Stopping on Nine: Evidence of Heuristic Managerial Decision-Making in Major League Baseball Pitcher Substitutions[1]

David C. Phillips
Hope College

Abstract
Behavioral biases may affect managers' decisions regarding employee workloads with subsequent implications for productivity. Using data on managers' pitcher substitution decisions in Major League Baseball, I find evidence that both salience of round numbers and overweighting present (certain) versus future (uncertain) outcomes affect managers' decisions. In recent years, managers remove starting pitchers more often when the next pitch will result in a pitch count ending in zero. Unlike counts ending in 9, pitch counts ending in 8 do not exhibit higher substitution rates despite the fact that pitchers usually throw multiple pitches before the manager's next decision opportunity. These facts imply that managers make decisions subject to both the salience of round numbers and severe overweighting of immediate (certain) outcomes. However, the observed biases in decision-making disappear when the stakes are high. I find no evidence of disproportionate substitutions on the 9's in tie games. As a result, the final digit of the pitch count does not discontinuously affect the probability that the team ultimately wins the game. Managers act in a manner consistent with rational inattention, displaying severe decision-making biases but only when the stakes are low.

[1] E-mail: phillipsd@hope.edu. I have benefitted from conversations with Peter Boumgarden and assistance with data from Josh Coulter. All remaining errors and omissions remain the responsibility of the author.

1. Introduction

The existence of systematic biases in decision-making have been well-documented in the lab (Kahneman and Tversky, 1979). A large body of evidence also demonstrates sub-optimal heuristic decision-making by buyers (Lacetera, et. al. 2012; Chetty, et. al. 2009), sellers (List, 2003; List, 2004), and workers (Fehr and Goette, 2007) in real-life markets. A considerable portion of this literature focuses on whether such biases disappear either as the decision-maker gains experience in the environment (List, 2003; List, 2004) or when the stakes are high (Pope and Schweitzer, 2011). However, the existing literature has rarely examined the existence and persistence of such biases in decision-making by managers. Managers tend to be experienced in their field and face high-stakes decisions regarding employees' workload that can affect rates of burnout, turnover, and workplace injury (Nahrang, et. al. 2011). However, data limitations generally prevent detailed analysis of managerial decisions.

In the present study, I test whether Major League Baseball managers exhibit heuristic biases when making decisions regarding player substitutions. Managers must weigh performance, fatigue, and injury risk when faced with such decisions (Bradbury and Forman, 2012; Tango et. al., 2007). I find that experienced decision-makers respond to these tradeoffs in a manner indicating the salience of round numbers. Over the past decade, managers have removed starting pitchers 2 percentage points (6% of the mean) more often when the total number of pitches thrown will immediately cross a round-number threshold if the pitcher is left in the game. Despite the fact that pitchers usually throw more than one pitch between substitution opportunities, the observed change in strategy is discontinuous. An increase in substitution occurs when the pitch count ends in 9 but not when it ends in 8. This latter fact indicates that, in addition to responding to round numbers, managers heavily discount uncertain events in the near future relative to immediate certain events. Theoretically, such effects could be driven either by a non-expected utility preference for certainty (Kahneman and Tversky, 1979; Camerer and Kunreuther, 1989) or severe discounting of the future (Strotz, 1955); in the present context the two bear considerable similarities (Halevy, 2008). I develop a simple theoretical model of optimal substitution strategy to investigate the extent of these biases implied by the data. For the model to be consistent with the empirical results, managers' must employ both round number salience and severe present bias, discounting the future around 90% *per pitch*. Thus, I find that experienced managers can exhibit severe behavioral biases similar to those documented in buyers, sellers, and workers.

I extend the main results in two directions. First, I demonstrate that managers' behavioral biases disappear in high-stakes situations. When the score is tied, managers' decisions have a greater likelihood of affecting the final outcome of the game, and I find no evidence of disproportionate substitutions on round numbers in tie games. Consistent with this result, I find no effect of round-number substitutions on the probability of winning games. This result contrasts with, for instance, evidence that golfers retain biases in high-stakes situation (Pope and Schweitzer, 2011). Baseball managers exhibit very large decision-making biases but appear to substitute more consistent decision-making in high-stakes situations. The present results are thus consistent with rational inattention by managers rather than sub-optimal decision-making per se (Caplin and Dean, 2015). Second, disproportionate substitution on pitch counts ending in 9 is new, appearing only in the 2000s. Improved data and a focus on pitcher injuries and pitch count data in recent years has generated rather than alleviated these biases. This result suggests that greater focus on data-driven decision-making can actually encourage the use of heuristics.

In addition to the studies already discussed, the present paper relates to two major strands of the existing literature. First, it contributes to the growing literature on the importance of round numbers to decision-makers in real-life situations, so-called left-digit bias. In addition to the classic importance of round numbers in prices (i.e. "99 cents") the importance of round numbers has been observed for used car odometers (Lacetera, et. al. 2012) as well as baseball batters and SAT takers (Pope and Simonsohn, 2011). My results confirm the importance of round numbers for a new group of decision-makers: managers. The present study also relates to the large literature using data from professional sports to test rationality in decision-making. Among other examples, football coaches' 4th down decisions (Romer, 2006), marathon runners' effort levels (Allen, et. al. 2014), and soccer players' penalty kicks (Chiappori, et. al. 2002) exhibit sub-optimal decision-making and/or non-neoclassical preferences. The present study furthers the argument that detailed data from professional sports can be used to test important questions regarding optimal decision-making in high-stakes situations.

2. Theory and Discussion

2.1. Removing Pitchers as a Discrete-Time Stopping Problem

While the focus of this paper is empirical, theory helps demonstrate what behavioral biases are consistent with observed decisions by managers. In particular, I will demonstrate that the observed results require a model of decision-making including very high time discounting in addition to round number salience. Consider a manager who at the end of each plate appearance must decide whether to leave a starting pitcher with pitch count p in the game or replace him with a reliever. From the point of theory, the risk-neutral manager faces a discrete time stopping problem similar to the standard labor market search model (McCall, 1970) but with a finite horizon and non-stationarity. Suppose that the reliever is an option with unknown future value but certain current value (e.g. because talent depends on environmental factors such as weather). After each plate appearance, reliever productivity y is drawn randomly and independently from a normal distribution with mean μ and standard deviation σ. The manager observes this value after it has been drawn. In Major League Baseball, once a pitcher is removed from the game he may not return. I will simplify this fact and assume that after switching the manager must stay with the reliever at productivity y for the remainder of the game. Denote as R and S the managers' options to choose the reliever or the starter, respectively. Then, a manager will have a value of choosing the reliever as:

$$V_p^R(y) = y + \beta V_{p+1}^R(y)$$

The parameter β represents the per-pitch discount rate of the manager, though it can also be interpreted as the extra discount which the manager applies to uncertain versus certain outcomes. For simplicity, I will assume that the game stops predictably after T pitches. Thus, $V_T^R = 0$ and geometric series results can be used to obtain:

$$V_p^R(y_R) = \left(\frac{1 - \beta^{T-p}}{1 - \beta}\right) y \qquad (1)$$

The manager compares this value of choosing the reliever to staying with the starter. The starter has flow value $g(p)$ that due to fatigue and injury risk depends on the pitch count. By remaining with the starter, the manager also retains the option value of choosing either the starter or reliever in the future. Thus, if the each plate appearance included only one pitch the value of the starter would be stated as:

$$V_p^S = g(p+1) + \beta E\left[\max\left\{V_{P+1}^R\left(y_{p+1}\right), V_{p+1}^S\right\}\right]$$

However, managers do not generally replace pitchers in the middle of plate appearances, and the number of pitches in a plate appearance is itself unknown ex-ante. In general, a plate appearance will last i pitches with probability λ_i, and the chosen pitcher will throw pitches $p+1$ through $p+i$. Thus, the true value function for staying with the starter is slightly more complicated, adding up the relevant value functions for different possible plate appearance lengths:

$$V_p^S = \sum_{i=1}^{k} \lambda_i \left\{ \left[\sum_{j=1}^{i} \beta^{j-1} g(p+j) \right] + \beta^i E\left[\max\{V_{P+i}^R(y_{p+i}), V_{p+i}^S\}\right] \right\} \quad (2)$$

The manager must then decide whether he prefers (2) or (3), the reliever or the starter. As is standard in the literature, this strategy will be defined by a cutoff rule. Define $\bar{y}_p$ as the reservation productivity for the reliever such that the starter remains if $y < \bar{y}_p$ and the manager substitutes in the reliever if $y \geq \bar{y}_p$. This cutoff is defined by an indifference condition:

$$V_p^R(\bar{y}_p) = V_p^S = \left(\frac{1-\beta^{T-p}}{1-\beta}\right)\bar{y}_p \quad (3)$$

Equations (2) and (3) can be combined with the assumption of normality of y to demonstrate:

$$\bar{y}_p = \left(\frac{1-\beta}{1-\beta^{T-p}}\right) \sum_{i=1}^{k} \lambda_i \left\{ \left[\sum_{j=1}^{i} \beta^{j-1} g(p+j)\right] + \beta^i \left(\frac{1-\beta^{T-p-i}}{1-\beta}\right) \left[\mu + \sigma \frac{\phi\left(\frac{\bar{y}_{p+i}-\mu}{\sigma}\right)}{1-\Phi\left(\frac{\bar{y}_{p+i}-\mu}{\sigma}\right)}\right] \right\} \quad (4)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution. See Appendix for the mathematical details. Equation (4) defines the manager's strategy recursively with $\bar{y}_p$ depending on $\bar{y}_{p+1}, \bar{y}_{p+2}, \ldots, \bar{y}_{p+k}$. Given the finite time horizon, $\bar{y}_p = 0$ for $p > T$. This provides the initial condition needed to define the sequence of reservation values back to all possible pitch counts. It is then straightforward to calculate the probability the manager leaves the starting pitcher in the game as $\Phi\left(\frac{\bar{y}_p-\mu}{\sigma}\right)$.

2.2. Neoclassical Managerial Decisions

The heuristic decision-making of the manager can be embodied in the starting pitcher's as yet undefined flow value, $g(p)$. I will consider two ways in which this value may change throughout the game: fatigue and injury risk.

$$g(p) = y_S(p) - c(p)$$

The first term $y_S(p)$ denotes the starting pitcher's productivity which declines smoothly with the pitch count. For concreteness suppose that the pitcher's productivity takes a quadratic form

$$y_S = y_0 - Ap^2$$

A rational manager will also likely consider the pitcher's injury risk to accumulate smoothly over the course of the game. If we define the manager's decision to leave the pitcher in the game after i pitches as $S(i) = I\{y_p < \bar{y}_p\}$ then injury costs for an entire game in which a pitcher throws $p - 1$ pitches can be characterized as $C\left(\sum_{i=1}^{p-1} S(i)\right) = c_1 + c_0(p-1)$. For concreteness I suppose that $C(\cdot)$ is linear. Then, deciding to leave the starter in for one more pitcher generates a constant cost of c_0. Thus, a rational manager viewing fatigue and injury risks as smoothly varying considers a flow value:

$$g(p) = y_0 - Ap^2 - c_o$$

The left pane of Figure 1.a. simulates an example using the optimal decision rule from equation (4) for a manager who views fatigue and injury risk in this manner. The probability that the starter remains in the game declines smoothly with the pitch count as the pitcher fatigues.

2.3. Heuristic Managerial Decisions

I wish to examine how this decision rule changes when the manager uses a decision heuristic based on round values of the pitch count. Suppose as before that the manager considers fatigue using a smooth, quadratic function of the pitch count. However, suppose that he only considers injury risk when it is observationally salient due to an upcoming round-numbered pitch count. This can be captured by a piecewise linear, discontinuous cost function:

$$C\left(\sum_{i=1}^{p} S(i)\right) = c_1 + c_0 * 10 * \left\lfloor \frac{p}{10} \right\rfloor$$

where $\lfloor \cdot \rfloor$ is the floor function. In this case the marginal injury cost of another pitch with vary discontinuously, taking on a value of $10 * c_0$ for round numbers and 0 otherwise. Thus, in this case:

$$g(p) = \begin{cases} y_0 - Ap^2 - 10c_0 & if \quad p - 10 * \left\lfloor \frac{p}{10} \right\rfloor = 0 \\ y_0 - Ap^2 & if \quad p - 10 * \left\lfloor \frac{p}{10} \right\rfloor \neq 0 \end{cases}$$

The right pane of Figure 1.a. displays how the manager's decision changes when using this heuristic. As expected, the decision rule becomes discontinuous, jumping at round values. Perhaps less obviously, the decision rule only jumps continuously from one direction. The probability that the manager removes the starter declines continuously and accelerates as the pitch count approaches a round number. The forward-looking manager using a round number heuristic rationally anticipates that a pitcher with a pitch count of 118 will likely cross the 120 pitch threshold if left in the game. Thus, the probability of leaving the starter in falls rapidly but continuously before a round number, only jumping once the threshold has been reached.

