sense our results are a testimony of the power of quantitative methods to support managerial decision making. Note as well that both Cashman (New York Yankees) and Beane (Oakland A's) still work for their first employers and are both among the highest paid GMs in the industry. It appears therefore that both firms realize the edge they possess by employing these managers. Beane, in particular, is an interesting case as according to our estimates, the A's are one of the most input deprived teams in MLB. This raises the question why none of the large revenue franchises (e.g., the Los Angeles Dodgers) has outbid Oakland for Beane's services.

4.2.3 Are personal characteristics correlated with managerial performance?

In this section we investigate the link between managerial performance and personal characteristics. We follow Goodall et al. (2011) by replacing all spell and manager effects in equation (2) with variables to proxy for the personal characteristics of both the GM and coach, i.e. we estimate

$$y_{ijt} = \gamma_i - \gamma_j + \beta_l(l_{it} - l_{jt}) + \beta_k(k_{it} - k_{jt}) + \omega_i - \omega_j + \beta_m X_{mi} - \beta_m X_{mj} + \beta_x X_{ui} - \beta_x X_{uj} + \varepsilon_{ijt}. \qquad (4)$$

The vector of the upper and middle manager personal traits (X_{mi} and X_{ui}) includes education, playing experience, the log and log square of age and managerial experience, and a dummy to capture whether a manager has a minority background. Both education and playing experience enter as categorical variables, with the levels indicated in Table 3. The interpretation of these estimates is therefore relative to their reference category, which is "non-professional player" for playing experience and "never attended college" for education. We estimate two specifications for each outcome variable, one with and one without instrumenting the payroll variables. We further include team fixed effects in all models as reported in Table 9.

<Insert Table 9 around here>

Experience as a high quality worker has a positive effect for coaches, whereas for GMs we fail to find a clear monotonic relationship between playing quality and managerial performance. Our result on coaches is consistent with the expert leadership hypothesis. This theory maintains that former high ability workers excel in management positions, because they may for example improve productivity by teaching workers more advanced skills.[19] The absence of a clear link for GMs, suggests that at higher levels in the

[19] See Goodall and Progrebna (2015) or Goodall et al. (2011) for more detailed theoretical underpinnings of this hypothesis.

organizational hierarchy specific technical skills matter less for managerial success. Consistent with the 'Peter Principle,' it may not be wise to promote top coaches to the top management levels. [20] An important caveat here is that our specification does not correct for labor market selection. Former All Stars could have access to more promising coaching positions, which makes it easier for them to generate a positive impact on team production.

At both levels our results show a positive association between college education and managerial performance. Even then, the estimated effects for coaches are somewhat less significant and slightly smaller in size than those for GMs. This finding has been documented quite a few times for upper managers,[21] but contradicts some of the scarce evidence on the link between middle management and education presented by Mair (2005).

For GMs, we further find that younger and less experienced managers tend to perform better than their older, more experienced counterparts, albeit at a declining rate. This points again to the rise of the young and inexperienced, yet analytically-minded GMs who entered the industry in the latter half of our sample. We also report (some) positive estimates for GMs with a minority background. This may be the result of barriers to entry in the GM labor market. However, given the limited number of individuals these results are identified from, we shy away from strong conclusions at this point.

4.3 Analyzing match and manager qualities

In this section we focus on analyzing our estimates of match and manager qualities. Throughout this section we restrict our sample to spells without caretaker managers and with at least 65 observations. In doing so we prevent extreme estimation results on very short spells from driving our results. As indicated in Table 1, we end up retaining approximately 61% of all spells. [22] To facilitate the interpretation of estimated coefficients, we also standardize all estimated manager, team and match effects by subtracting the population mean and dividing by the standard deviation.

4.3.1 Are high quality managers matching with high quality co-workers?

A central result in the literature on matching in employer-employee settings is that high quality firms tend to attract high quality workers, i.e. there is assortative matching in the labor market (e.g., Abowd et al., 1999; Bandiera et al., 2015). It seems natural to extend this hypothesis to co-workers and ask whether

[20] See Benson et al (2015) for a recent empirical analysis showing support for the Peter Principle in sales management.
[21] Among others Chevalier and Ellison (1999) show this result for mutual fund managers.
[22] This number is chosen to coincide with about half a season worth of observations. We have tried various (lower and higher) values for this cut-off, which did not lead to significantly different results.

high quality managers tend to match with each other across different hierarchical layers. In order to explore this issue we run the following regressions:

$$\hat{\mu}_m = \alpha_u \hat{\upsilon}_u + \alpha_i \widehat{\omega}_i + \alpha_l \widetilde{l_{it}} + \beta_m X_m + \beta_u X_u + \varepsilon_{miu} \quad (5)$$

$$\hat{\upsilon}_u = \alpha_m \hat{\mu}_m + \alpha_i \widehat{\omega}_i + \alpha_l \widetilde{l_{it}} + \beta_m X_m + \beta_u X_u + \varepsilon_{uim}. \quad (6)$$

In equation (5) and (6) we analyze the data at the level of the individual match between coach m, GM u and team i. The variables $\widehat{\omega}_i$, $\hat{\upsilon}_u$ and $\hat{\mu}_m$ represent the standardized estimates of firm, upper and middle manager qualities, respectively. We further include the log of total player payroll relative to the yearly industry average to assess the quality of the firm workforce. Two vectors of control variables, X_m and X_u, contain observable manager characteristics (education, age, playing and management experience) for the middle and upper manager.

