Do males differ from females in the way they set and meet goals? An analysis of marathon runners*

Abstract:

We examine strategies adopted by people completing a well-defined but difficult task: running a marathon. We find strong evidence that males relative to females focus on, and care more about, beating a round number time (e.g., 4:00 hours). Our evidence suggests that setting a round number goal helps males perform better in the race. We find that these males are more likely to run at constant pace (the optimal way to run a marathon) and have more energy to speed up ("kick") towards the end of the race. In contrast, females are less likely to set round number times as goals, and run more conservatively at the start, pace themselves better throughout the race, and finish with a stronger kick, irrespective of whether their finish time is close to a round number. Our results also suggest that males benefit more from planning than do females, since we find that less experienced males are more likely to start the race too aggressively and slow down considerably ("bonk") towards the end of the race. Our results have implications for organizations because they suggest that the sexes can subconsciously differ in the strategies they adopt to complete a task and the goals they use to evaluate their performance.

All my life I've had one dream: To achieve many goals – Homer Simpson.

Blessed is he who expects nothing, for he shall never be disappointed – Alexander Pope.

I. Introduction

The objective of our research is to examine whether males and females differ in the internal, implicit goals they set to complete a task that has an uncertain outcome. This is important because much of the prior research on goal-setting has focused on males, has not distinguished males from females, or implicitly assumed that there are no differences.[1] However, other research suggests that males differ from females in their level of risk aversion and competitiveness.[2] These inherent differences could affect strategic goal-setting. If there are differences in the types of implicit goals set by males versus females when completing a task, then this has implications for managing and motivating individuals within organizations that could have been overlooked by prior research.

The task we focus on is running a marathon. The marathon is a 42.195-kilometer (26.2 mile) race popular with recreational runners.[3] The marathon setting offers several advantages for investigating decision-making while completing a task. First, there is no ambiguity over whether the participants want to complete the task. All runners enter the race with the objective of finishing. Second, marathon participants are comprised of a broad cross-section of individuals of different ages and backgrounds and therefore our inferences are more generalizable than experimental settings that typically have fewer and less diverse participants. Third, running a

---

[1] For example, Pope and Simohnson (2011) study major league (all male) baseball players; and Pope and Schweitzer (2011) examine male golfers; studies that do not distinguish the genders include Heath et al (1999) tests of goals as reference in various lab settings and Allen et al (2017) study of marathon runners.

[2] For example, Barber and Odean (2001) conclude that males are far more risk-seeking than females when it comes to their stock investment strategies. One possible explanation for why is that testosterone level influence risk taking and competitive behavior (e.g., Dreber et al (2009)). Croson and Gneezy (2009) review experimental research that is consistent with differences in risk preferences across genders.

[3] There were approximately 803 marathons held in the U.S. and Canada during 2018, with 503,073 finishers - http://findmymarathon.com/statistics.php

marathon is an all-consuming, self-contained activity. This enables us to draw stronger inferences than other field settings in which individuals are often multi-tasking. Fourth, marathon organizers provide each runner with a chip that records his or her start time and split times throughout the race.[4] This allows us to document the runner's strategic choices while completing the task as opposed to just the outcome. Finally, other than a few elite competitors, marathon participants do not receive financial gain, or broad outside recognition, from their performance. Therefore, monetary incentives do not drive behavior.[5] Thus, we view marathon running as a powerful setting for identifying whether genders differ in their strategies and implicit goals when completing a task that requires considerable motivation and effort.

There are many different goals that runners could set for themselves when running a marathon. Probably the most important, especially for the first-time marathon runner, is completing the race. Our sample, by construction, consists of people who finished the marathon and so achieved this goal. However, beyond the goal of finishing, runners could aim to finish within a particular timeframe. The time goal they set will depend on their health, age, physical fitness, training, experience, willpower and motivation. Runners could also set higher-level goals such as beating a prior marathon time, or the time of another person. We do not have survey information on the idiosyncratic internal goals set by the runners before the race. Our data consists of race outcomes (finish times) and we use it to investigate whether males and females differ in their strategies to meet one specific goal: beating a round number finish time (e.g., 3:00, 3:30, 4:00 or 4:30 hours).

[4] Split times refer to the amount of time it takes to complete a specific distance. As discussed later, in our sample the split is approximately each 10 kilometer interval of the marathon.

[5] For example in 2012, the Chicago Marathon paid a total of $500,000 to the top ten finishers of each gender. This means that approximately one half of a percent of total finishers received prize money. http://www.bestroadraces.com/bank_of_america_chicago/marathon

We focus on round number times because Allen et al (2017) show that marathon runners 'bunch' at round number times in a manner consistent with them viewing these times as goals. However, they do not examine whether genders differ in the use of the goal. This benefit of examining exact round-numbers is that the bunching appears to be inconsistent with the goal of running the race in as fast a time as possible since one would expect a relatively smooth distribution of finish times, reflecting the vast array or skill, training, and physical ability of runners. In addition, the round-number finish time is not an explicit goal set by race organizers, and it is not meaningful in and of itself, (the numbers 3:00 and 4:00, for example, are just numbers on a clock).[6] So focusing on round-number times means that any gender differences that we observe must be due to psychological differences rather than physiological/biological differences, as discussed subsequently. That is, the physical ability of a runner who finishes the race in 3 hours and 59 minutes is unlikely to differ materially from a runner of the same sex who runs 4 hours and 5 minutes.

We conduct our analysis using the 'full-split' sample of marathon runners from Allen et al. (2017) that provides pace data over the course of the race (i.e., the time of the runner at the 10km, 20km, 30km, 40km, and 42.195km marks). This sample includes marathons from 2002-2013 and has over 700,000 finishers with approximately 60 percent of the runners being male. We first examine whether the genders differ in their likelihood of setting a round number as a goal. We find that 18 percent of men and 14 percent of women have times within 4 minutes of a round number. Our statistical tests reveal that the excess mass for men is 12% and for women is

[6] Allen et al (2017) remove finish times that could be due to the goal of qualifying for Boston and find that this goal is not driving the clustering since the clustering is still strongly observable in the data. In section 6 we remove all runners whose nearest round number finish time is also their Boston qualifying time. We find the tenor of all results holds with this restriction.

6%, with these differences being significant, suggesting that males are more likely than females to use specific, implicit time goals to evaluate their performance.

We then examine how the strategies adopted by runners correspond to use of a round number goal. Elite runners know that the optimal strategy for running a marathon is to run at a constant pace. Medical research suggests that deviating from this strategy can cause the runner to run a less efficient (i.e., utilize more energy) and more risky (i.e., greater chance of injury) race, with the expected outcome of a slower finish time. If a runner's primary goal is to finish the race in the best possible time, then he or she should run the race at a constant pace. We define an optimal pacing strategy as one in which the running pace does not deviate by more than 5% over the first 40kms of the race. Utilizing this strategy requires willpower and planning and is unlikely to happen by chance. Therefore, if runners plan to beat a round number time, then we expect to observe a greater magnitude of bunching for runners using the optimal running strategy, regardless of gender.

Our results suggest that males who meet round number finish times behave more "optimally" and/or better "plan" than other males. Specifically, we find that the bunching of finish times at round numbers is driven by men who run an optimal race. This result is particularly evident for faster finish times (e.g., 3:00 hours), suggesting that where the margin for error is higher and the effort level more intense, males engage in greater pre-race planning. It is also consistent with more experienced and competitive males being more forward-thinking in their racing strategy. In contrast, although females are unconditionally more likely than males to run an optimal race, optimal and non-optimal females do not differ in the amount of bunching at round-number finish times. This result suggests that a woman's strategic choice to run an optimal race is not motivated by the desire to beat a round number time goal.

We next investigate whether runners who meet round number times engage in a more short-term, potentially costlier, and more effort-intense strategy to meet their goal. Specifically, because we have data on how quickly runners complete the last 2.195 kilometers of the race we can determine whether a runner sped-up to meet a time goal. If a runner has any energy in reserve then he or she should exert all remaining energy over the final portion of the race to finish in the fastest possible time. However, if the runner does not care about the finish time or is in easy reach of their time goal, then there is no need to exert extra effort and speed up.[7] Consistent with males caring more about round number times, we find that they are more likely to speed up towards the end of the race to meet their time goals and that this tendency is strongest among males who run optimally. In contrast, we find that the likelihood that a female speeds up is only marginally affected by her closeness to beating a round number time. These results suggest that males relative to females will push themselves harder when it will help them meet their implicit but arbitrary time goals.

Our results also highlight that training and experience appear to be more important for race performance for male runners. Specifically we find that males are significantly more likely than females to slow down considerably towards the end of the race, a phenomenon colloquially known as "bonking." We find that 36 percent of males versus 21 percent of females bonk. This result suggests that the typical male runs too aggressively in the early parts of the race, increasing the likelihood of obtaining a strong finish time (should he be able to keep up the pace), but greatly increasing the risk of failure (obtaining a poor race time). Further, we find that this result is not driven by biological/physiological differences between males and female

[7] A person who aims to run a marathon in under 4:00 hours will need to run at a constant pace of 9 minute 9 second miles (5.58 minutes per kilometer) and hit the 40km mark by 3:47:26 so that they can then run the last 2.2kms (1.36 miles) in 12 minutes and 28 seconds. Therefore, a runner can determine whether a reasonable level of accuracy at the 40km mark how feasible it is to meet a time goal.

because more experienced male runners are less likely to bonk, suggesting that a better trained male is less likely to make the mistake of going out too aggressively.

Our results provide two contributions to the literature. First, our study highlights that in the marathon setting, in which individuals are motivated by their own implicit goals, males and females adopt different strategies to complete the task (i.e., finish the race). Females appear to be more concerned with finishing the race within a broader time bandwidth and adopt a conservative but scientifically supported approach to ensure they complete the task with less effort and risk of injury. The actual time (in terms of the exact minute that they finish) appears to be a less important goal. In contrast, a larger proportion of males appear to set themselves a very tight and specific finish time (e.g., under 3:00 hours or under 4:00 hours) and will strategize both before and during the race to ensure that they meet their goal. However, setting ambitious goals comes with risk: males are more likely than females to go out too aggressively and “bonk,” with the result being a poor race outcome.

More broadly, our results suggest that in the absence of clear evaluation metrics, genders can differ in their perception of what it means to “successfully” complete a task. This finding has implications for research on performance evaluation and decision making by individuals within an organization (e.g. Locke and Latham 2002; Latham and Locke 2007; Presslee et al 2013; Feichter et al 2018). Specifically, our results suggests that males could be more likely than females to set themselves specific targets even when management might prefer them to try to optimize on other dimensions.

The remainder of the paper has four sections. The next section provides our predictions. Section 3 provides our research design and section 4 and 5 discuss our data and our empirical results, as well as a discussion of the potential implications of our results. Section 6 concludes.