For comparison, consider a much less patient manager. Figure 1.b. shows the same results for a present-focused manager with a discount rate of 0.1 instead of the more patient manager in Figure 1.a. who has a discount rate of 0.9. A per-pitch discount rate of 0.1 is extremely present-focused. In this instance, the probability that a pitcher remains in the game jumps discontinuously twice. As before, the probability jumps up after the pitcher crosses the round number threshold and incurs the heuristic injury cost. However, the probability of remaining in the game also jumps down on pitch counts ending in 9 relative to behavior on prior pitches. In this case, the manager either does not place significant value on events beyond the next pitch. Thus, he responds suddenly when a round number is immediately on the horizon but no sooner. Note that β is a per pitch discount rate, such that even a rate of 0.9 indicates strong time preferences. To obtain the pattern shown in the right pane of Figure 1.b. the manager must be extremely present focused (or overweight certain compared to uncertain consequences). Figure 1.c. underscores this fact by showing a simulation in which the manager has a discount rate of 0.3. Even with this still extremely low discount rate, the manager foresees crossing the 10-pitch threshold enough that the probability the starter drops noticeably at pitch counts ending in 8 in addition to those ending in 9.

Altogether, the theory provides clear empirical predictions. Heuristic thinking by managers regarding injury risk can generate discontinuities in the probability that a starting pitcher remains in the game. When the manager is considering leaving the pitcher in for a pitch that crosses a round number threshold, the current pitch count will end in 9 and thus jumps in managerial decisions will be expected at this point. However, for the probability that the starter remains in the game to "dip" at round numbers, managers must employ severe time-discounting (or a very low weight on uncertain relative to certain outcomes) in addition to round-number heuristics.

3. Data and Context

I use data from Retrosheet,[2] a database of all events in all games in Major League Baseball. Event records include the sequence of pitches thrown (e.g. swinging strike, ball, pickoff throw), the outcome of each plate appearance (e.g. single, home run, walk), and the

---

[2] The information used here was obtained free of charge from and is copyrighted by Retrosheet. Interested parties may contact Retrosheet at "www.retrosheet.org".

current score. Using pitch sequence data, I calculate the number of pitches for each plate appearance using the number of events recorded by Retrosheet in the pitch sequence, excluding non-pitches (plays not involving the batter, catcher blocks, pickoff throws, indications of the runner moving on the pitch, and indications of no pitch thrown e.g. due to balks). The pitch count then represents the running total of this value for a particular pitcher over the course of the game. I compute pitch counts for all pitchers of all Major League Baseball games for the years 1992-2012. I exclude a small number of games for which pitch sequence data is missing (< 2%). I limit the sample to starting pitchers and plate appearances during which the pitch count is at least 100 because the probability of being removed from the game is low below this point.

I measure the manager's decision and final outcomes directly. The data list the pitcher for each plate appearance, and I measure a dummy variable for whether the pitcher in the next plate appearance versus the same team is the same as the current pitcher. The data also provide multiple measures of the success of the team, in particular the team's pitchers, following the current plate appearance. I measure whether the team ultimately wins the game as the final goal of the team. Scoring runs is the intermediate outcome to this end, and I can measure how many runs the team's pitchers allow following a particular plate appearance. Finally, as the most immediate outcome I can measure success by the pitcher's team in the next plate appearance via dummy variables for whether the next batter succeeds in reaching base (by hit, walk, or otherwise) and how many bases the batter attains with a hit.

Table 1 shows summary statistics from the data. The first pane displays the criteria for sample selection. The first column shows that the data include nearly 4 million plays. The second column demonstrates that I select 175,965 plate appearances for which the original pitcher is in the game with a pitch count of at least 100. Thus, by definition my sample includes only starting pitchers and these players have already thrown an average of 109 pitches. Managers are much less likely to allow the pitchers in my sample to remain in the game. On average, a current pitcher has a 91% chance of remaining in the game. In my sample, pitchers remain only 73% of the time because I am selecting a group of plays with starting pitchers who have already thrown many pitches. The third and fourth columns split the sample into pitchers with pitch counts ending in nine or a number other than nine. A preliminary comparison shows that pitchers with pitch counts ending in 9 remain in the game 67% of the time, which is less than the 73% rate of pitchers with counts ending in a digit other than 9. However, this does not control for the fact that pitchers with pitch counts ending in nine have by construction thrown more pitches on average (113 vs. 109). Thus, this estimate should not be interpreted as evidence of bias, but it does provide a first indication of the eventual results.

The second pane of Table 1 shows the different final outcomes that I will measure. Teams in the study sample eventually win 57% of their games because teams tend to only allow pitchers to accumulate pitch counts over 100 if the team is doing well. But there is no observable difference when the pitch counts end in 9. The opposing team does appear to perform somewhat worse when the current pitcher's count ends in 9. Opposing teams score 1.31 rather than 1.42 runs in the remainder of the game. They reach base 33.2% of the time instead of 33.8% of the time (a 6 point drop in on-base percentage). They average .394 bases from hits rather than .404 bases (essentially, a drop of 10 points in slugging average). However, measuring true effects on performance will require seeing if these outcomes deviate from what would be predicted by a smooth function of the pitch count.

4. Empirical Strategy

I will measure whether managers' decisions to remove a pitcher incorporate a heuristic based on the pitch count rather than choosing optimally based on all observable information. In particular, I will test whether managers disproportionately remove pitchers with pitch counts for which the final digit is 9. As discussed in the theory section below, such behavior would relate to at least two behavioral biases. A discontinuous difference in behavior for 109 versus 110 pitches likely indicates so-called left digit bias in which decision-makers focus on leftmost digits, in this case implicitly believing that a pitcher's health or performance will be exceptionally harmed by crossing round numbers. Given that nearly all plate appearances in Major League Baseball require more than one pitch from the pitcher (87% in the present sample), stopping disproportionately on numbers ending in 9 but not those ending in 8 also demonstrates severe focus on the present. Even if a manager wishes to use a round number pitch count rule, a pitch count ending in 8 will most likely cross the same injury risk threshold as one ending in 9. Only if managers use heuristic rules and maintain a very high *per-pitch* discount rate will pitch counts ending in 9 differ discontinuously from both counts ending in 0 and counts ending in 8. Thus, disproportionate removal of pitchers on pitch counts ending in 9 demonstrates that managers exhibit both left digit bias and severe focus on immediate consequences of their actions.