<Insert Table 10 around here>

In Table 10 we report OLS estimates with bootstrapped standard errors for equations (5) and (6). In the top panel we perform a univariate regression of estimated GM effects on coach effects. We then gradually add more controls in the lower panels, which report separate regressions for the coach and GM effects.

The results we obtain for the RE and FE approach differ quite substantially. We find positive and significant results in all RE specifications, which clearly supports the notion of assortative matching among middle and upper managers. This finding does not extend to the FE estimates, as these indicate insignificant or even negative associations between manager effects. The explanatory power for these specifications is also far below the RE models, providing further evidence of a weak association between estimated manager effects. We further confirm the well documented positive association between firm and manager quality, but fail to uncover any significant results on the association between relative worker quality and managerial effects. This seems to indicate that better managers tend to work at more efficient teams, rather than with teams who invest heavily in playing talent. This finding is consistent with our observation that the top GMs and coaches have not necessarily been employed by the large market MLB teams.

The RE results in Table 10 echo the findings in a recent paper by Spanos (2015) which explores assortative matching across organizational layers using administrative data on French workers and firms. Our paper clearly diverges from Spanos (2015) in the way quality is estimated, since he gauges worker and firm quality by employing a FE approach to estimate individual salary regressions without controls for match quality. In our view, both results are best interpreted as complimentary, as we test for assortative matching between co-workers in different hierarchical positions explicitly taking match quality into account, albeit in a much more specific environment than Spanos (2015).

4.3.2 Are high quality managers achieving high match qualities?

In this section we analyze the quality of the cooperation between upper and middle managers within spells. While we have established above that high quality managers tend to cooperate, this does not mean that they are also cooperating more efficiently, i.e. generating higher match quality. To investigate this question we perform the following regression:

$$\hat{\sigma}_{mui} = \alpha_u \hat{v}_u + \alpha_m \hat{\mu}_m + \alpha_i \widehat{\omega}_i + \alpha_l \widetilde{l_{it}} + \beta_m X_m + \beta_u X_u + \varepsilon_{miu}. \quad (7)$$

Here the dependent variable is the standardized estimate of match quality ($\hat{\sigma}_{mui}$) for the pairing between GM u and coach m at team i. We are primarily interested in the estimates of α_m and α_u, i.e. the correlation of match quality and managerial quality. The control variables are the same as featured in equation (5) and (6) above.

<Insert Table 11 around here>

Table 11 shows that all estimates we obtain for α_m and α_u are positive, highly significant and robust to the inclusion of personal characteristics. This clearly supports the idea that higher quality managers are associated with higher spell values. In other words, talented managers not only tend to work with other high quality managers, but they also tend to cooperate better. Table 11 is therefore consistent with the view that managerial skills at different organizational levels are complements, rather than substitutes. The cooperation between talented managers need not result in a destructive war of egos for control over firm policies, strategy, etc. On the contrary, talented managers excel even more when working with other talented managers. An alternative explanation may be that manager skills are positively correlated to matching ability. Following this reasoning, better managers could be simply more capable of recognizing and obtaining (e.g. through successful networking) superior cooperation opportunities.

4.3.3 Is matching on personal characteristics important for performance and match quality?

We now turn our attention towards observable personal characteristics. First, we analyze whether matching on education, playing experience, age and management experience has a significant association with firm performance. In other words, are firms with 'mismatched' pairings of middle and upper managers underperforming with respect to firms employing a heterogeneous manager duo? Based on the existing literature it is not a priori clear which answer we should expect to find. On the one hand, researchers in the labor economics literature have long regarded worker mismatch as contributing to job dissatisfaction, high worker turnover and low wages.[23] On the other hand, researchers in strategic management claim that diversity among a firm's management team benefits firm performance as it could, for example, help to avoid group think.[24] To shed light on this issue, we revisit the estimation of equation (4), and now add measures for the closeness of the match between the upper and middle manager in terms of each observable characteristic described above (M_{mui}):

$$y_{ijt} = \gamma_i - \gamma_j + \beta_l(l_{it} - l_{jt}) + \beta_k(k_{it} - k_{jt}) + \omega_i - \omega_j + \beta_m X_{mi} - \beta_m X_{mj} + \beta_u X_{ui} - \beta_u X_{uj} + \beta_{mu} M_{mui} - \beta_{mu} M_{muj} + \varepsilon_{ijt} \quad (8)$$

For both age and experience we enter the log of the absolute value of the difference between both managers expressed in days. For education, we introduce separate dummy variables for manager pairings which identify the degree of a match by specifying when a match differs by zero, one or two categories.[25] Hence, a college educated GM working with a coach that never attended college, would differ by two categories in education. We take a similar approach for playing experience, although we allow for up to three categories of difference. As before, we estimate specification (8) using both simple OLS and wage instruments for both output measures. The input measures and observable characteristics contained in (l_{it}, k_{it}, X_{mi} and X_{ui}) are the same as those listed above.

<Insert Table 12 around here>

To preserve readability, Table 12 only reports the results for the matching variables. The estimates suggest that a strong mismatch in education (two category difference) is associated with lower firm performance. This result speaks to the idea that managers need to be able to "communicate effectively with each other," which may be harder when educational backgrounds are very distinct. Think for example about the potential conflict in strategic vision between an analytically minded, college-educated GM and a non-educated coach that subscribes to a traditional strategic approach. Despite this, differences in age, working experience and managerial experience are not correlated with lower managerial performance. Note however that we also fail to document any significant positive effect from managerial diversity in contrast to Carpenter (2002).

[23] See e.g. Allen and van der Velden (2001).
[24] See among others Carpenter (2002).
[25] For this purpose, we drop the distinction of a top 100 vs. regular college match.