2. Predictions

2.1 Setting goals in a marathon

We assume that all marathon runners have one clear goal – to finish the race. If a runner's only motivation is to meet this goal, then, in order to reduce the chance of injury, he or she should run/walk at a leisurely pace that ensures a finish within the timeframe allowed (usually six to eight hours). However, most marathon runners appear to be exerting effort when they run, suggesting that they have other goals beyond just finishing. Given the lack of external payoffs associated with their performance, any goals beyond finishing must be purely intrinsic and are not a result of financial incentives or contracts.[8]

Consistent with Markle et al (2018) we assume that runners define these goals in terms of their actual final finish time, as opposed to their relative placement in the race.[9] The finish time is a summary measure of performance and reflects choices made leading up to, and during the race. It has the added benefit of being objective, readily observable, and allows us a clean comparison of the goals used by each of the genders. While there are a variety of possible finish time goals runners could have, we investigate whether males and females differ in their use of the exact round-number finish time goal. We focus on this goal because Allen et al. (2017) document that runners "bunch," that is a greater than expected number of finishers, at round-number times, suggesting that these times are used as reference points.[10] The benefits of focusing

[8] This feature differs from the majority of field studies that involved environments with explicit financial payoffs (see Camerer 2000 and Barberis 2012 for a review). Most races require an entry fee (e.g. $150 for the 2012 Chicago Marathon) as well as other training and travel expenses, therefore most runners will lose money from the experience. However, extrinsic goals do exist in marathons. The first extrinsic goal is placing in the race and receiving a financial reward. This applies to only a few very fast individuals and is unlikely to motivate the majority of runners. The second extrinsic goal is obtaining a time to qualify for the Boston Marathon. We know the age of racers and their finish times, when we exclude individuals who qualify for Boston results are similar.

[9] In a survey of marathon runners, Markle et al. (2018) find evidence that pre-race goals are expressed in terms of finish times.

[10] Heath et al. (1999) contend that goals function as reference points and that individuals will act in a way that is consistent with the predictions of Prospect Theory. Prospect Theory predicts that individuals behave differently at

on this goal is that (i) we can compare the bunching of finish times for males and females and determine whether both genders care equally about this goal, (ii) it allows us to examine whether the running strategies of the genders is impacted by having a round number time goal and (iii) we can control for biological differences between sexes by comparing runners of the same sex who are clustered at round numbers to those that are not clustered at round numbers finish times. Thus, focusing on round numbers as goals enables us to make stronger inferences about whether gender differences we observed are due to psychological differences in competitiveness and risk sensitivities rather than physical or biological differences.

2.2 Gender differences in goal setting

Prior research suggests that males differ from females in their levels of competitiveness and risk sensitivities (see Crosson and Gneezy 2009 for a review). Specifically, the results suggest that males are generally more competitive (Gneezy et al 2003), and females more risk averse (e.g. Eckel and Grossman 2008, Barber and Odean 2001). These differences are attributed to a variety of factors including biological differences (Dreber et al 2009), emotional differences, social preferences, and environmental context (Corson and Gneezy 2009). If females are less competitive, then almost by definition, they will be less likely to want to set goals that they have to beat. Likewise if females are more risk averse, then they will prefer not to set a goal, because this creates an unnecessary risk: the possibility of missing the reference point. Running the race in a more "rational" manner "of doing the best job I can, given my training and physical attributes" avoids forming a reference point and so makes for a less stressful race.[11]

---

reference points and argues that because individuals are loss averse they will work harder when they are close to their goal relative to when they are in easy reach of their goal. Prospect Theory also predicts that individuals have diminishing sensitivity and will exert less effort the further away they are from their goal and increase effort (at an increasing rate) as they approach it.

[11] This logic regarding round time reference point is similar to the discussion of goal setting and risk in Byrnes et al (1999). There they discuss how the act of attempting to meet a goal qualifies as risk-taking when "a) the behavior in question could lead to more than one outcomes, and b) some of the outcomes are undesirable (pg. 367)."

In order to determine whether there the genders differ in the use of round-number time goals, we compare the distribution of male and female finish times and test whether there is a difference in 'bunching' around exact round-number finish times (Allen et al 2017). Our first prediction is:

H1: The bunching at round number finish times is greater for males than for females.

Note that the null hypothesis predicts no differences in the bunching of males versus females at round number time. This could occur if (i) males and females do not actually differ in their preferences regarding goal setting or (ii) there is a selection bias in marathons towards females who have similar preferences to males.[12]

2.2 Within race strategies

Given the length of the marathon, runners can maximize the likelihood of meeting their goals by planning in advance. Physiological research suggests that an "even-pacing' strategy is the optimal approach for long distance races. In their review of pacing research, Abbiss and Laursen (2008) discuss results showing that an athlete's performance is negatively impacted when their velocity or power drops below their physiological limits at any point during an endurance event. Their research also shows that this effect holds even when an athlete tries to make up for lost time by speeding up at the end of the race. Utilizing this even-pace strategy stops the runner from starting the race too fast (and getting overly tired too early in the race) or too slow (and having to exert too much energy at the end of the race).[13] We assume that runners

[12] There is literature that questions whether males and females differ in competitiveness and risk preferences. For example, Tanaka et al. (2010) find, in their study of Vietnamese villagers, that after controlling for education and income levels, there are no differences in the risk aversion of genders. In a more targeted study, Atkinson et. al (2003) examine a subsample of mutual fund managers and find that after controlling for experience and other characteristics the two genders have similar levels of performance and risk preferences.

[13] For the more academically minded we can also describe an optimal strategy in terms of how the body uses energy. Energy is supplied to the body by complex chemical reactions. The end result is a compound called adenosine triphosphate (ATP). ATP can be produced aerobically (with the use of oxygen) or anaerobically - using ATP stored in the muscle cells. Anaerobic production is quick, and readily available and is primarily what is used for sprinting,

who are able to execute the even-pace strategy (we term an "optimal" running strategy) are better planners than those who are not.

We examine whether the likelihood of running an optimal race is associated with round number bunching within each gender. Specifically, we compare the strategies of females/males who bunch at round numbers to people of the same sex who do not bunch at round numbers. This allows us to control for general differences between the sexes that could be due to biological factors. [14] We expect that runners are more likely to run an optimal race when they have a specific round number goal because they are likely to have planned in advanced to meet the goal.

H2a: Male runners who beat round number finish times are more likely to run an optimal race than other male runners, because this is the best running strategy for achieving their goal.

H2b: Female runners who beat round number finish times are more likely to run an optimal race than other female runners, because this is the best running strategy for achieving their goal.

Our null hypothesis is that since an even-pace running strategy is optimal, all runners should engage in this strategy whether or not they have a round number time goal.

Our third prediction concerns end of race effort. The benefit of focusing on the final

---

but it is available in limited quantity. Aerobic production produces large amounts of ATP but takes longer to produce. When running, the body uses both aerobic and anaerobically produced ATP. A by-product of anaerobic energy production is lactic acid. Lactic acid can also be used to produce ATP but the body needs sufficient oxygen in the blood to do so. As the athlete pushes harder a point is reached where the body can no longer process lactic acid fast enough, and so lactate begins to accumulate at an increasing rate in the blood. This point is called lactate threshold (LH) and when this is reached breathing becomes labored and a burning sensation is felt in the muscles. A runner can only sustain this pace for a limited time until discomfort slows him down. The "optimal" pace for long distance runners is where they stay in "Zones 1 through 3," where they are using aerobically produced ATP and are close to their aerobic threshold but are below their lactate threshold (see Bernhardt (2006, Chapter 2). Running even only slightly faster than the lactate threshold over a long distance can result in the runner "hitting the wall" and needing to slow down considerably.

[14] One benefit of examining the relationship between round number goals and pacing strategy is that it allows us to determine whether psychological factors impact pacing strategies. Prior research (e.g., March et al (2011), Deaner et al. 2014, Santos-Lozano et al 2013, and Smyth 2018) find that females tend to run more optimally. However, this research focuses on physiological explanations for the differences. One exception is Hubble and Zhao (2016) who use corral times for finishers of the Houston Marathon and find that men are more likely to miss the corral time than females. They attribute this to males' greater over-confidence, however other explanations such as a greater desire to avoid starting with slower runners could also drive the result.

interval is that there is little uncertainty as to whether the runner will finish the race. As a result, the expectation is that any runners with remaining energy reserves will speed-up, or ‘kick’, over the final interval (last two kilometers) in order to finish in as fast a time as possible. However, Allen et al (2017) provide evidence that the ‘kick’ is contingent on how big of a risk there is of missing a round number finish time. They argue that this is consistent with prospect theory’s prediction that individuals engage in costly and effort inducing actions to avoid a sense of loss from missing a reference point. If males are more competitive and less risk averse than females and set round number times as implicit goals, then we expect males to feel a greater sense of loss from missing a round number than do females:

H3: Males are more likely than females to speed up at the end of the race to meet or beat a round-number time goal.

The null hypothesis is that all runners who have energy remaining at the end of the race will speed up, regardless of whether they have a round-number time goal.

3. Research Design

3.1 Bunching around specific round-number finish times.

We use the same methodology as Allen et al (2017) to test for bunching in the distribution of marathon finish times. This methodology stems from literature examining individual taxpayer responses to “kinks” in the tax code (e.g. Saez 2010, Kleven and Waseem 2013). If runners use the exact round-number finish times to evaluate their performance, and they are able to perfectly adjust their effort to meet them, then we would expect to see a “hole” in the distribution of finish times at round numbers, because runners near the times will experience a strict gain in utility by beating them.[15] However, in our setting, as in Chetty et al (2011), the

[15] This is similar to the “notch” described in Kleven and Waseem (2013), where they find a sharp discontinuity at a region in the Pakistani tax code where taxpayers can increase both leisure and consumption by moving to a region just below the discontinuity.

observed bunching is likely to be diffuse rather than a point mass. Runners are unlikely to be able to perfectly control their effort levels over the course of the race. They may underestimate the amount of energy they have left, incorrectly calculate the required pace to meet the benchmark, or build a cushion into their pacing that causes them to beat the round number by more than a small amount. As a result, rather than seeing a sharp increase in runners just beating the round number finish times and then an immediate drop, we expect to see a bunching of finish times around them. This will reflect runners who attempt to meet the goal time and just miss, as well as those who beat it by just a small amount.

To calculate whether there is significant bunching at round-numbers, we need a counterfactual distribution of finish times. If runners do not use round-numbers as goals then we expect the distribution of finish times to be approximately normal reflecting the distribution of ability across the sample. Consistent with Allen et al. (2017), we set the bunching region as up to four minutes faster than the exact round-number finish time and compare to the actual number of finishers in that region to the counterfactual.[16] We refer to the percentage difference between the actual and counterfactual as 'excess mass' which is our prime measure of bunching. We then test the significance of the excess mass using the bootstrap technique described in Chetty et al (2011). Appendix A provides a detailed description of the approach.