I will use polynomial regressions to measure whether managers disproportionately remove pitchers at the end of plate appearance when the pitch count ends in 9. Consider the following regression:

$$Y_{itg} = \alpha + \beta N(P_{itg}) + \rho\left(P_{itg}\right) + \epsilon_{itg} \quad (5)$$

Define p_{itg} as the pitch count for pitcher i after plate appearance t of game g. Y_{itg} is an outcome of interest, for instance a dummy variable for whether the pitcher remains in the game for the next plate appearance. The treatment variable $N(p_{itg}) = I\left\{p_{itg} - 10 * \left\lfloor\frac{p_{itg}}{10}\right\rfloor = 9\right\}$ is a dummy for whether the pitch count ends in nine. Finally, the function $\rho(p_{itg})$ controls for a polynomial of the pitch count. I will focus on β, which measures whether the managers' treatment of situations in which the pitch count ends in 9 deviates from their standard practice as captured by a smooth function of the pitch count. If β is non-zero, then this would indicate that managers treat pitch counts ending in nine discontinuously differently from other similar situations.

Local random variation in pitch counts can justify the use of equation (5). Pitch counts are not randomly assigned in general. A pitcher may have a low pitch count because that pitcher has been successful in getting outs with few pitches or simply because it is early in the game. However, within a narrow range of pitch counts, the final digit of the pitch count should be as good as randomly assigned, depending on random variation in the outcome of the previous at bat. Consider the final pane of Table 1, which displays outcomes in the game leading up to the plate appearance in question. In the analysis sample, up to the plate appearance in question a pitcher with a pitch count ending in 9 has given up on average 2.21 runs. This is identical to the runs given up by pitchers with counts ending in digits other than nine. Conditional on a smooth function of the pitch count, runs scored by the pitcher's own team and performance of the most recent opposing team's batter also appear similar. Controlling for a polynomial in the pitch count, pitchers with pitch counts ending in 9 and pitchers with pitch counts ending in other digits have performed similarly and face similar game situations. Thus, we can be confident that any differences in manager behavior regarding these pitchers reflects response to the pitch count itself rather than differences in the quality of the pitcher or the game situation.

5. Results

5.1. Managers' Decisions to Remove Pitchers

Managers disproportionately remove starting pitchers from games when the final digit of their pitch count equals 9. Figure 2 displays the main evidence for this claim. The probability that the pitcher remains in the game declines with the final digit of the pitch count as fatigue and injury risk increase. A quartic polynomial fitted curve displays the downward trend. The sample probability that a pitcher remains in the game, given the final digit of the pitch, generally matches the fitted line closely with one exception. A large gap between the 0.67 probability that a pitcher remains in the game with a pitch count ending in 9 and the predicted rate of 0.69 provides evidence of heuristic decision-making by managers. The first two columns of Table 2 quantify this gap. Using a regression as in equation (5) with linear control for pitch count, I find that managers are 1.6 percentage points less likely to leave in a pitcher with a pitch count ending in 9. Controlling for a quartic polynomial of pitch counts diminishes this gap to 1.1 percentage points, which remains statistically significant at the 5% level. This discontinuous jump in managerial decisions when pitch counts end in 9 provides strong evidence that managers use a heuristic, attempting to avoid having pitchers cross 10-pitch thresholds in the pitch count. This tendency is so strong that a manager is actually more likely to leave a pitcher in the game when the pitch count ends in 0 than when it ends in 9. That is, a manager will stay with a more fatigued pitcher over a less fatigued pitcher despite worse performance and greater injury risk.

5.2. Differences across Time

To confirm that heuristic managerial decision-making drives the observed jump in the probability that a pitcher remains in the game, I test whether such effects have become stronger as teams have paid close attention to pitch counts. The popular press has widely documented that teams have placed greater emphasis on pitch counts as time has progressed (e.g. Posnanski, 2009). Such changes are evident in the data. Both low and high pitch count games have become less common. In 1992, 90 percent of starting pitchers ultimately threw at least 125 pitches and only 10 percent threw less than 69 pitches. In 2012, these percentiles narrowed to 114 and 78, respectively. The reduced dispersion in pitch counts matches popular accounts that teams have generally become more likely to use pitch counts to make decisions. At the top end of the distribution, this change accelerates in the year 2001. As shown in Figure 3, the number of pitchers throwing at least 123 pitches per game (1 s.d. above the mean in 1992) has dropped over time but most noticeably dropped in the middle of my sample in 2001.

Thus, to confirm whether heuristic managerial decision-making drives my results, I test whether the gap observed for pitch counts ending in 9 widens over time. Figure 4 displays these results. The higher (green and orange) collection of plotted points shows the probability of a starting pitcher remaining in the game by pitch count for the 1992-2000 baseball seasons. The solid orange points representing pitch counts ending in 9 do not noticeably deviate from the general trend. On the other hand, the lower (blue and red) collection of points covering the 2001-2012 seasons do exhibit downward jumps for pitch counts ending in 9. Column (3) of Table 2 tests this difference formally by modifying equation (5) to interact the ending-in-nine dummy and the pitch count polynomial with a dummy for the 2001-2012 timeframe. The coefficient of -0.000 on the ending-in-nine dummy indicates that prior to 2001 managers only stay with pitchers with pitch counts ending in nine < 0.1 percentage points less often, a statistically insignificant difference. No bias exists during the early period during which pitch counts were less emphasized. On the other hand, the gap between 9 and other ending digits is 2.3 percentage points during later years. As expected, measured heuristic decision-making by managers has grown as Major League Baseball has paid closer attention to pitch counts.

5.3. Close Games

Having identified heuristic rules in managers' decisions regarding pitcher substitution, I now test whether such rules perform optimally. If pitcher fatigue, performance, and injury risk all change smoothly with the pitch count, then optimal pitcher substitution probabilities would also vary smoothly. Heuristics would be sub-optimal. However, the magnitude of the penalty for using a heuristic will depend on when it is used. If managers aim to win baseball games then it matters little if they use heuristics when one team is far ahead and the probability of changing the outcome of the game is low. Figure 5 tests whether managers act "rationally inattentive" in this way, using heuristics only in situations of limited importance. Both panes in Figure 5 consider the period 2001-2012 during which heuristic decision-making appears. The first pane of Figure 5 shows managers' decisions when the score is tied and thus events can greatly affect the probability of winning. In tied games, managers' decisions do not jump discontinuously at pitch counts ending in nine. As shown in the second pane, all such heuristic decision-making is concentrated in situations in which the score is not tied. The final column of Table 2 formalizes this idea. Pitch counts ending in 9 result in a 2.7 percentage point decrease in the probability that starting pitcher remains in the game when the game is not tied, but that gap almost completely closes for tie games. Managers use heuristic decision-making, but they only do so when closer attention to detail is unlikely to change the ultimate outcome.