We examine the robustness of these results by estimating an adapted version of the match quality model in equation (7), adding both the level of, and matching on, observable characteristics (X_{mi} and M_{mui} respectively), i.e.

$$\hat{\sigma}_{mui} = \alpha_u \hat{v}_u + \alpha_m \hat{\mu}_m + \alpha_i \widehat{\omega}_i + \alpha_l \widetilde{l_{it}} + \beta_m X_{mi} + \beta_u X_{ui} + \beta_{mu} M_{mui} + \varepsilon_{ijt}. \quad (9)$$

To avoid multicolinearity issues we enter characteristics individually and include either our estimates of manager and team qualities ($\hat{v}_u$, $\hat{\mu}_m$ and $\widehat{\omega}_i$) or the level of the observable characteristics (X_{mi} and X_{ui}).

<Insert Table 13 around here>

The results of this exercise, shown in Table 13, largely confirm the picture we obtained above. Apart from a strong mismatch on education, no matching on observable characteristics relates significantly to match quality.

4.3.4 Are well matched spells longer?

Finally, we ask whether spells with high match quality result in longer tenures. Intuitively, we would expect well matched managers to have an incentive to sustain their current cooperation. Likewise, a team owner will have little to gain in breaking up a fruitful partnership. To check for this possibility, we regress match duration (d_{miu}) on our estimates of match quality. We insert controls for manager and firm qualities, player payroll and match on observable manager characteristics. This results in estimating equation (10),

$$d_{miu} = \alpha_s \hat{\sigma}_{mui} + \alpha_u \hat{v}_u + \alpha_m \hat{\mu}_m + \alpha_i \widehat{\omega}_i + \alpha_l \widetilde{l_{it}} + \beta_{mu} M_{mui} + \varepsilon_{miu}. \quad (10)$$

Here we measure the total duration of a spell as the log number of games played by team i under the $m-u$ manager pair.[26] We estimate the model in (10) both with standard OLS and as a proportional (Cox) hazard model.

<Insert Table 14 around here>

As shown in Table 14, we find a significant and positive estimate for match quality in all specifications.[27] Our results lend strong support to the hypothesis set out above and as such, helps to validate our estimation of match quality. We further infer that coaching quality correlates strongly to spell duration, yet GM quality does not significantly impact duration. This is an intuitive result, since the ability

[26] We believe this measures spell length more accurately than calendar days, because days during October-March (the MLB off-season) are not comparable to days during the season (April-September).

[27] We also estimate this model excluding spells which are active at the end of our sampling period, but the results are completely similar.

to end a spell typically lies with the GM. Hence, underperforming middle managers are held accountable much faster than upper managers.

5 CONCLUSION

In this paper we present the first attempt to account for match quality in the estimation of the performance contributions of upper and middle level managers. Allowing for match quality reduces the effect estimates for both levels of management, but we still find that management generates a statistically significant and economically important impact on output. Furthermore, match quality between manager pairings and firms turns out to be a significant determinant of firm performance and spell duration. Our results are also consistent with assortative matching in the managerial labor market, as high quality managers tend to match across hierarchical levels and, on average, these pairings also achieve higher match qualities. We finally uncover that large mismatches in education between manager pairs is associated with lower firm performance, which is not the case for a host of other characteristics. We believe our results can inform organizations with respect to their hiring strategies of management teams partnered across hierarchical levels.

6 REFERENCES

Abowd, J. M., Kramarz, F., & Margolis, D. M. (1999). High Wage Workers and High Wage Firms. *Econometrica, 67*(2), 251-333.

Allen, J., & van der Velden, R. (2001). Educational Mismatches versus Skill Mismatches: Effects on Wages, Job Satisfaction and On-the-Job Search. *Oxford Economic Papers, 53*(3), 434-452.

Bandiera, O., Guiso, L., Prat, A., & Sadun, R. (2015). Matching Firms, Managers and Incentives. *Journal of Labor Economics, 33*(3), 623-681. doi:10.1086/679672

Benson, A., Li, D., & Shue, K. (2015). *Do Promotion Tournaments Produce Bad Managers? Evidence of the "Peter Principle" in Sales.* Working paper.

Bertrand, M., & Schoar, A. (2003). Managing With Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics, CXVIII*(4), 1169-1208.

Bloom, M. (1999). The Performance Effects of Pay Dispersion on Individuals and Organizations. *Academy of Management Journal, 42*(1), 25-40.

Cameron, D. (2014, March 4). *The Cost of a Win in the 2014 Off-Season*. Retrieved from FanGraphs: http://www.fangraphs.com/blogs/the-cost-of-a-win-in-the-2014-off-season/

Carpenter, M. A. (2002). The implications of strategy and social context for the relationship between top management team heterogeneity and firm performance. *Strategic Management Journal, 23*(3), 275-284.

Chevalier, J., & Ellison, G. (1999). Are Some Mutual Fund Managers Better Than Others? Cross-Sectional Patterns in Behavior and Performance. *The Journal of Finance, LIV*(3), 875-899.

Crocker, A., & Eckardt, R. (2014). A Multilevel Investigation of Individual- and Unit-Level Human Capital Complimentarities. *Journal of Management, 40*(2), 509–530. doi:10.1177/0149206313511862

Depken II, C. A. (2000). Wage Disparity and Team Performance: Evidence from Major League Baseball. *Economics Letters, 67*, 87-92.

Falato, A., Li, D., & Milbourn, T. (2015). Which Skills Matter in the Market for CEOs? Evidence from Pay for CEO Credentials. *Management Science*.