3.2 Constant pace running for an optimal race.

We use constant pacing as the baseline for our evaluation of whether a runner exhibits an optimal pacing strategy (e.g., Santos-Lozano et al (2014), March et al (2011), Hubble and Zhao (2016), Deaner et al (2015)). We acknowledge that for most runners, running at a constant effort level will not result in a perfectly constant pace over the course of the race. Indeed, given the

[16] We use this bunching window for consistency with Allen et al (2017). The results are robust to extending the window to include finishers who just 'miss' by up to one minute as well.

physiological demands of the marathon, most runners will exhibit a slowdown over the second half of the race even when they keep their effort levels constant. Still, the closer they are to maintaining a constant pace, the more likely they are attempting to engage in the optimal strategy.

To construct our measure of optimality we examine how much the runners pace per-kilometer varies over the course of the race. Similar to Santos-Lozano et al (2014), we calculate the coefficient of variation (CV) for each runner over each of the first four 10 kilometer intervals. We define the runner as having run an optimal race when the coefficient of variation is less than five percent. That is, the runner's pace in a given interval does not deviate by more than five percent from the pace in any other interval. [17] This allows for the fact that differences in elevation profiles (or other external conditions) can cause the pace to vary between intervals even while the runner is exhibiting an overall constant effort level.

We also use split data to calculate the 'pace ratio' for each interval. The pace ratio is defined as: the pace-per-kilometer for the current interval divided by the pace-per-kilometer for the first 10km of the race. For example, if over the interval from the Halfway point to 30km the runner runs a pace of 6:30 per-km, and their pace over the first 10km was 6:20 per-km, then he or she will have a pace ratio of 1.02 (6:30/6:20 = 1.0161) for that interval. The closer the pace ratio is to 1, the closer the runner is to the optimal strategy. Examining the pace ratio allows us to determine not only whether the runner ran an optimal race overall, but also the running strategies adopted during the race.

3.3 Likelihood of speeding up over the final interval.

As in Allen et al (2017), we examine whether there is a difference in the propensity to

[17] We use a variety of different cut-offs for optimality ranging from 3% to 10%. The results are consistent regardless of the measure choses.

speed up at the end of the race based on a runner's proximity to the round number goal. We define a runner as being in a "loss" position when maintaining r current pace at the 40 kilometer point of the race will result in them missing the nearest round number time. If a runners view the round-number finish time as a goal, then a runner in a "loss" position (projected to miss the round-number goal) will be more likely to speed up than a runner in a "gain" position (projected to beat the round-number goal). And runners in a small "loss" position (projected to just miss) will speed up more than runners in a large "loss" position (projected to miss by a large amount). Running faster involves a small increase in risk for the runner, because speeding up at the end of a marathon can increase the chance of exhaustion and injury, both of which could impact performance. In addition, speeding up may be painful because the body is more likely to meet the lactic threshold. In contrast, a runner who has beaten the goal will have no incentive to speed up since the gain from beating the benchmark by a slightly greater amount is unlikely to outweigh the physical pain of the effort.

We compare the pace-per-kilometer over the final 2.19 kilometers to the pace-per-kilometer for the first 40 kilometers. If the pace over the final interval is faster, then we classify the runner as having sped up. We then partition the runners into groups based on their projected distance from the nearest round number finish time, and subtract the gender average for speeding up across the sample. We call this the 'conditional probability' of speeding up. Thus, our measure captures the incremental likelihood of speeding up, relative to expected likelihood for the gender.

4. **Data and sample construction**

We use the sample of marathon runners from Allen et al (2017) to perform our analysis. Because of our desire to measure the runners' pacing decisions we restrict our analysis to the ‘full-split’ sample that has complete 10-km, half-marathon, 30-km and 40 km split times, as well as the actual finish time. We eliminate observations when the gender of the runner is missing, the finish time is less than two or more than eight hours, and the calculated pace-per-kilometer for any interval is faster than the current 10 kilometer world record pace (two minutes and thirty-six seconds per kilometer). This leaves us with a sample of 763,903 observations.

Table 1 provides the marathons and years included in the sample, as well as the percentage of finishers in each race that are female. We have marathons from a wide variety of places and over many years, which increases our confidence that our findings are not due to gender differences in subpopulations. One point to note is that the US has a far stronger representation of female runners than other countries. However, the proportion of runners at international events is less than ten percent of the sample (65,873) and so cultural differences are unlikely drive our results.

Table 2 provides descriptive statistics for the sample as a whole and by gender. The finishers have an average age of 39 and an average finish time of 4:40:26. A comparison of the genders reveals that marathon running is more popular with male runners with 61% percent of the participants being male; the average male is older than the average female (40 versus 36), and the average male runs the marathon approximately 30 minutes faster than the average female. [18]

5. Empirical Results

5.1 Bunching at Round Number Finish Times

Figure 1 Panel A provides the distribution of finish times in minute bins for all runners

---

[18] The sample statistics compare favorably for the demographics for all marathons in the U.S. For example in 2010 the average marathon finish time was 4:38:25, the average finisher age was 38.8, and males composed 58.8% of the finishers. http://www.marathonguide.com/features/Articles/2010RecapOverview.cfm

(in orange). Consistent with Allen et al (2017) the bunching around the round number finish times is clearly visible. Figure 1 Panel B displays the total distribution but here we show the distribution of finish times for females (in grey). The incremental orange shading in each minute bin are the male finish times. The visual comparison of Panel A to Panel B highlights that the bunching at round numbers appears to be driven by males. The female distribution has a much smoother shape with the only visible spikes being at 4:00 and 4:30 hours.

Table 3 provides our more formal analysis of whether the male and female distributions differ with respect to their bunching at round-number finish times. Panel A indicates that approximately 18 percent of males versus 14 percent of females are bunched at round number finish times (i.e., their finish time is within 0 to beating a round number time by 4 minutes). However, this does not take into account the expected level of bunching. Table 3 Panel B and Panel C provides the excess mass, as a percentage of the counterfactual finishers, around each of the round-number goals, for males and females respectively. We use the same round-number finish times as Allen et al (2017) but add in 5:30 because females could potentially use this as a goal. The excess mass is significantly greater than zero for seven of the nine round number finish times for the males versus five for females. In addition, the magnitudes of the excess masses are consistently larger for males, with the only exception being at 3:20 where females are more highly clustered (9.3%). Men do not appear to view 3:20 as a goal, preferring 3:10 (9.1%) or 3:30 (13.5%).

Table 3 Panel D consolidates the total excess mass for male and female finishers identified in the bunching regions. For this analysis we take each runner in the bunching region and calculate his or her distance from the round number finish time by subtracting the round number time from the actual finish time. Panel D indicates that there are approximately 12 percent more males and 6 percent more females than expected based on the counter-factual. The difference of 6 percent is significant (Z-Statistic of 4.92). This result suggests that while both genders appear to exhibit some bunching at round-number finish times, males are significantly more likely to do so. This result is consistent with *H1* and suggests that males are more focused on round number finish times and use them as implicit goals.

5.2 Optimal runners and round number time goals

Our next test examines differences in pacing behavior of males and females and how they relates to round-number time goals. Table 4 Panel A provides descriptive statistics on the proportion of optimal runners that are male versus female. Consistent with prior research, females are more likely than males to run at a constant pace (30% versus 24%). Optimal runners also appear to be better runners, since their times are significantly faster than other runners (42 minutes faster for males and 46 minutes faster for females). This suggests that optimal runners have engage in more training and are likely to be more experienced and better pre-race planners than the average runner. The results also highlight that proportionally more women are better at running an optimal race. This unconditional difference across the genders could be due to both physical and psychological reasons.

Table 4 Panel B investigates whether proportionally more of the optimal runners bunch at round-number finish times. The Panel indicates that 32% of males at the bunching regions are running an optimal race and represent 24.5% (27,129/110,694) of the total males who are running optimally. The panel also indicates that 36% of females at bunching regions are running an optimal race and represent 17% (15,307/89,658) of the total females who are running an optimal race. Thus although more females run an optimal race, proportionally fewer (17%) of them are clustered at round number times than males (24.5%).

Figure 2 provides a visual representation of male clustering at round numbers separated by whether the male ran an optimal race. Panel A provides the male distribution of marathon times, and inserts the distribution of non-optimal males (in grey). The orange area represents the optimal males and there are two points worth noting from the figure: (i) optimal males are clustering more strongly at round number finish times (e.g., 4:00 hours and 3:00 hours) and (ii) optimal male runners are skewed towards the faster end of the distribution. Below three hours there are few kinks in the non-optimal distribution, and all the kinks are being driven by the optimal males. Panels B and C further highlight these differences. Panel B provides the distribution of optimal males and Panel C of non-optimal males. The clustering of finish times are more extreme for the optimal males and are skewed at times below four hours. Thus faster males, who are likely to be more experienced and serious about running, use better strategies than other males to run a marathon.

Figure 3 provides the same distributions as Figure 2 for females. Panel A provides the distribution for all females, with non-optimal females inserted on to the figure. Similar to the male distribution, optimal females tend to be running faster. There is some evidence that optimal females view 4:00 hours and 4:30 as a goal (see Panel B), but the rest of the distribution shows little other evidence of bunching. There is almost no evidence of bunching in the non-optimal female distribution, suggesting that these females do not view round numbers as goals (see Figure 3 Panel C).

Table 4 Panel C provides our statistical tests to determine whether more optimal runners cluster at round number times. The results indicate that the consolidated excess mass of 16.1% of optimal male runners is larger than that of non-optimal male runners (9.9%) and difference of 6.20% is statistically significant (Z=4.30). The results for females however, give a different picture. The consolidated excess mass for optimal females is not statistically different from non-optimal females (6.4% versus 5.5%). Therefore, females do not plan to run at an optimal pace to ensure that they beat a round number time. The results in Table 4 therefore support prediction *H2a* for males. Males who bunch at round number finish times are more likely to have run the race at an optimal pace, consistent with planning and setting the round number finish time as a goal. However, the results do not support *H2b* for females. The female likelihood of running an optimal pace is not influenced by whether they have a round number goal.

5.3 The final "kick" and round number time goals

While the prior analysis focusing on pacing up to the final interval, we next examine whether males and females exert additional effort at the end of the race and run faster to beat round number time goals. We begin our analysis in Table 5 by providing descriptive evidence on how optimal versus non-optimal runners pace themselves during the race. Recall that the pace ratio is equal to 1 when the runner is running at the same speed over the given interval as they ran the first 10km (within a 5 percent margin of error). By definition, optimal runners do a good job of pacing as can be seen in both panel A (for Males) and Panel B (for females).

When we compare genders, we find some striking differences. First, regardless of whether the runner exhibits an optimal strategy, female runners start out relatively slower than males. For example, the pace ratio for optimal females is lower than for optimal males over the entire first half of the race. However, by starting out relatively slowly females are able to conserve energy and exhibit a larger propensity to kick (40% vs. 26%) over the final interval. While this more conservative strategy allows females to retain more energy and exhibit a lower deterioration in effort, the larger kick may indicate they held too much in reserve and could have finished with a faster time.