5.4. Team Performance

A more direct test of optimality in manager decision making can be conducted by examining direct measures of pitching success: limiting the number of runs scored by the other team and winning the game. Table 3 displays results of a similar empirical strategy applied to these outcomes. The first two columns of Table 3 shows results for runs scored by the other team from the current point in the game to the end of the game. Column (1) shows no difference in opponent run-scoring following pitch counts ending in 9 versus other digits. Focusing on the 2001-2012 period similarly shows no statistically significant difference in runs scored by the other team after a pitch count ending in 9. However, the measured effects are rather large. If the opposing team's score increases by 0.03 for a 0.022 decrease in the probability that the pitcher remains in the game and effects are linear, then replacing a starting pitcher early would increase run scoring by 1.36 runs. Similarly, I measure statistically insignificant but noisily measured effects for winning the game in columns (3) and (4). Altogether, I find no evidence that heuristic decision-making by managers affects overall team performance, though the measured effects are noisy.

6. Conclusion

This study demonstrates the existence of experienced managers who use heuristics in determining their employees' workloads. Other things equal, Major League Baseball managers disproportionately remove pitchers from games when the pitch count ends in 9 but not when they end in 8. I develop a simple theoretical model of optimal substitution strategy. To match the empirical results, managers must maintain not only a round-number heuristic but also an extreme discount rate applied to future/uncertain events relative to immediate/certain events. Despite exhibiting such large behavioral biases, managers' sub-optimal choices have no measurable effect on the probability of actually winning games. Consistent with this fact, I find that managers do not display heuristic decision-making in tie games, leaving sub-optimal decisions for low-stakes situations. Altogether, the results establish a context in which experienced decision-makers systematically make errors but only in low-stakes situations.

This study supports the idea that the same people may display optimal rational decision-making in some contexts but not in others (e.g. Kahneman, 2011). While the previous literature demonstrates that biases can disappear with experience (List, 2003), I demonstrate that the stakes may influence use of heuristics. In this sense, decision-makers may be rationally-inattentive, only exerting the effort required of consistent rationality when the returns to such decisions are high. Other evidence does exist for sub-optimal decision-making even in the face of high stakes and experienced decision-makers (Pope and Schweitzer, 2011). The present study complicates existing results, indicating that the tendency of behavioral biases to disappear when the stakes are high may depend on the context.

References

Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2014). "Reference-dependent preferences: Evidence from marathon runners." NBER Working Paper No. 20343.

Bradbury, J. C., & Forman, S. L. (2012). "The impact of pitch counts and days of rest on performance among major-league baseball pitchers." *The Journal of Strength & Conditioning Research*, *26*(5), 1181-1187.

Camerer, C. F., & Kunreuther, H. (1989). "Decision processes for low probability events: Policy implications." *Journal of Policy Analysis and Management*, *8*(4), 565-592.

Caplin, Andrew, and Mark Dean (2015) "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review*, 105(7): 2183-2203.

Chiappori, P. A., Levitt, S., & Groseclose, T. (2002). "Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer." .*American Economic Review*, 1138-1151.

Chetty, R., Looney, A., & Kroft, K. (2009). "Salience and Taxation: Theory and Evidence." *The American Economic Review*, *99*(4), 1145.

Fehr, E. & Goette, L. (2007). "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *The American Economic Review*, *97*(1), 298-317.

Halevy, Y. (2008). "Strotz meets Allais: Diminishing impatience and the certainty effect." *The American Economic Review*, 1145-1162.

Kahneman, D. (2011) *Thinking, Fast and Slow.* New York: Farrar, Straus, and Giroux.

Kahneman, D., & Tversky, A. (1979). "Prospect theory: An analysis of decision under risk." *Econometrica: Journal of the Econometric Society*, 263-291.

Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic Thinking and Limited Attention in the Car Market. *American Economic Review*, *102*(5), 2206-2236.

List, John A. (2003) "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, 118(1): 41–71.

List, J. A. (2004). "Neoclassical theory versus prospect theory: Evidence from the marketplace." *Econometrica*, *72*(2), 615-625.

McCall, J. J. (1970). "Economics of information and job search." *The Quarterly Journal of Economics*, 113-126.

Nahrgang, J. D.; Morgeson, F. P.; Hofmann, D. A. (2011) “Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes.” *Journal of Applied Psychology*, Vol 96(1), 71-94.

Pope, D., & Simonsohn, U. (2011). “Round numbers as goals evidence from baseball, SAT takers, and the lab.” *Psychological science*, *22*(1), 71-79.

Pope, D. G., & Schweitzer, M. E. (2011). “Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes.” *The American Economic Review*, *101*(1), 129-157.

Posnanski, J. (2009) “Talkin’ Pitch Counts and Nolan Ryan’s Crusade, with Bill James.” *Sports Illustrated*, 6/15/2009.

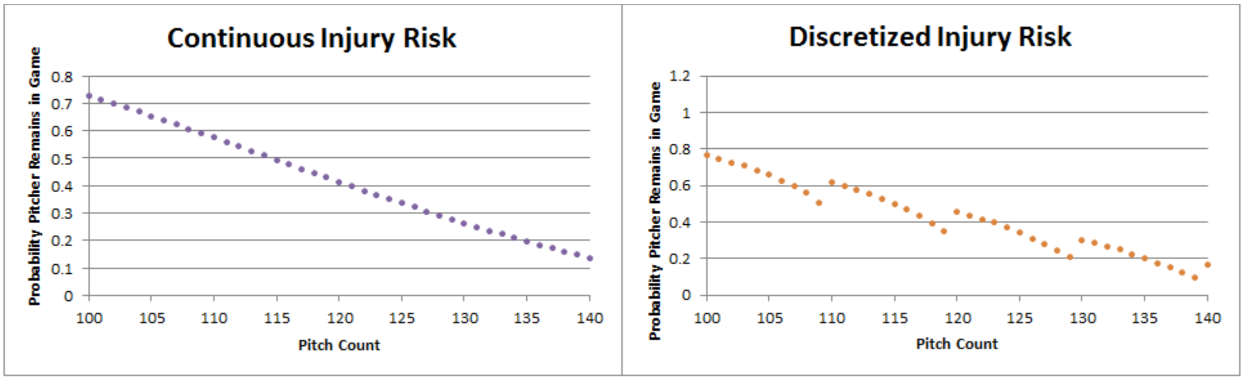
Romer, D. (2006). “Do firms maximize? Evidence from professional football.” *Journal of Political Economy*, *114*(2), 340-365.