Goff, B. (2013). Contributions of Managerial Levels: Comparing MLB and NFL. *Managerial and Decision Economics, 34*, 428-436.

Goldfarb, A., & Xiao, M. (2011). Who thinks about competition? Managerial ability and strategic entry in US local telephone markets. *American Economic Review, 101*, 3130-3161.

Goodall, A. H., & Pogrebna, G. (2015). Expert leaders in a fast-moving environment. *The Leadership Quarterly, 26*, 123-142. doi:10.1016/j.leaqua.2014.07.009

Goodall, A. H., Kahn, L. M., & Oswald, A. J. (2011). Why do leaders matter? A study of expert knowledge in a superstar setting. *Journal of Economic Behavior & Organization, 77*, 265-284.

Hakes, J. K., & Sauer, R. D. (2006). An Economic Evaluation of the Moneyball Hypothesis. *Journal of Economic Perspectives, 20*(6), 173-185.

Hambrick, D. C., & Mason, P. A. (1984). Upper Echelons: The Organization as a Reflection of Its Top Managers. *The Academy of Management Review, 9*(2), 193-206.

Hendricks, K. B., Hora, M., & Singhal, V. R. (2014). An Empirical Investigation on the Appointments of Supply Chain and Operations Management Executives. *Management Science, 61*(7), 1562 - 1583.

Jackson, K. C. (2013). Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. *The Review of Economics and Statistics, 95*(4), 1096-1116.

Kahn, L. M. (2000). The Sports Business as a Labor Market Laboratory. *Journal of Economic Perspectives, 14*(3), 75-94.

Kaplan, S. N., Klebanov, M. M., & Sorensen, M. (2012). Which CEO Characteristics and Abilities Matter? *The Journal of Finance, LXVII*(3), 973-1007.

Lazear, E. P., Shaw, K. L., & Stanton, C. T. (2015). The Value of Bosses. *Journal of Labor Economics*.

Lewis, M. (2004). *Moneyball: the art of winning an unfair game.* W. W. Norton & Company.

Mair, J. (2005). Exploring the Determinants of Unit Performance: The Role of Middle Managers in Stimulating Profit Growth. *Group & Organization Management, 30*, 263-288.

Malmendier, U., Tate, G., & Yan, J. (2011). Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies. *The Journal of Finance, LXVI*(5), 1687-1732.

Mollick, E. (2012). People and Process, Suits and Innovators: the Role of Individuals in Firm Performance. *Strategic Management Journal, 33*, 1001-1015. doi:10.1002/smj.1958

Owan, H., Takahashi, S., Tsuru, T., & Uehara, K. (2014). *Finding good managers: an econometric case study of a large Japanese auto dealership.* Working paper.

Peeters, T., & Szymanski, S. (2014). Financial Fair Play in European football. *Economic Policy, 29*(78), 343-390. doi:10.1111/1468-0327.12031

Pollard, R., & Pollard, G. (2005). Long-term trends in home advantage in professional team sports in North America and England (1876 – 2003). *Journal of Sport Sciences, 23*(5), 337-350.

Raes, A. M., Heijltjes, M. G., Glunk, U., & Roe, R. A. (2011). The Interface of the Top Management Team and Middle Managers: A Process Model. *Academy of Management Review, 36*(1), 102-126.

Siebert, S. W., & Zubanov, N. (2010). Management Economics in a Large Retail Company. *Management Science, 56*(8), 1398-1414.

Spanos, G. (2014). *Sorting Within and Across French Production Hierarchies.* Working paper.

Szymanski, S. (2003). The Economic Design of Sporting Contests. *Journal of Economic Literature, 41*(4), 1137-1187.

Terviö, M. (2008). The Difference That CEOs Make: An Assignment Model Approach. *American Economic Review, 98*(3), 642-668.

Wolfe, R., Wright, P. M., & Smart, D. L. (2006). Radical HRM Innovation and Competitive Advantage: the Moneyball Story. *Human Resource Management, 45*(1), 111-145.

Wooldridge, B., Schmid, T., & Floyd, S. W. (2008). The Middle Management Perspective on Strategy Process: Contributions, Synthesis, and Future Research. *Journal of Management, 34*(6), 1190-1221.

7 TABLES AND FIGURES

Table 1: Manager, team and spell identification in dataset

	Coach		GM		Team		Spell	
	#	% total	#	% total	#	% total	#	% total
Total	147		107		31		345	
Movers	56	38.1%	27	25.2%	-		-	
At firm with movers	139	94.6%	95	88.8%	28		-	
> 1 co-worker	144	98.0%	103	96.3%	-		-	
Identifiable	136	92.5%	91	85.0%	28	90.3%	291	84.3%
Reference category	31	21.1%	14	13.1%	-		81	23.5%
Effects in final	105	71.4%	77	72.0%	28	90.3%	210	60.9%

Table 2: Summary statistics – manager effects estimation sample

	Obs.	Mean	Std. Dev.	Min	Max
Output					
Game result	84534	0.5	0.5	0	1
Run difference	84534	0	4.366	-27	27
Inputs					
Total salary ('09 $)	84534	71.0m	35.1m	12.1m	228.0m
Std. dev. salary ('09 $)	84534	3.05m	1.46m	0.29m	8.82m
Stadium capacity	84534	47680	7569	33851	76273
Stadium age	84534	27.8	25.4	1	101
Instruments					
NL/AL	84534	0.53	0.50	0	1
Year	84534	2001.4	6.7	1989	2012
MSA population	84534	7.0m	5.6m	1.6m	23.4m
# Teams in MSA	84534	1.31	0.46	1	2
Tenure in MSA	84534	58.3	39.5	2	139
Fixed effects					
Team id	84534	14.4	8.0	1	28
GM id	84534	38.6	22.5	1	78
Coach id	84534	50.2	30.4	1	106
Spell id	84534	152.2	83.9	1	291

Note: Sources used for data set are http://www.shrpsports.com/mlb/ (game results), http://www.ballparks.com/ (stadium data), http://content.usatoday.com/sportsdata/baseball/mlb/salaries/team (payroll data) and http://www.baseball-reference.com/ (manager data).