Males, on the other hand, start out relatively faster in the early portions of the race. Optimal males speed up in the second interval (pace ratio of 1.01), and even the non-optimal males are more likely to maintain pace or speed up in the second interval than comparable females (37% vs. 19%). The optimal males maintain the relatively faster pace until the 30km point, whereas the non-optimal males pace drops below the non-optimal females after the half-way point and stays there for the remainder of the race. While this strategy increases the potential for a faster time, it also increases the risk that males will expend energy too quickly resulting in a sub-optimal outcome. Consistent with this being the case, we find that both optimal and non-optimal paced males slow down by more than females as the race progresses. This is consistent with the aforementioned lower likelihood of a 'kick' over the final interval.

We next examine how proximity to a round-number finish time affects end-of-race effort. Figure 4 provides the conditional probability of speeding up based on proximity to beating or missing the nearest round number finish time. For example, the bin labelled (-0:01) includes all runners projected at the 40km mark to ***beat*** the nearest round number benchmark by 00:01:00 to 00:01:59. The 0:00 bin includes all runners projected to ***beat*** the benchmark by less than one minute. Thus if runners care about meeting a round number time, then we would expect the greatest evidence of speeding up to be evident in the 0:00 bin as these runners face the greatest uncertainty as to whether they will meet the round number goal.

Consistent with the bunching results, regardless of whether the male is running an optimal or a non-optimal race, he will speed up when he is close to missing the goal. However, men who are running an optimal race show a much greater sensitivity to missing the round number time than non-optimal men. For example, there is a marked increase in effort in the miss by 0:01-0:59 bin (conditional probability of 4.3%), a peak in the beat by 0:00-0:59 bin (4.4%)

and an immediate decrease when they are projected to beat by a larger amount (-0.3% conditional probability in the beat by 1:00-1:59 bin). Thus, the likelihood of a male increasing effort at the end of the race is highly contingent on how close he is to meeting a round-number goal.

Females, on the other hand, do show an increase in the conditional probability of speeding up when they are in danger of missing a round number, but it is still significantly less than comparable male runners. For example, in Figure 4 Panel A, we see that the conditional probability is only 3.1% for females in the Miss by 0:01-0:59 bin, vs. 4.6% for males. This is reinforced when we partition the female runners based on whether they run an optimal race (Figure 4 Panel C). In contrast to the males, both groups exhibit only small changes in effort around the round numbers. In addition, the similarity between the two groups reinforces the contention that female race strategy is not driven by a fixation on a round number goal.

In summary, males appear significantly more likely to use an exact round-number finish time as a goal. This increased fixation drives their effort levels at the end of the race. Given the previously documented result that they are more likely to have expended the majority of their energy prior to the final interval, this increased contingent effort relative to females indicates that they are engaging in even more relative 'pain' to meet their goal.

5.4 Additional Analysis

We perform two additional sets of analyses based on our findings discussed above. Our first analysis concerns strategies adopted at the beginning of the race and their implications for race performance. Recall that in Table 5 we find that males start the race relatively faster than females. Going fast at the start is a competitive strategy that will lead to a fast finish time, should the person be able to maintain the pace. However, if the pace is too fast, then this strategy increases the possibility of slowing down considerably, or bonking later in the race.[19] Section 5.4.1 investigates whether the male strategy of starting out relatively fast is a biological advantage or evidence of males competitiveness. Specifically, if males have a physical advantage of going out relatively fast, then both sexes should show similar rates of "bonking." In contrast, if males go out too fast because they are competitive and less risk averse, then we should observe a greater proportion of males bonking at the end of the race.

Our second analysis concerns the role of the Boston Marathon qualifying times. Boston qualifying times are often set as round number times and so we evaluate whether the bunching in Figure 1 is driven by the Boston Marathon "explicit" goal rather than an "implicit" or internal goal set by the runner.

5.4.1 Competitiveness and the Likelihood of Bonking

Table 6 provides our initial analysis of bonking. Following prior research (Buman et al 2008), we classify a runner as having "bonked" when he or she slows down by more than 20 percent during the 30 – 40 km interval relative to the prior 30 km interval. Most runners who bonk will do so around the 30 kilometer, or mile 20, and so this measure should capture runners that have had a serious slowdown in pace due to fatigue or injury/cramping as a consequence of poor race planning (i.e., going out too fast). Panel A reveals that 36 percent of males "bonk" compared to 21% of females. Not only are males more like to bonk than females, they also slow down by a greater amount when they do (33% slowdown for males and 30% for females). This evidence suggests that males are more likely to go out too fast, consistent with greater

[19] Bonking describes the biological condition where the body's stored energy is insufficient to feed the muscles. When this occurs, the runner must slowdown considerably and take nutrition and rest to regain the ability to run.

competitiveness.[20] Interestingly, the results also indicate that bonkers are slower runners even before their effort deterioration (the implied finish times of 4:24 for males and 5:00 for females are slower than for non-bonkers by around 13 minutes). We explore this finding in more detail in Table 7.

Table 7 examines whether experience reduces the likelihood of competitive and risky behavior that leads to bonking. We expect that more experienced runners are better planners and more likely to run an optimal race. Our proxy for "experience" is a runner's speed at the beginning of the race (the first 10 kilometers). Our assumption is that runners who start out relatively quickly have engaged in more training and so better understands his or her limitations. We partition the sample into four groups based on each runners' split-time for the first 10 kilometers of the race. *Elites* are finishers whose 10k split is in approximately the top quintile for the gender. This is a split of less than 45/(52) minutes for males/(females). *Runners* are finishers whose 10k split time places them in approximately the second quintile for the gender. This is a 10k split time between 45/(52) and 55/(62) minutes. *Joggers* are finishers with a 10km split time between 55/(62) and 93/(100) minutes, and *walkers* are those with a split time slower than 93/(100) minutes.

Table 7 Panel A presents the results for males and Panel B females. Consistent with speed being an indicator of planning and experience, we see that the proportion of optimal runners is highest among the elites (36% for males and 46% for females) and declines monotonically across speed categories (e.g., for joggers it is 19% for males and 23% for females). The results also indicate that the proportion of runners that bonk is lower among elites

[20] Buman et al (2008) provide survey evidence from approximately 300 marathon finishers and finds that men are more likely to report bonking during the race. However, they do not evaluate whether this is due to physiological and psychological reasons.

relative to runners and joggers (for males: 26% for elites versus 35% and 38% and for females: 15% for elites versus 21% and 22%). All differences between categories are significant within each gender. Thus elite males and females who are likely to be more experienced are less likely to bonk.

Panel C of Table 7 examines the difference in the likelihood of running and optimal race and bonking across categories for males versus females. All differences are significant, with the exception of the optimal walkers, and the differences are greatest for the jogger category (16%). Joggers are the slower 10K runners, but they are also likely to be less experienced. In addition, this is the group in which physiological differences between the genders are likely to be lowest. The fact that we find the largest difference in bonking for males in this group, suggest that psychological differences (male competitiveness) contribute to the observed differences in bonking.

In summary, the results in Table 7 suggest that males' increased propensity to 'bonk' appears to be driven by both biological and psychological factors. Consistent with biological differences, we find that across all categories males are more likely to bonk than females. However, consistent with psychological factors also playing a role, we find that it is the slower males who are more likely to bonk. These males (joggers) are likely to be less experienced but appear to start the race too aggressively consistent with higher competitiveness and lower risk aversion. In contrast, faster males are less likely to bonk, suggesting that planning and experience combats the male's competitive tendency to go out too aggressively.

5.4.2 Effect of Boston Marathon Qualification

A potential confounding factor in our analysis is that several of the round number finish times that we call implicit goals also correspond to the Boston Marathon Qualifying times for certain runners. For example, the 2019 qualifying time for 35-39 year-old men is 3 hours and 10 minutes. If the bunching around that time simply reflects an effort of those runners to obtain the extrinsic reward of qualifying for the Boston Marathon, then this would significantly change the interpretation of our results. In order to ensure that this is not the case we re-ran the analysis after dropping all finishers whose qualifying time is the same as the nearest round number benchmark. We find that all the results hold.

5.5 Discussion and Implications of Marathon Results for Organizations

Our results suggest that male and female runners differ in the use of internal goals to evaluate their performance in completing the marathon. Further, this difference in goal setting is associated with differences in strategic behavior during the task itself. These results have important implications. First, they suggest that females will be more likely to focus on goals that are related to process, as opposed to outcomes. This may allow females to be more flexible in the strategies they use, and potentially increase the likelihood of an optimal outcome when faced with an uncertain task. However, our results also suggest that this increased focus on the process could lead females to behave too conservatively (i.e., have too much stored energy when they finish the race). This is consistent with prior research that finds greater risk aversion in female experimental participants.

Second, our results suggest that men are motivated by use of a specific internal output goal. If this is the case with other output oriented internal goals, then our results suggest that in circumstances where it matters, such reference points should be made explicit to females. For example, consider the case of evaluating individuals using a performance metric that is a continuous measure. If the marathon dynamic holds, then male workers will be more likely to adjust their effort to ensure they receive a round number score from the metric. Females on the other hand, may not alter their effort to hit an arbitrary implicit benchmark. Therefore, managers should make the reference point explicit in order to encourage both males and females to exert similar effort at these points.

Third, if the use of an implicit goal leads to suboptimal outcomes, then leaders in organization need to ensure that such metrics do not get overweighed in performance evaluation. For example, overweighing the importance of obtaining an average teaching score from a group of students of “6.00” out of 7.00 could encourage male professors to devote extra effort to pandering to students immediately before the evaluations are submitted (i.e., kicking at the end of the race to hit a round number). Similarly, a focus on round number earnings targets, (for example of $1.50 a share), could result in executives making suboptimal investment decisions or engaging in earnings manipulation to meet the target.

Fourth, if males subconsciously focus on specific implicit goals more than females, then this could cause male managers to disproportionately reward male employees who work to achieve these implicit reference points. We encourage future research to examine whether this behavior can explain gender differences in pay and other compensation differentials.

Finally, our results can help inform the significant accounting literature that shows firms will report financial numbers that just ‘beat,’ specific numerical benchmarks (e.g. Burghstahler and Dichev 1997, Dechow et. al 2003, Degeorge et. al. 1999, Beaver et. al. 2007, Burghstahler & Chuk 2011, Jacob & Jorgensen 2007, Stice et al. 2018). This is generally attributed to managerial ‘manipulation,’ either real or accrual based, and explained by more traditional capital market based explanations. Our results suggest that this may simply reflect differences in effort levels at the reference points. It also suggests differences in gender composition of managers could result in differences in the propensity to meet and beat certain reference points.

6. Conclusion

In this study we investigate strategic choices and goal-setting behavior for genders when completing a difficult task that is voluntary and requires self-motivation – a marathon. Our evidence suggests that a larger proportion of males set themselves the implicit goal of beating a round number finish time (such as 3:00 hours, 3:30, or 4:00 hours). We find that the bunching at round number finish times identified by Allen et al (2017) in the marathon distribution are predominantly driven by male runners. We find that the clustering of females at round number times is about half the magnitude of that of males, and occurs at fewer round finish times. The stronger results for males are consistent with males being more competitive than females and setting themselves identifiable goals to beat even when they do not need to. The weaker results for females suggest that women set a goal of either completing the race within a broader timeframe (and hence do not view beating a round number time as special), or avoiding the use of round number goal because they do not want the competitive stress of having to beat this goal, or are unaware that beating a round number time is a performance metric that males value.