Strotz, R. H. (1955). “Myopia and inconsistency in dynamic utility maximization.” *The Review of Economic Studies*, 165-180.

Tango, T.M., M.G. Lichtman, and A.E. Dolphin (2007) “The Book: Playing the Percentages in Baseball.” *Potomac Books Inc.*

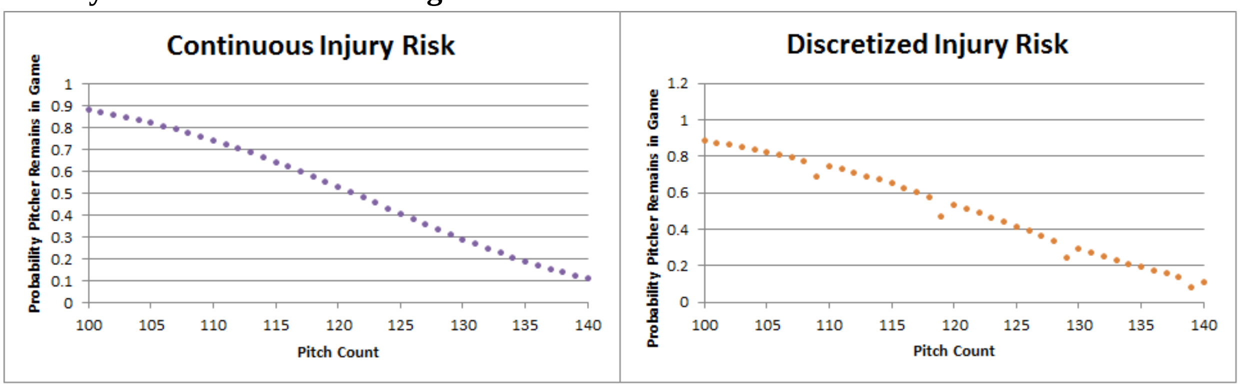
Figure 1. Theory-Based Simulation of Heuristic Managerial Decision-Making

a. Patient Manager



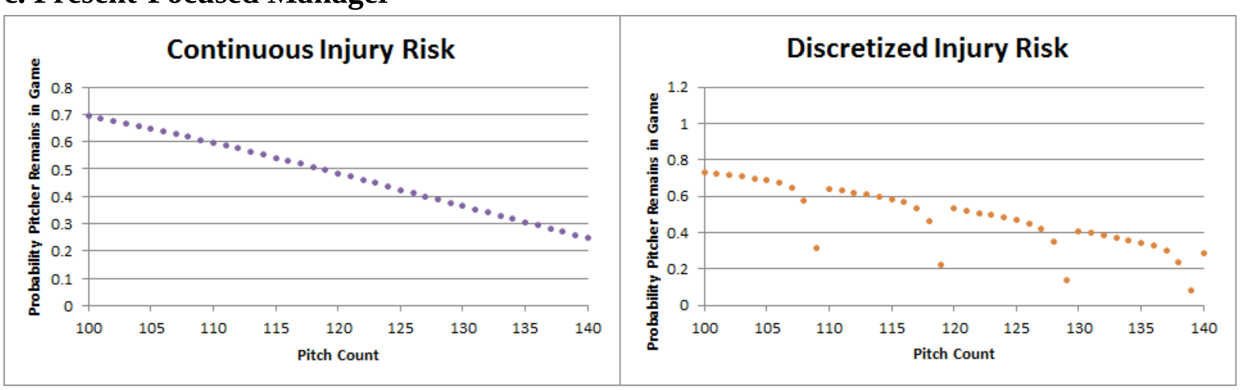
$$A = 0.001, y_0 = 30, T = 160, \mu = 20, \sigma = 3, c_o = 1, \beta = 0.9$$

b. Very Present-Focused Manager



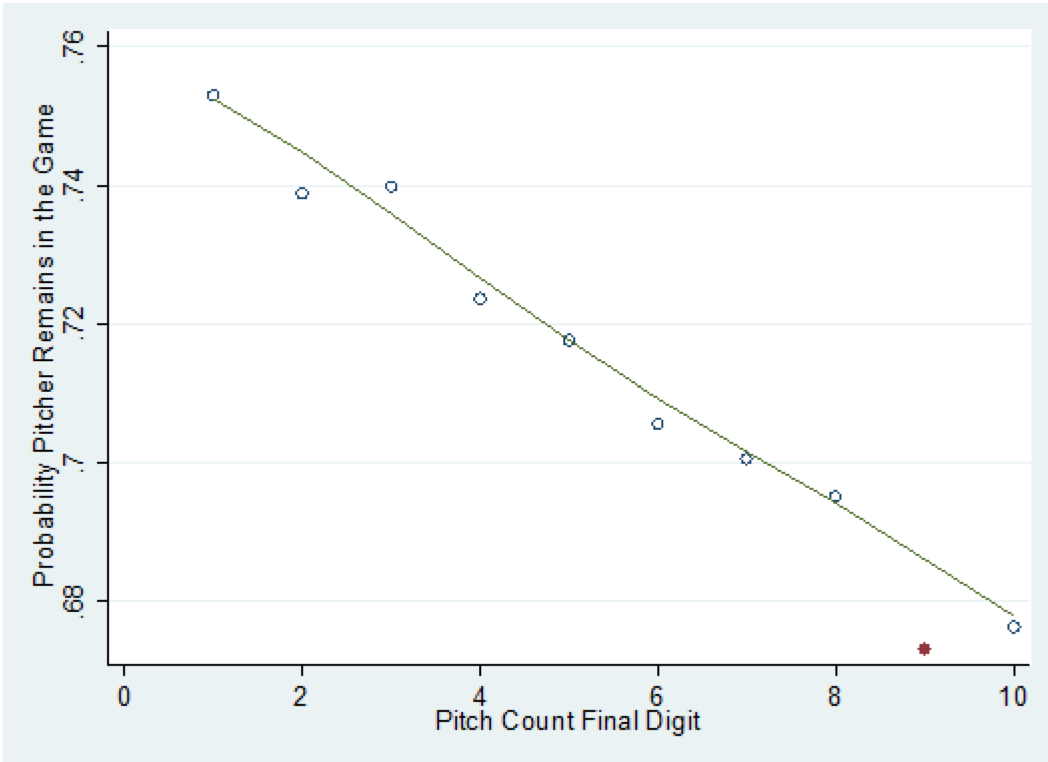
$$A = 0.001, y_0 = 35, T = 160, \mu = 20, \sigma = 4, c_o = 1, \beta = 0.1$$