Table 3: Summary statistics – manager personal characteristics

	Coach					GM				
Char.:	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Col. Att.	105	0.714	0.454	0	1	74	0.865	0.344	0	1
Col. Grad.	105	0.390	0.490	0	1	74	0.797	0.405	0	1
Col. top 100						74	0.189	0.394	0	1
Pro player	105	0.971	0.167	0	1	76	0.421	0.497	0	1
MLB player	105	0.781	0.416	0	1	76	0.197	0.401	0	1
MLB All Star	105	0.324	0.470	0	1	76	0.052	0.223	0	1
Age (days)	105	18538	2283	13018	25598	75	17556	3034	11649	26255
Exp. (days)	105	2141	2393	87	10124	76	3752	2178	728	10049
Minority	105	0.152	0.361	0	1	77	0.078	0.270	0	1

Table 4: Regression results – 1st stage

Game Result	**NO FE**	**TEAM FE**	**COACH FE**	**GM FE**	**ALL FE**	**SPELL FE**
Home advantage	0.073***	0.073***	0.073***	0.073***	0.073***	0.073***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Team payroll	0.093***	0.085***	0.059***	0.061***	0.054***	0.056***
	(0.008)	(0.009)	(0.011)	(0.012)	(0.013)	(0.015)
Std. dev. team payroll	-0.035***	-0.039***	-0.025**	-0.040***	-0.030**	-0.022*
	(0.008)	(0.009)	(0.010)	(0.010)	(0.012)	(0.012)
Stadium capacity	-0.014	-0.036*	-0.041	-0.083***	-0.091**	-0.088**
	(0.012)	(0.022)	(0.031)	(0.029)	(0.038)	(0.042)
Stadium age	-0.001	-0.000	0.003	-0.001	0.004	0.003
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
Constant	0.463***	0.463***	0.463***	0.463***	0.463***	0.464***
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Team FE	No	Yes	Yes	Yes	Yes	No
F-test	-	6.517***	2.323***	1.877***	0.949	-
Coach FE	No	No	Yes	No	Yes	No
F-test	-	-	3.145***	-	2.167***	-
GM FE	No	No	No	Yes	Yes	No
F-test	-	-	-	3.616***	2.292***	-
Spell FE	No	No	No	No	No	Yes
Observations	84,534	84,534	84,534	84,534	84,534	84,534
R-squared	0.011	0.015	0.021	0.022	0.026	0.029
Adj. R-sq.	0.011	0.014	0.019	0.019	0.022	0.023

***: significant at 1%-, ** significant at 5%-, * significant at 10%-level.

Table 5: Correlation coefficients between spell effects under baseline and alternative models

	IV Equivalent	Check 1	Check 2	Check 3	Check 4	Check 5
Win% Spell FE	0.9806	0.9720	0.9968	1.0000	0.9993	0.8290
Win% Spell IV FE	-	0.9824	0.9972	1.0000	0.9985	0.8055
Run Dif. Spell FE	0.9686	0.9578	0.9963	1.0000	0.9995	0.8601
Run Dif. Spell IV FE	-	0.9845	0.9968	1.0000	0.9988	0.8385

Note 1: Check 1 adds coach and GM age and experience; 2 excludes playoff games; 3 includes interaction between home indicator and stadium capacity/age; 4 excludes standard deviation of team payroll; 5 adds additional MSA level input measures.
Note 2: Values of 1.000 are larger than 0.9999, but do not equal one.
Note 3: All are significant at the 1% level.

Table 6: Impact of managerial layers and firm effects under different estimation methods

Output		Obs.	Fixed Effects				Random Effects							
							No Spell Effects				Spell Effects			
			Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.
Win %	Team	28	-0.12%	5.08%	-14.96%	9.64%	0.00%	0.86%	-2.43%	1.34%	0.00%	0.68%	-1.89%	1.00%
	Coach	105	1.69%	5.49%	-14.30%	17.38%	1.04%	2.04%	-4.21%	5.28%	0.98%	1.43%	-2.78%	4.09%
	GM	77	1.33%	5.84%	-7.99%	22.32%	0.67%	1.89%	-3.23%	5.90%	0.62%	1.39%	-2.23%	5.23%
	Match	210	-0.03%	2.02%	-10.05%	8.50%					0.03%	0.97%	-2.72%	2.95%
Run Dif.	Team	28	-0.101	0.681	-1.450	1.549	0.000	0.087	-0.236	0.135	0.000	0.053	-0.143	0.079
	Coach	105	0.126	0.526	-1.812	1.178	0.095	0.187	-0.361	0.515	0.085	0.105	-0.181	0.363
	GM	77	0.062	0.736	-1.777	1.977	0.046	0.223	-0.392	0.662	0.044	0.166	-0.295	0.565
	Match	210	0.003	0.178	-0.805	0.763					0.003	0.132	-0.296	0.432