We also investigate whether people who bunch at round number finish times engage in better running strategies, consistent with pre-race planning. We find strong evidence that males appear to engage in strategic planning to meet their goal. Specifically, we document that males who bunch at round number times are more likely than other men, to have run the race at a constant pace. Running at a constant pace is the optimal way to run a marathon when the object is to obtain a strong and specific finish time. We also find that males who run at a constant pace are more likely than other males to speed up or "kick" at the end of the race to ensure that they will meet their time goal. These results are consistent with males setting round number times as reference points and wanting to avoid a sense of loss from missing their goal.

In contrast, our evidence for females suggests that a larger proportion of females run the race at a constant pace, consistent with an optimal race strategy. However, we find the motivation to run at a constant pace is not strongly influenced by round number finish times. Specifically, there is no difference in bunching between females who run an optimal and non-optimal pace. In addition, the women show less propensity to "kick" at the end of the race to meet a round number finish time. Our results provide support for research that claims gender differences in psychological preferences contribute to observed differences in behavior. Furthermore, our results suggest that strategies adopted by the genders to complete a task can differ in their level of risk and focus. We hope that these insights will help future research better understand variation in performance across the genders.

References

Abbiss, C. R., & Laursen, P. B. (2008). Describing and Understanding Pacing Strategies during Atheltic Competition. *Sports Medicine, 38*(3), 239-252.

Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2017). Reference-Dependent Preferences: Evidence from Marathon Runners. *Management Science, 63*(6), 1657-1672.

Atkinson, S. M., Baird, S. B., & Frye, M. B. (2003). Do Female Mutual Fund Managers Manage Differently? *The Journal of Financial research, XXVI*(1), 1-18.

Barber, B. M., & Odean, T. (2001). Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics, February*, 261-292.

Barberis, N. C. (2013). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspecitives, 27*(1), 173-196.

Beaver, W. H., McNichols, M. F., & Nelson, K. K. (2007). An alternative interpretation of the discontinuity in earnings distributions. *Review of Accounting Studies, 12*, 525-556.

Bernhardt, G. (2006). *Training plans for multisport athletes.* (2nd ed.). Boulder, CO, USA: Velo Press.

Buman, M. P., Brewer, B. W., Cornelius, A. E., Van Raalte, J. L., & Petipas, A. J. (2008). Hitting the wall in the marathon: Phenomenological characteristics and associations with expectancy, gender, and running history. *Psychology of Sport and Exercise, 9*, 177-190.

Burghstahler, D., & Chuk, E. (2011). What Have We Learned About Earnings Management? Correcting Disinformation about Discontinuities. Working paper.

Burghstahler, D., & Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics, 24*, 99-126.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender Differences in Risk Taking: A Meta-Analysis. *Psychological Bulletin, 125*(3), 367-383.

Camerer, C. (1998). Prospect Theory in the wild: Evidence from the field. In D. Kahneman, & A. (. Tversky, *Choices, Values and Frames.*

Chetty, R., Freidman, J. N., Olsen, T., & Pistraferri, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence From Danish Tax Records. *The Quarterly Journal of Economics*, 749-804.

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature, 47*(2), 448-474.

Dechow, P. M., Richardson, S. A., & Tuna, I. (2003). Whay Are Earnings Kinky? An Examination of the Earnings Management Explanation. *Review of Accounting Studies, 8*, 355-384.

Dreber, A., Apicella, C. L., Eisenberg, D. T., Garcia, J. R., Zamore, R. S., Lum, J. K., & Campbell, B. (2009). The 7R polymorphism in the dopamine receptor D4 gene (DRD4) is associated with financial risk taking in men. *Evolution and Human Behavior, 30*, 85-92.

Eckel, C. C., & Grossman, P. J. (2008). Men, Women and Risk Aversion: Experimental Evidence. *Handbook of Experimental Economics Results, 1*, 1061-1073.

Feichter, C., Grabner, I., & Moers, F. (2018). Target Setting In Multi-Divisional Firms: State of the Art and Avenues for Future Research. *Journal of Management Accounting Research, 30*(3), 29-54.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competative Environments: Gender Differences. *The Quarterly Journal of Economics*, *118*(3), 1049-1074.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as Reference Points. *Cognitive Psychology, 38,* 79-109.

Hubble, C., & Zhao, J. (2016). Gender differences in marathon pacing and performance prediction. *Journal of Sports Analytics, 2*, 19-36.

Jacob, J., & Jorgensen, B. N. (2007). Earnings management and accounting income aggregation. *Journal of Accounting and Economics, 43*, 369-390.

Kleven, H. J., & Waseem, M. (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence From Pakistan. *The Quarterly Journal of Economics*, 669-723.

Latham, G. P., & Locke, E. A. (2007). New Developments in and Directions for Goal-Setting Research. *European Psychologist*, *12*(4), 290-300.

Locke, E. P., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation. *American Psychologist, 57*(9), 705-717.

March, D. S., Vanderburgh, P. M., Titlebaum, P. J., & Hoops, M. L. (2011). Age, Sex, and Finish Time as Determinants of Pacing in the Marathon. *Journal of Strength and Conditioning Research, 25*(2), 386-391.

Markle, A., Wu, G., White, R., & Sackett, A. (2018). Goals as reference points in marathon running: A novel test of reference dependence. *Journal of Risk and Uncertainty, 56*(1), 19-50.

Pope, D. G., & Schweitzer, M. E. (2011). Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes. *American Economic Review, 101*, 129-157.

Pope, D., & Simonsohn, U. (2011). Round Numbers as Goals: Evidence from Baseball, SAT Takers, and the Lab. *Psychological Science*, *22*(1), 71-79.

Presslee, A., Vance, T. W., & Webb, R. A. (2013). The Effects of Reward Type on Exmployee Goal Setting, Goal Commitment, and Performance. *The Accounting Review, 88*(5), 1805-1831.

Saez, E. (2010). Do Taxpayers Bunch at Kink Points. *American Economic Journal: Economic Policy, 2*(August), 180-212.

Santos-Lozano, A., Collado, P. S., Foster, C., Lucia, A., & Garatachea, N. (2014). Influence on Sex and Level on Marathon Pacing Strategy. Insights from the New York City Race. *Inernational Journal of Sports Medicine, 35*, 933-938.

Smyth, B. (2018). Fast starters and slow finishers: A large-scale data analysis of pacing at the beginning and end of the marathon for recreational runners. *Journal of Sports Analytics, 4*(3), 229-242.

Stice, D., Stice, E. K., Stice, H., & Stice-Lawrence, L. (2019). The Power of Numbers: Base-Ten Threshold Effects in Reported Revenue. Working Paper.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and TIme Preferences: Linking Experimental and Household Survey Data from Vietnam. *The American Economic Review, 100*(1), 557-571.

Appendix: Description of measures and methodology

Bonking

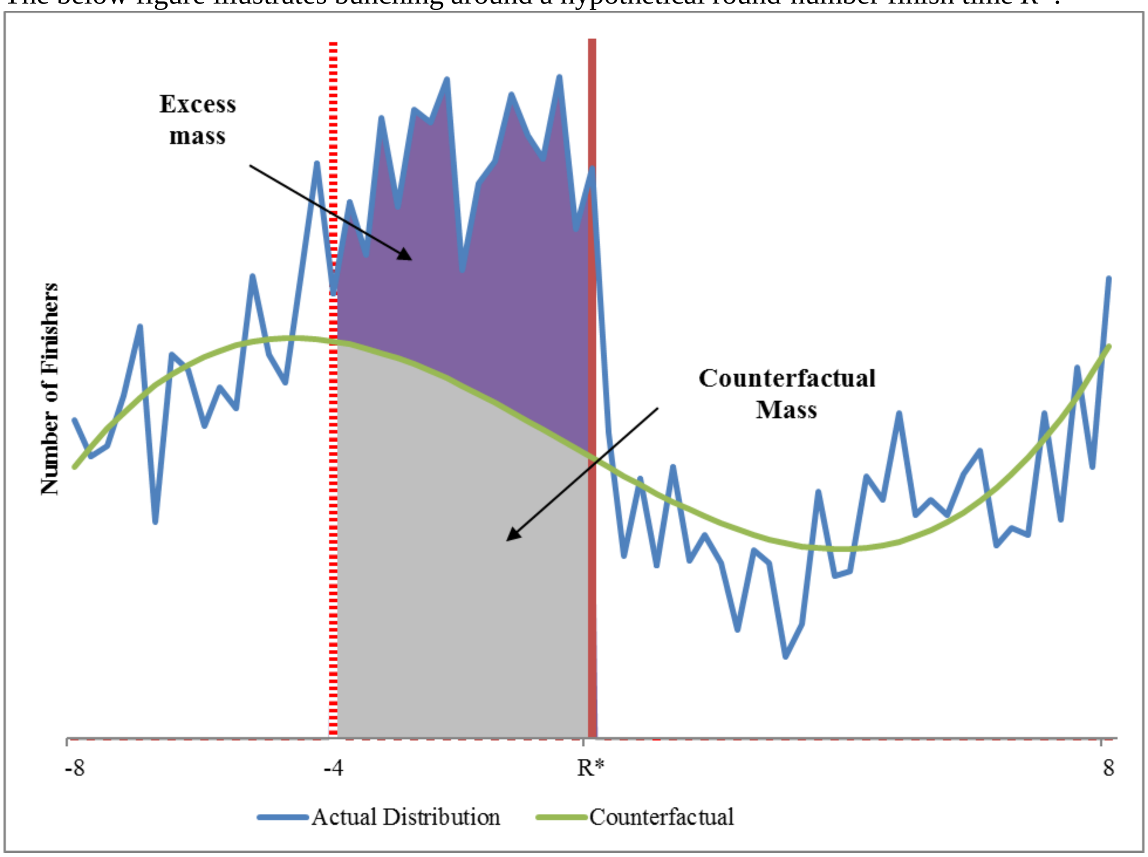
To determine if a runner bonked we compare their pace from the 30km – 40km point in the race to the pace run for the first 30km. If the runner slows down by 20% or more during the 30-40km interval they are classified as having bonked.

Bunching or excess mass percentage

We calculate the bunching around a round-number finish time utilizing the methodology described in Chetty et al (2011). We first set the *evaluation window* around each finish time. This window includes the proposed *bunching area* as well as a set amount of time on each side. The *bunching area* is then determined based on a visual inspection of actual distribution. Consistent with Allen et al (2017) we set the evaluation window to extend for 8 minutes before and after the round number finish time, and bunching area to start at the round number finish time and extend up to four minutes faster. We then remove the actual finishers from the *bunching area* and fit a fifth degree polynomial through the region to calculate the *counterfactual* number of finishers. We then subtract the *counterfactual* from the actual number of finishers in the bunching region to come up with the *excess* finishers in the bunching region. We then calculate significance using the methodology described in Chetty et al. (2011).