c. Present-Focused Manager



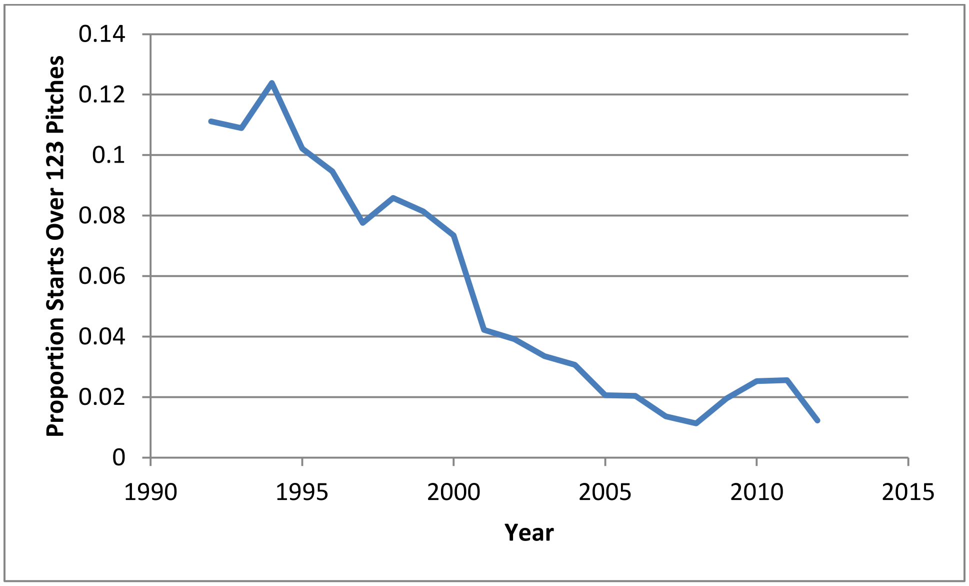
$$A = 0.001, y_0 = 35, T = 160, \mu = 20, \sigma = 8, c_o = 1, \beta = 0.3$$

Figure 2. Probability Manager Leaves Pitcher in Game, 1992-2012



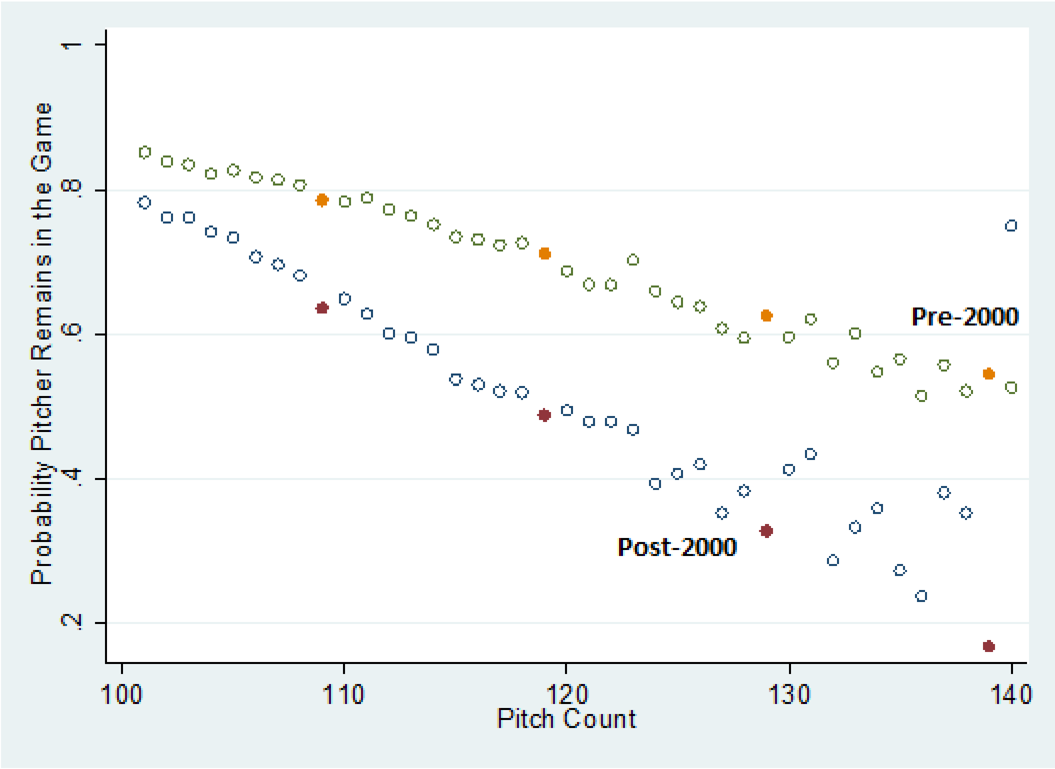
The plotted points show sample means by pitch count for all plate appearances 1992-2012 in which a starting pitcher is in the game and the pitch count is between 101 and 140. The fitted curve results from a regression of a starter remains dummy on a quartic polynomial of the pitch count. Source: author's calculations using Retrosheet data

Figure 3. Number of High Pitch Count Games by Year



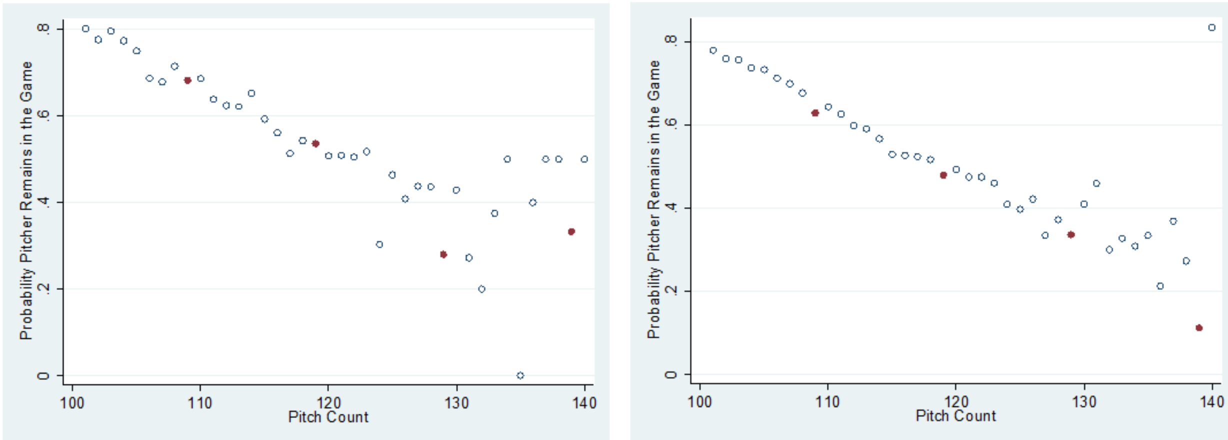
Source: author's calculations using Retrosheet data