Table 7: Correlation of individual effects across estimation methods

Correlation Individual Effects		**Win%**	**Run Dif.**	**Obs.**
FE - RE no spell	Coach	0.7518***	0.7517***	105
	GM	0.6592***	0.6225***	77
	Team	0.4061	0.2794	28
FE - RE spell	Coach	0.7020***	0.6804***	105
	GM	0.6146***	0.5760***	77
	Team	0.3083	0.1117	28
	Spell	0.4023***	0.4377***	210
RE spell - RE no spell	Coach	0.9816***	0.9611***	105
	GM	0.9821***	0.9702***	77
	Team	0.9703***	0.9299***	28

Table 8: Top 10 manager contributions to club winning percentage and run differential

Coach	Win%		Run Dif.		GM	Win%		Run Dif.	
	RE	FE	RE	FE		RE	FE	RE	FE
Lou Piniella	4.09%	8.25%	0.36	0.47	Brian Cashman	5.23%	14.01%	0.56	1.43
Mike Scioscia	3.97%	7.04%	0.29	0.74	Theo Epstein	3.71%	13.20%	0.39	1.36
Jack McKeon	3.76%	8.47%	0.28	0.65	Billy Beane	3.69%	22.32%	0.40	1.98
Joe Torre	3.71%	3.52%	0.29	0.44	Pat Gillick	3.48%	4.92%	0.39	0.57
Ron Washington	3.59%	17.38%	0.21	1.10	John Schuerholz	2.70%	10.98%	0.25	0.86
Fredi Gonzalez	3.40%	5.23%	0.22	0.48	Jon Daniels	2.38%	-3.98%	0.30	0.88
Willie Randolph	3.34%	6.34%	0.23	0.44	Brian Sabean	2.36%	7.48%	0.20	0.35
Davey Johnson	3.23%	6.78%	0.28	0.52	Ruben Amaro Jr.	2.26%	7.78%	0.21	0.70
Tony La Russa	3.22%	9.81%	0.25	0.80	Bob Watson	2.23%	7.09%	0.27	0.86
Buck Showalter	3.11%	10.62%	0.27	1.08	Roland Hemond	2.22%	8.01%	0.25	0.83

Table 9: Regression results – manager characteristics

	Game Result				**Run Differential**			
	Coach + GM		**Coach + GM**		**Coach + GM**		**Coach + GM**	
Characteristics:	**Coach**	**GM**	**Coach**	**GM**	**Coach**	**GM**	**Coach**	**GM**
Professional Player	0.026	0.018***	0.040*	0.018***	0.222	0.140***	0.340*	0.146**
	(0.016)	(0.005)	(0.023)	(0.007)	(0.143)	(0.043)	(0.202)	(0.061)
MLB Player	0.009	-0.005	0.023	-0.009	0.095	-0.055	0.215	-0.095
	(0.016)	(0.007)	(0.023)	(0.010)	(0.143)	(0.059)	(0.202)	(0.084)
MLB All-Star	0.033**	0.023**	0.047**	0.020	0.312**	0.214***	0.440**	0.188
	(0.016)	(0.009)	(0.023)	(0.013)	(0.144)	(0.082)	(0.202)	(0.116)
College not graduated	0.016***	0.018*	0.016**	0.022	0.153***	0.292***	0.150**	0.336***
	(0.005)	(0.010)	(0.008)	(0.014)	(0.046)	(0.089)	(0.066)	(0.126)
College graduated	0.012**	0.015**	0.012	0.022**	0.129***	0.163***	0.127*	0.222***
	(0.005)	(0.006)	(0.008)	(0.009)	(0.048)	(0.052)	(0.067)	(0.074)
College Top 100		0.020***		0.025**		0.156**		0.199**
		(0.008)		(0.011)		(0.065)		(0.093)
Minority	-0.004	0.014	-0.007	0.034***	-0.101**	0.155*	-0.128*	0.344***
	(0.006)	(0.009)	(0.008)	(0.013)	(0.051)	(0.079)	(0.072)	(0.109)
Age	0.263	-2.224**	0.193	-2.233*	-3.737	-23.93***	-5.055	-24.25**
	(1.356)	(0.889)	(1.924)	(1.259)	(11.628)	(7.874)	(16.503)	(11.146)
Age squared	-0.010	0.111**	-0.006	0.111*	0.221	1.195***	0.288	1.211**
	(0.069)	(0.046)	(0.098)	(0.065)	(0.591)	(0.405)	(0.839)	(0.573)
Experience	0.013*	-0.079**	0.008	-0.083*	0.144**	-0.557**	0.099	-0.601
	(0.007)	(0.031)	(0.010)	(0.045)	(0.061)	(0.269)	(0.086)	(0.382)
Experience squared	-0.001	0.006***	-0.000	0.006**	-0.009*	0.040**	-0.005	0.045*
	(0.001)	(0.002)	(0.001)	(0.003)	(0.005)	(0.018)	(0.007)	(0.026)
Inputs:								
Home advantage	0.079***		0.079***		0.261***		0.261***	
	(0.004)		(0.006)		(0.038)		(0.054)	
Total salary	0.071***		0.008		0.604***		0.001	
	(0.009)		(0.016)		(0.076)		(0.139)	
Std. dev. total salary	-0.029***		-0.001		-0.261***		-0.007	
	(0.008)		(0.017)		(0.072)		(0.147)	
Stadium capacity	-0.075***		-0.067**		-0.576***		-0.498*	
	(0.021)		(0.030)		(0.187)		(0.265)	
Stadium age	0.001		-0.003		-0.001		-0.038	
	(0.002)		(0.003)		(0.020)		(0.028)	
Constant	0.461***		0.461***		-0.130		-0.130	
	(0.056)		(0.086)		(0.373)		(0.537)	
Team FE	yes		yes		yes		yes	
IV wage	no		yes		no		yes	
Observations	104,660		104,660		104,660		104,660	
R-squared	0.018		0.017		0.015		0.014	