Our main variable of interest is in the narrative if the *excess-mass percentage.* This is equal to the excess number of finishers in the *bunching area,* divided by the sum of the estimated number of finishers in the same area as implied by the counter-factual. We then use the estimated standard deviations for these variables to determine whether the *excess mass percentage* is significantly different from zero using a z-test.

The below figure illustrates bunching around a hypothetical round-number finish time *R**:



Consolidated excess mass percentage

In order to get a clear picture of the excess mass around all of the potential round-number time goals in one summary measure, we calculate the difference between the actual finish time for each runner from the set of proposed reference points (for the round number reference points this is the set of 3:00, 3:10, 3:20, 3:30, 3:40, 4:00, 4:30, 5:00, 5:30 and 6:00. We then take the difference between the actual finish time from each potential reference point, and find the one that is the closest. That distance is then used to determine the *consolidated* distribution finish times around the reference point. For example, if a runner has an actual/implied finish time of 3:57:45 , the closest round number is reference point is 4:00, and they are classified as being -0:02 away from the round number goal (beat by up to 2 minutes) in the consolidated distribution. We then calculate the counterfactual distribution around this ‘consolidated’ round number using the same methodology described above.

Implied finish time

To calculate a runner’s implied finish time, we take their actual time at the 40km point of the race and multiply it by 1.055 (42.1949km/40km). For the bonking testing it is calculated based on the time at the 30km point of the race.

Optimal Race

We determine whether the runner ran an optimal race using the coefficient of variation (CV) of their paces calculated over the course of the race. This measure shows the relative variability of the runners pace over the course of the race. We calculate the CV as:

$$CV = \left(\frac{1}{N}\sum(Pace\ for\ Interval - Pace\ for\ Entire\ race)^2\right) X\ 100$$

If CV is less than 5% we classify the runner as having run an optimal race. We do not include the pace over the final 2.19 km in this calculation in order to remove the effect of the final kick.

Pace Ratio

To test pacing behavior we calculate a runner’s pace ratio for the interval under evaluation. This number is calculated as:

$$\frac{\text{Average pace per km for the race up to the evaluation interval}}{\text{Average Pace per Km for the first 10km}}$$

For example, if we want the pace ratio for the second to the second to last interval (30-40km) of the race we calculate it as:

$$\frac{\text{Average pace per km for the interval from 30-40 km of the race}}{\text{Average Pace per km for the first 10km}}$$

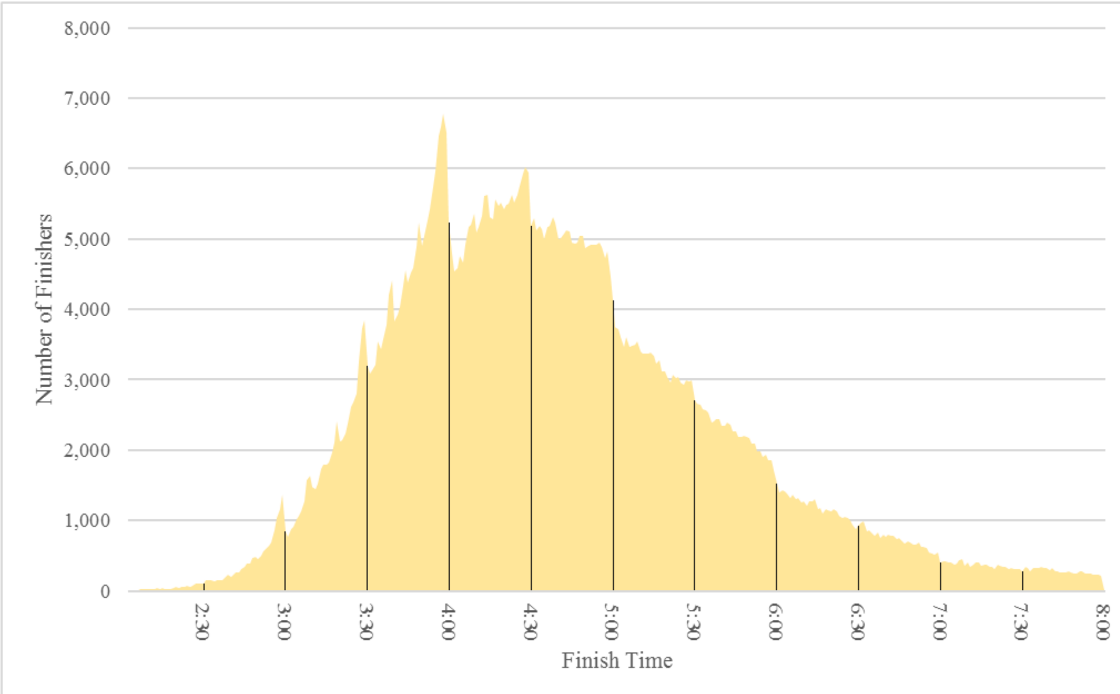
A pace ratio of 1 indicates runner maintained the same average pace for the evaluation interval
Pace ratio > 1 indicates the runner sped up over the evaluation interval
Pace ratio < 1 indicates the runner slowed down over the evaluation interval

'Speed up' over final interval

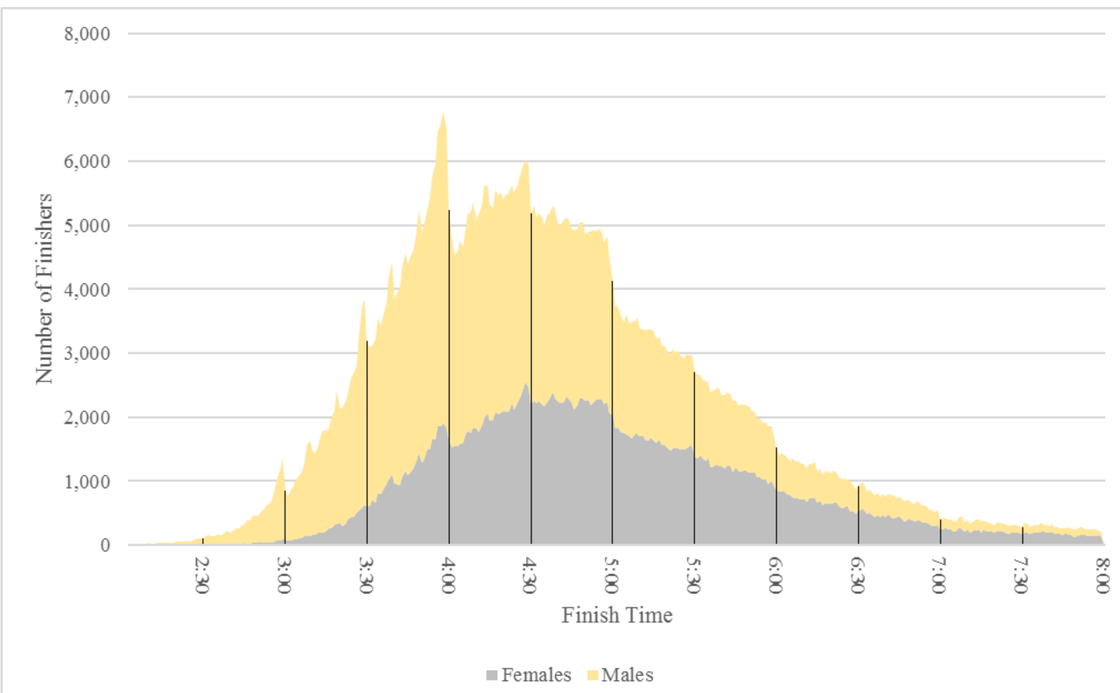
We compare a runner's pace over the final 2.19 km [(actual finish time – time to 40km)/2.19499] to their average pace over the previous 40km (time to 40km/40). If the pace over the final interval is faster they are classified as speeding up.

Figure 1: Distribution of Finish times for the full sample and for males and females.

Panel A: *Finish times: All Runners*



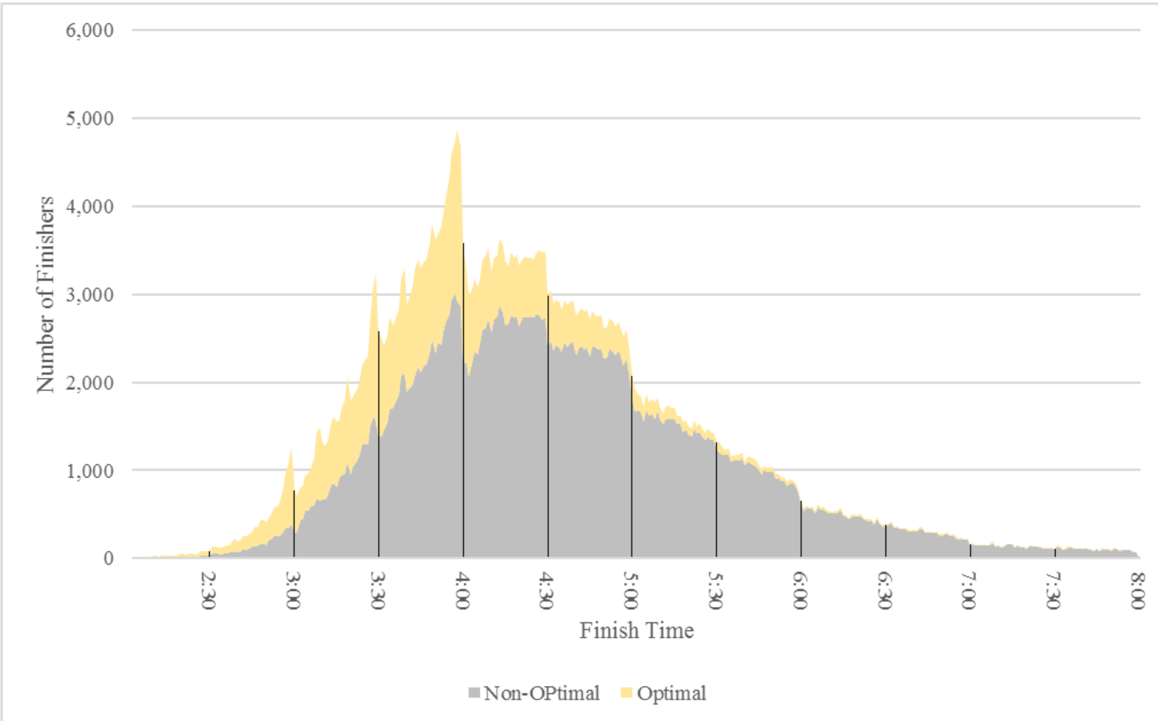
Panel B: *Finish Times partitioned on Gender*



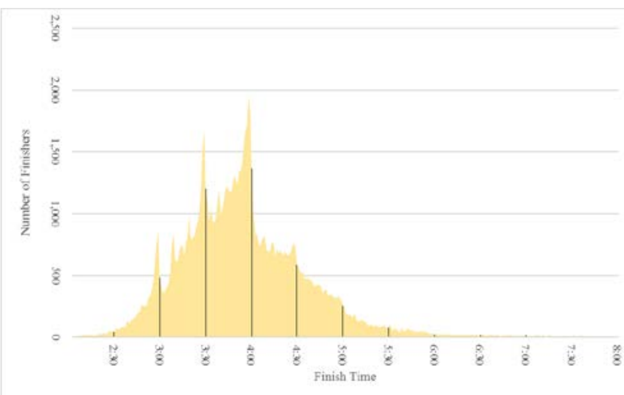
This figure plots the distribution of marathon finishers in 1 minute increments. Panel A shows all finishers. Panel B shows the breakdown of the total sample between male and female finishers. The **Black** bars represent every 30 minute interval. See Table1 for sample construction.