Figure 4. Probability Manager Leaves Pitcher in Game, 2001-2012 vs. 1992-2000



The plotted points show sample means by pitch count for all plate appearances in which a starting pitcher is in the game and the pitch count is between 101 and 140. The green and orange (upper) set of points shows 1992-2000 with pitch counts ending in nine orange (solid). The blue and red (lower) set of points shows 2001-2012. Source: author's calculations using Retrosheet data.

Figure 5. Probability Manager Leaves Pitcher in Game, 2001-2012



Tie Game **Not Tie Game**

The plotted points show sample means by pitch count for all plate appearances 2001-2012 in which a starting pitcher is in the game and the pitch count is between 101 and 140. Solid red points show pitch counts ending in 9. Source: author's calculations using Retrosheet data.

Table 1. Summary Statistics

	All Plays	Study Sample	Study Sample, Not Nines	Study Sample, Nines	Difference	Difference (Adjusted for Pitch Count)
Starter	0.66	1.00	1.00	1.00	--	--
Pitch Count	40	109	109	113	4***	--
Pitcher Remains	0.91	0.73	0.73	0.67	-0.057***	-0.011**
Wins Game	0.48	0.57	0.57	0.57	0.004	0.000
Opponent Runs to End of Game	2.52	1.41	1.42	1.31	-0.110***	-0.001
Next Batter On Base	0.343	0.337	0.338	0.332	-0.006	-0.007
Next Batter Number of Bases	0.415	0.406	0.407	0.399	-0.008	-0.008
Runs Allowed Previously	2.39	2.21	2.21	2.21	-0.003	-0.004
Runs Scored Previously	2.42	3.30	3.30	3.34	0.046*	-0.021
Current Batter On Base	0.343	0.342	0.342	0.337	-0.006	-0.006
Current Batter Number of Bases	0.415	0.405	0.405	0.394	-0.011	-0.007
N	3,843,509	175,965	164,346	11,619		

Statistical significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively using standard errors clustered by game. The study sample includes all plate appearances in which the starting pitcher has remained in the game through the plate appearance in question and has a pitch count of at least 100.

Table 2. Manager's Decisions to Keep or Remove Pitcher

Dependent Variable:	(1) Linear Stays In	(2) Quartic Stays In	(3) Time Period Stays In	(4) Close Games Stays In
Pitch Count Ends in Nine	-0.016***	-0.011**	-0.000	-0.027***
	(0.005)	(0.005)	(0.006)	(0.007)
Post X Ends in Nine	--	--	-0.023**	--
			(0.009)	
Tie Game X Ends in Nine	--	--	--	0.026
				(0.019)
Pitch Count Linear	Yes	No	No	No
Pitch Count Quartic	No	Yes	Yes	Yes
Pitch Count Quartic X Post	No	No	Yes	No
Pitch Count Quartic X Tied	No	No	No	Yes
N	175,965	175,965	175,965	90,359
Sample	All Years	All Years	All Years	Post Period

Statistical significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively. Standard errors clustered by game are in parentheses. The sample includes all plate appearances in which the starter has remained in the game thus far with a pitch count of at least 100.

Table 3. Team Performance Following Manager's Decision

Dependent Variable:	(1) Runs to End	(2) Runs to End	(3) Team Wins	(4) Team Wins
Pitch Count Ends in Nine	-0.00	0.03	0.000	0.001
	(0.02)	(0.02)	(0.004)	(0.006)
Pitch Count Quartic	Yes	Yes	Yes	Yes
N	175,965	90,359	175,965	90,359
Sample	All Years	Post Period	All Years	Post Period

Statistical significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively. Standard errors clustered by game are in parentheses. The sample includes all plate appearances with a pitch count of at least 100.

Theory Appendix

To derive the reservation productivity formula, start with the value function for a starting pitcher:

$$V_p^S = \sum_{i=1}^{k} \lambda_i \left\{ \left[\sum_{j=1}^{i} \beta^{j-1} g(p+j) \right] + \beta^i E\left[\max\left\{ V_{P+i}^R\left(y_{p+i}\right), V_{p+i}^S \right\} \right] \right\}$$

Express future values as excess returns relative to the value of remaining with the starter.

$$V_p^S - \sum_{i=1}^{k} \lambda_i \beta^i V_{p+i}^S = \sum_{i=1}^{k} \lambda_i \left\{ \left[\sum_{j=1}^{i} \beta^{j-1} g(p+j) \right] + \beta^i E\left[\max\left\{ V_{P+i}^R\left(y_{p+i}\right) - V_{p+i}^S, 0 \right\} \right] \right\}$$

Substitute in from (4) the definition of the reserve value as the value function for a starter and directly substitute in the value function of a reliever from (3):

$$V_p^S - \sum_{i=1}^{k} \lambda_i \beta^i V_{p+i}^S = \sum_{i=1}^{k} \lambda_i \left\{ \left[\sum_{j=1}^{i} \beta^{j-1} g(p+j) \right] + \beta^i \left(\frac{1-\beta^{T-p-i}}{1-\beta} \right) E\left[max\left(y_{p+i} - \bar{y}_{p+i}, 0 \right) \right] \right\}$$

Since $y_p \sim N(\mu, \sigma)$, then $y_{p+i} - \bar{y}_{p+i} \sim N(\mu - \bar{y}_{p+i}, \sigma)$ which makes $E\left[max\left(y_{p+i} - \bar{y}_{p+i}, 0\right)\right]$ just the expected value of a truncated normal distribution. Using well-known results for the expected value of a truncated normal and minor algebra yields equation (4).