***: significant at 1%-, ** significant at 5%-, * significant at 10%-level.

Table 10: Regression results – assortative matching model

Dep. Var.	Coach Effect	Team Effect	Rel. Payroll	Obs. Char.	R-squared	Obs.
GM RE Win%	0.270*** (0.051)			No	0.073	210
GM RE Run Dif.	0.322*** (0.066)			No	0.104	210
GM FE Win%	-0.184*** (0.062)			No	0.034	210
GM FE Run Dif.	-0.096 (0.066)			No	0.009	210
GM RE Win%	0.211*** (0.055)	0.282*** (0.057)	0.498 (1.051)	No	0.152	210
GM RE Run Dif.	0.267*** (0.074)	0.296*** (0.056)	0.784 (1.059)	No	0.194	210
GM RE Win%	0.250*** (0.066)	0.186** (0.084)	0.187 (1.540)	Yes	0.259	204
GM RE Run Dif.	0.296*** (0.073)	0.228*** (0.072)	-0.080 (1.178)	Yes	0.310	204
Dep. Var.	**GM Effect**	**Team Effect**	**Rel. Payroll**	**Obs. Char.**	**R-squared**	**Obs.**
Coach RE Win%	0.227*** (0.081)	0.137* (0.073)	-0.626 (1.131)	No	0.090	210
Coach RE Run Dif.	0.294*** (0.071)	0.093 (0.072)	-0.955 (1.398)	No	0.113	210
Coach RE Win%	0.239*** (0.054)	0.259*** (0.091)	-1.144 (1.221)	Yes	0.289	204
Coach RE Run Dif.	0.282*** (0.053)	0.242*** (0.063)	-1.495 (1.355)	Yes	0.336	204

***: significant at 1%-, ** significant at 5%-, * significant at 10%-level.

Note: Observable characteristics account for GM/Coach age, experience, education and professional playing experience.

Table 11: Regression results – spell quality model

Spell quality	(I)		(II)	
	Win%	Run. Dif.	Win%	Run. Dif.
Coach Effect	0.375***	0.400***	0.368***	0.397***
	(0.065)	(0.077)	(0.072)	(0.074)
GM Effect	0.294***	0.251***	0.315***	0.249***
	(0.070)	(0.063)	(0.088)	(0.076)
Team Effect	0.075	0.088**	0.094	0.105*
	(0.052)	(0.045)	(0.069)	(0.062)
Player Payroll	2.034**	1.712**	1.957**	1.322
	(0.935)	(0.847)	(0.845)	(1.025)
Constant	-2.038**	-1.715**	-1.203	2.570
	(0.935)	(0.857)	(4.876)	(5.915)
Obs. Char.	No	No	Yes	Yes
Observations	210	210	204	204
R-squared	0.337	0.338	0.379	0.373

***: significant at 1%-, ** significant at 5%-, * significant at 10%-level.

Note: Observable characteristics account for GM/Coach age, experience, education and professional playing experience.

Table 12: Production function results – matching on observable characteristics

Match on:		**Win%**	**Run Dif.**	**Win%**	**Run Dif.**
Education	1 Cat.	0.007	0.100	-0.000	0.031
		(0.011)	(0.092)	(0.015)	(0.130)
	2 Cat.	-0.015**	-0.146**	-0.016*	-0.147*
		(0.007)	(0.057)	(0.009)	(0.080)
Playing level	1 Cat.	-0.008	-0.121**	-0.003	-0.071
		(0.007)	(0.058)	(0.010)	(0.082)
	2 Cat.	0.002	-0.051	0.011	0.032
		(0.010)	(0.084)	(0.014)	(0.118)
	3 Cat.	-0.004	-0.059	0.004	0.025
		(0.014)	(0.125)	(0.020)	(0.176)
Experience		-0.000	0.001	0.000	0.005
		(0.002)	(0.016)	(0.003)	(0.022)
Age		0.001	0.009	0.002	0.016
		(0.002)	(0.018)	(0.003)	(0.025)
Controls for:					
Inputs		Yes	Yes	Yes	Yes
Level Obs. Char.		Yes	Yes	Yes	Yes
Team FE		Yes	Yes	Yes	Yes
IV for wage		No	No	Yes	Yes
Observations		104,118	104,118	104,118	104,118
R-squared		0.019	0.015	0.017	0.014

** Significant at 5%-, * Significant at 10%-level.

Note: Categories in education match refer to no college/college attended/college graduated; categories in playing experience are no pro player/pro player/MLB player/All Star. Match is expressed as the number of ordered categories matched managers differ.