Figure 2: Distribution of finish times for Males and whether the runner ran an optimal race.

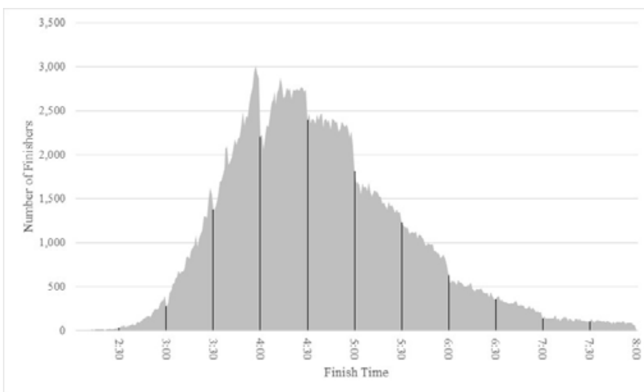
Panel A: *All males*



Panel B: *Optimal Males*



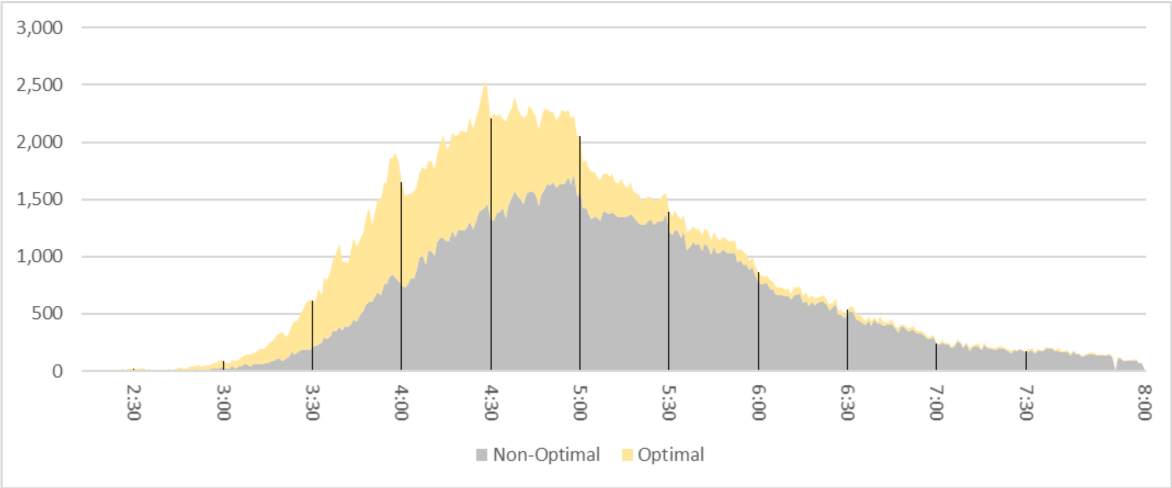
Panel C: *Non-Optimal Males*



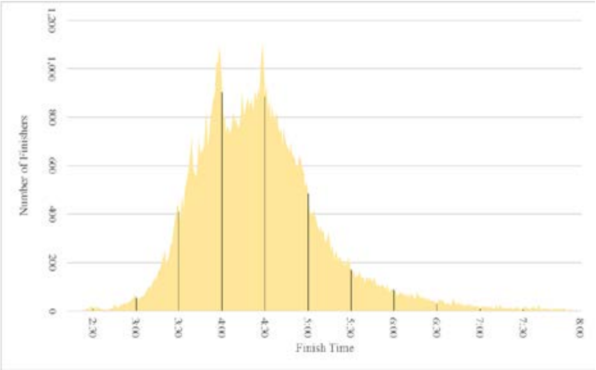
This figure presents the distribution of male finishers in one minute increments. The **Black** bars represent every 30 minute interval. Panel A shows the full sample partitioned on whether runner runs an optimal race. Panel B shows just the optimal finishers, and Panel C the non-Optimal. See the appendix for a definition of how we determine the runner ran an optimal race. See Table1 for sample construction.

Figure 3: Distribution of finish times for females and whether the runner ran an optimal race.

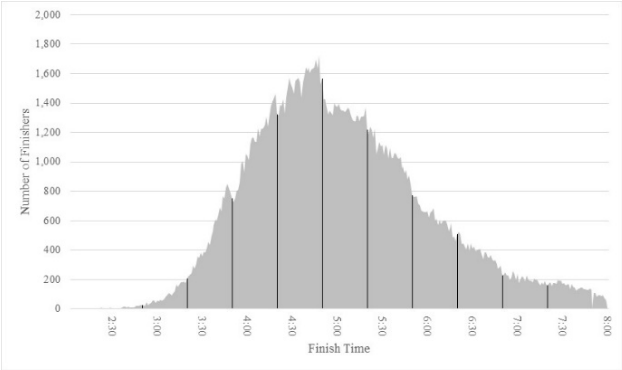
Panel A: *All females*



Panel B: *Optimal*



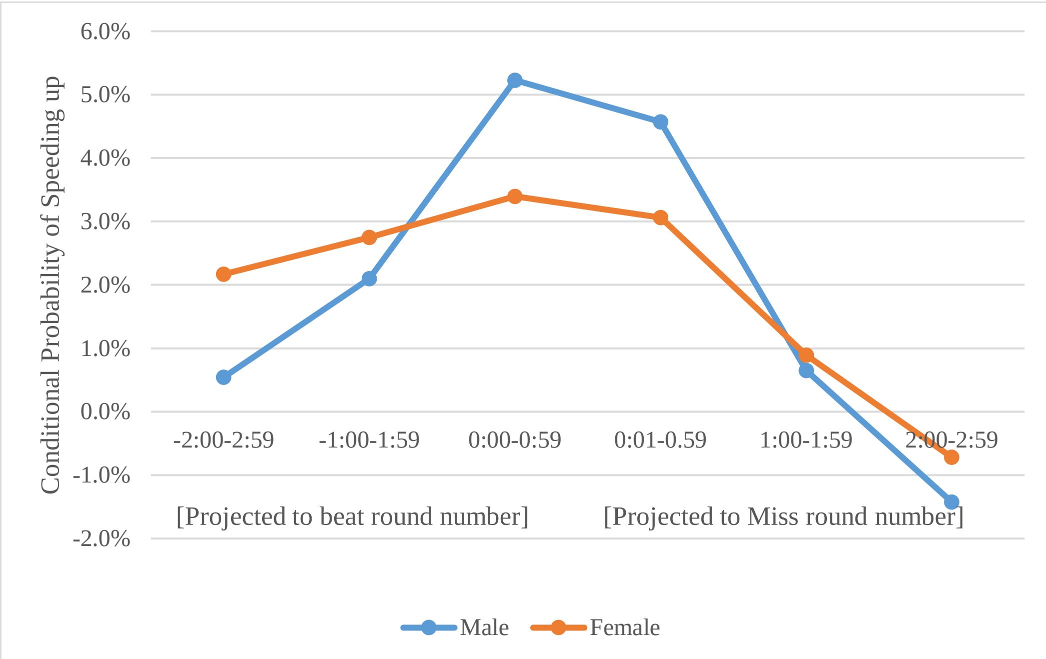
Panel C: *Non-Optimal*



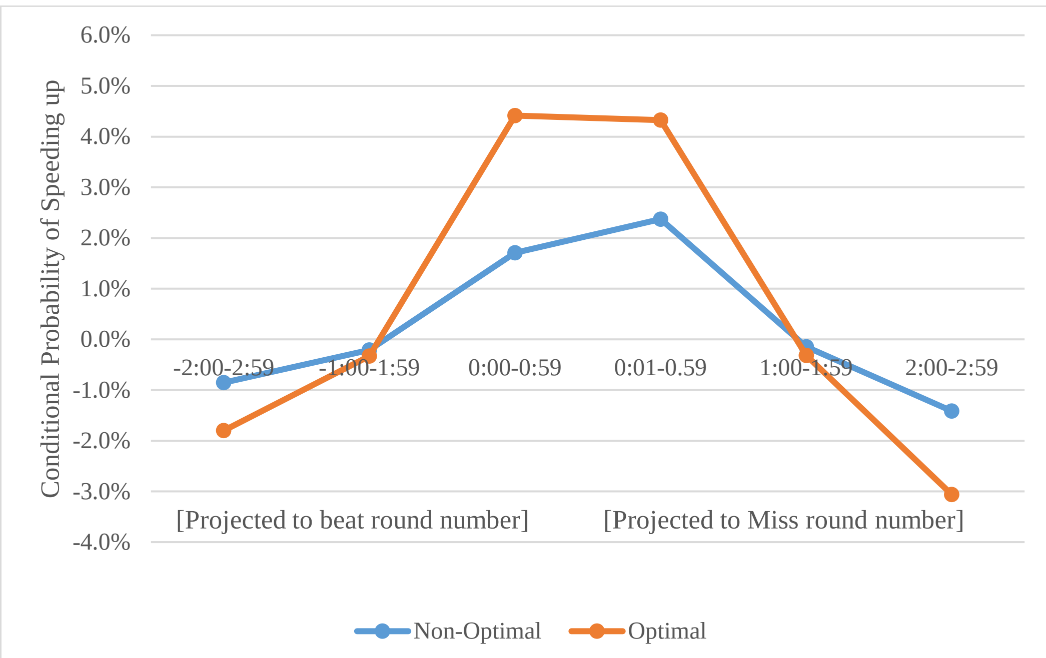
This figure presents the distribution of female finishers in one minute increments. The **Black** bars represent every 30 minute interval. Panel A shows the full sample partitioned on whether runner runs an optimal race. Panel B shows just the optimal finishers, and Panel C the non-Optimal. See the appendix for a definition of how we determine the runner ran an optimal race. See Table1 for sample construction.

Figure 4: End of race effort partitioned on gender and whether the runner ran an optimal race.