Table 13: Regression results – matching on observable characteristics

Spell quality	Education 1 cat	Education 2 cat	Playing level 1 cat	Playing level 2 cat	Playing level 3 cat	Age	Exp.	Level Char.	Man Team Eff.	R-sq.	Obs.
Win%	0.000	-0.005**						Yes	No	0.032	210
	(0.005)	(0.002)									
Run Dif.	-0.019	-0.059*						Yes	No	0.036	210
	(0.073)	(0.031)									
Win%	0.018	-0.258*						No	Yes	0.351	210
	(0.121)	(0.148)									
Run Dif.	0.046	-0.223						No	Yes	0.351	210
	(0.136)	(0.137)									
Win%			-0.001	-0.002	0.002			Yes	No	0.026	210
			(0.003)	(0.003)	(0.004)						
Run Dif.			-0.033	-0.033	0.001			Yes	No	0.028	210
			(0.047)	(0.040)	(0.049)						
Win%			-0.095	0.123	0.195			No	Yes	0.348	210
			(0.140)	(0.141)	(0.166)						
Run Dif.			-0.107	0.087	0.155			No	Yes	0.347	210
			(0.170)	(0.172)	(0.173)						
Win%						0.001		Yes	No	0.041	204
						(0.005)					
Run Dif.						0.034		Yes	No	0.035	204
						(0.073)					
Win%						0.057		No	Yes	0.342	204
						(0.061)					
Run Dif.						0.073		No	Yes	0.342	204
						(0.050)					
Win%							0.001	Yes	No	0.019	208
							(0.001)				
Run Dif.							0.013	Yes	No	0.018	208
							(0.010)				
Win%							-0.017	No	Yes	0.337	208
							(0.049)				
Run Dif.							-0.025	No	Yes	0.338	208
							(0.047)				

** Significant at 5%-, * Significant at 10%-level.

Note 1: Categories in education match refer to no college/college attended/college graduated; categories in playing experience are no pro player/pro player/MLB player/All Star. Match is expressed as the number of ordered categories matched managers differ.
Note 2: Level char is yes if the observed GM and coach characteristics are entered into the model on top of the match on the specified characteristic, no otherwise.

Table 14: Regression results – OLS and Cox-model for spell duration

Spell duration	(I)				(II)				(III)			
	Win%		Run. Dif.		Win%		Run. Dif.		Win%		Run. Dif.	
Spell Eff.	0.233***	0.668***	0.209***	0.672***	0.171**	0.765***	0.134*	0.767***	0.163**	0.750**	0.138**	0.742***
	(0.058)	(0.031)	(0.047)	(0.044)	(0.070)	(0.076)	(0.076)	(0.067)	(0.064)	(0.086)	(0.057)	(0.062)
Team Eff.					-0.019	1.035	-0.009	1.029	-0.007	1.074	-0.014	1.090
					(0.061)	(0.096)	(0.054)	(0.083)	(0.056)	(0.108)	(0.058)	(0.079)
Coach Eff.					0.133**	0.785***	0.136**	0.790**	0.121**	0.740***	0.125**	0.743***
					(0.057)	(0.066)	(0.064)	(0.074)	(0.058)	(0.071)	(0.050)	(0.071)
GM Eff.					0.023	0.847	0.028	0.849**	0.022	0.860	0.028	0.853
					(0.057)	(0.089)	(0.069)	(0.064)	(0.066)	(0.084)	(0.058)	(0.099)
Player Payroll					-0.088	2.899	-0.051	3.045	0.703	0.769	0.814	0.648
					(0.789)	(4.398)	(0.995)	(4.890)	(0.906)	(1.463)	(0.888)	(1.308)
Match Obs. Char.	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Est. Meth.	OLS	Cox	OLS	Cox	OLS	Cox	OLS	Cox	OLS	Cox	OLS	Cox
Obs.	210	210	210	210	210	210	210	210	204	204	204	204
(Pseudo) R-sq.	0.110	0.019	0.095	0.016	0.139	0.027	0.121	0.024	0.189	0.036	0.170	0.034

***: significant at 1%-, ** significant at 5%-, * significant at 10%-level.

Note: Observable characteristics account for the quality of the GM/Coach match with regard to age, experience, education and professional playing experience.

8 APPENDIX

This appendix provides additional results tables.

Table 1A: Regression results – 1st stage run difference model

Run Dif.	NO FE	TEAM FE	COACH FE	GM FE	ALL FE	SPELL FE
Home advantage	0.219***	0.219***	0.219***	0.219***	0.220***	0.216***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Team payroll	0.839***	0.719***	0.514***	0.481***	0.433***	0.420***
	(0.070)	(0.079)	(0.096)	(0.101)	(0.116)	(0.126)
Std. dev. team payroll	-0.279***	-0.321***	-0.243***	-0.315***	-0.210**	-0.146
	(0.073)	(0.077)	(0.089)	(0.087)	(0.099)	(0.106)
Stadium capacity	-0.047	-0.167	-0.328	-0.662**	-0.678**	-0.737**
	(0.102)	(0.190)	(0.273)	(0.260)	(0.337)	(0.375)
Stadium age	-0.012	-0.022	0.004	-0.015	0.028	0.041
	(0.014)	(0.021)	(0.027)	(0.027)	(0.034)	(0.038)
Constant	-0.110***	-0.110***	-0.110***	-0.109***	-0.110***	-0.108***
	(0.022)	(0.028)	(0.027)	(0.026)	(0.027)	(0.023)
Team FE	No	Yes	Yes	Yes	Yes	No
F-test	-	6.543***	2.335***	2.577***	1.826***	-
Coach FE	No	No	Yes	No	Yes	No
F-test	-	-	3.633***	-	2.389***	-
GM FE	No	No	No	Yes	Yes	No
F-test	-	-	-	4.609***	2.899***	-
Spell FE	No	No	No	No	No	Yes
Observations	84,534	84,534	84,534	84,534	84,534	84,534
R-squared	0.007	0.011	0.019	0.019	0.025	0.028
Adj. R-sq.	0.007	0.010	0.017	0.016	0.020	0.021

***: significant at 1%-, ** significant at 5%-, * significant at 10%-level.