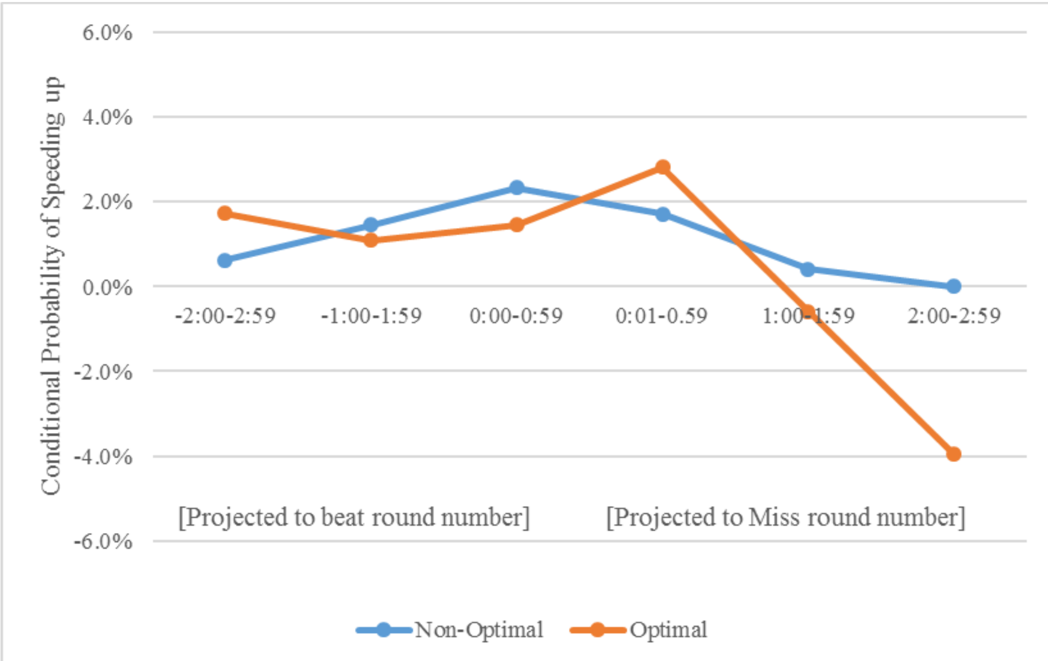
Panel A: *Total Sample partitioned on gender*



Panel B: ***Males:*** *Conditional probability of speeding up over the final interval*



Panel C: ***Females****: Conditional probability of speeding up over the final interval*



See Table1 for sample construction, and the appendix for variable definitions. For each runner the split time at the 40km is multiplied by 1.054875 in order to get the implied finish time. The closest round number finish time is then subtracted from the implied finish time to determine the 'distance' from that round number. Each runner is then placed into 1-minute bins based on their distance, with a distance of zero indicating an implied finish time exactly equal to the nearest round number, a negative distance indicating that they are projected to finish faster than the nearest round number, and a positive distance indicating they are projected to finish short of the nearest round number. The unconditional average for the each group is then subtracted from the percentage that sped up in each bin, to get the conditional probability of speeding up. So, for example, if 31% of men speed up in the 0:00-0:59 bin and the total percentage of men that speed up across the sample is 26%, then the conditional probability of speeding up for male runners in the 0:00-0:59 bin is 5%.

Table 1: List of marathons in sample

Marathon Name	Years	Total Finishers	Percentage Female (%)
Buenos Aires Marathon	1	5,024	19.29%
Chicago Marathon	8	255,333	43.03%
Eugene Marathon	1	1,658	44.93%
Frankfurt Marathon	2	17,930	15.96%
Hamburg Marathon	3	35,586	20.45%
Honolulu Marathon	6	109,588	46.08%
Houston Marathon	1	5,092	33.74%
Las Vegas Marathon	1	7,968	40.65%
Marine Corps Marathon	5	100,558	40.16%
New York City Marathon	6	209,041	35.28%
Nike Women's Marathon	2	8,792	91.53%
Poznan	1	5,633	12.07%
Toronto Waterfront Marathon	1	1,700	25.71%
U.S. Marathons	30	698,030	41.29%
International Marathons	8	65,873	18.56%
All Marathons	38	763,903	39.33%

We use the ‘full-split’ sample of marathon runners from Allen et al (2017) to perform our analysis. This sample that has complete 10-km, half-marathon, 30-km and 40 km split times, as well as the actual finish time. We eliminate observations when the gender of the runner is missing, the finish time is less than two or more than eight hours and the calculated pace-per-kilometer for any interval is faster than the current 10 kilometer world record pace (two minutes and thirty-six seconds per kilometer). This leaves us with a sample of 763,903 observations.

Table 2: Descriptive Statistics on Marathons

Variables	Total Sample	Subdivided by gender Male	Female	Difference*
Number of Finishers	763,903	463,433	300,470	162,963
Percentage Female	0.39			
Age**	38.64	40.11	36.29	3.82
Actual Finish Time	4:40:26	4:28:20	4:59:05	0:30:44
Pace per km for the entire race	0:06:39	0:06:22	0:07:05	0:00:44
Pace per mile for the entire race	0:10:42	0:10:14	0:11:25	0:01:10

See Table1 for sample construction.

*All between gender differences are significant at the 1% level.
Pace per km and pace per mile for the entire race are based on the Actual Finish Times.
******Age data is only available for 573,533 observations. 352,624 Male, and 220,909 female

Table 3: The bunching at round number finish times for males and females

Panel A: *Percentage of each gender that finishes in the bunching region*

	Male	Female	Diff
Percentage in bunching region	18%	14%	**4%**
N	84,449	42,719	

Panel B: ***Males:*** *Excess mass and magnitude of bunching at round numbers*

Time	Actual Finishers	Counterfactual Finishers	Excess Mass %
3:00	4,277	3,247	**31.7%**
3:10	5,330	4,888	**9.1%**
3:20	7,542	7,372	2.3%
3:30	11,980	10,554	**13.5%**
4:00	19,953	17,546	**13.7%**
4:30	14,807	13,914	**6.4%**
5:00	10,789	10,251	**5.3%**
5:30	6,141	6,030	1.8%
6:00	3,630	3,348	**8.4%**

Panel C: ***Females:*** *Excess mass and magnitude of bunching at round numbers*

Time	Actual Finishers	Counterfactual Finishers	Excess Mass %
3:00	307	267	**15.2%**
3:10	563	550	2.4%
3:20	1,223	1,119	**9.3%**
3:30	2,387	2,283	4.6%
4:00	7,933	7,215	**10.0%**
4:30	10,283	9,728	**5.7%**
5:00	9,352	9,187	1.8%
5:30	6,468	6,294	2.8%
6:00	4,203	4,021	**4.5%**

Panel D: *Excess mass and magnitude of bunching at round numbers:* ***Consolidated***

	Number of Finishers:		
Gender	**Actual**	**Counterfactual**	**Excess Mass % (Z-stat)**
Male	84,449	75,533	**11.8%**
			(16.29)
Female	42,719	40,365	**5.8%**
			(6.08)
Difference			**6.0%**
Z-Statistic			(4.92)

See Table 1 for sample construction, and the appendix for all variable definitions.

Panel A reports the actual number of finishers and percentage of each gender that appears within the bunching region of any round number time goal.

Panel B/(Panel C) reports the actual and counterfactual number of finishers, as well as the excess mass percentage for male/(female) runners in the bunching area around each round number time goal. It also tests whether the excess-mass percentage is significantly different from zero using a z-test.

Panel D reports the consolidated actual and counterfactual number of finishers, as well as the excess mass percentage for male and female runners. It also tests whether the between gender differences in consolidated excess mass are significantly different using a z-test.

For all panels **Bold** indicates percentage difference is significantly different from zero (for within gender) or the comparison group (for between gender) at the five percent level.

Table 4: The interaction of bunching at round number finish times and an optimal race

Panel A: *Percentage of runners who runs an optimal race (run at a constant pace)*

Males	**Total Sample**	**Optimal**	**Non-Optimal**	**Diff**
Percentage of Sample	100%	**24%**	76%	
Number of Finishers	463,433	110,694	352,739	
Avg. Coefficient of Variation	10.54%	3.15%	12.86%	9.71%
Avg. Finish Time	4:28:20	3:56:04	4:38:28	0:42:23
Females	**Total Sample**	**Optimal**	**Non-Optimal**	**Diff**
Percentage of Sample	100%	**30%**	70%	
Number of Finishers	300,470	89,658	210,812	
Avg. Coefficient of Variation	8.87%	3.28%	11.25%	7.97%
Avg. Finish Time	4:59:05	4:26:32	5:12:55	0:46:23

Panel B: *Percentage of runners in the bunching regions who runs an optimal race*

	Male		**Female**	
	N	**% of total**	**N**	**% of total**
Optimal	27,129	32%	15,307	36%
Non-Optimal	57,320	68%	27,412	64%
Total	84,449	100%	42,719	100%

Panel C: *An analysis of the consolidate bunching region (i.e. runners at round number finish times) and the excess mass explained by optimal and non-optimal runners*

	Number of Finishers:			
Group	**Actual**	**Counterfactual**	**Excess Mass % (Z-stat)**	**Difference (Z-Stat)**
Male: Optimal	27,129	23,373	**16.1%**	
			(12.89)	
Male: Non-Optimal	57,320	52,160	**9.9%**	**6.2%**
			(11.76)	(4.30)
Female: Optimal	15,307	14,385	**6.4%**	
			(4.06)	
Female: Non-Optimal	27,412	25,980	**5.5%**	0.9%
			(4.53)	(0.45)
Difference Between Males and Females				
Optimal			**9.7%**	
			(4.85)	
Non-Optimal			**4.4%**	
			(3.05)	

See Table1 for sample construction, and the appendix for variable definitions.

Panel A provides descriptive statistics for male and female runners based on whether they ran an optimal race. The difference in average coefficient of variation and average finish time for optimal and non-optimal runners is significantly different from zero for both genders.

Panel B reports the actual number of finishers and percentage of each gender that appears within the bunching region of any round number time goal partitioned on whether they ran an optimal race. The percentage difference is significantly different at the 1% level for both between, and within, gender comparisons.

Panel C reports the consolidated actual and counterfactual number of finishers, as well as the excess mass percentage for male and female runners portioned on whether they ran an optimal race. It also tests whether the within and between gender differences in consolidated excess mass are significantly different. **Bold** indicates the excess mass percentage is significantly different from zero at the five percent level for within group comparisons, or the comparison group (e.g. optimal male vs. non-optimal male) for between group comparisons.

Table 5: An analysis of the pacing ratio for optimal and non-optimal runners

Panel A: *Male Finishers*

	Total Sample	Optimal	Non-Optimal	Diff
Number of Observations	463,433	110,694	352,739	
Average Pace Ratio over each Interval				
The first 10km	1.00	1.00	1.00	-
10km - Halfway	0.99	1.01	0.98	0.02
Halfway - 30km	0.92	0.99	0.90	0.09
30km-40km	0.84	0.96	0.81	0.15
40km-End	0.89	0.98	0.86	0.12
Percentage of Runners that:				
Speed up over the 10km-Halfway interval	41%	55%	37%	19%
Speed up over the final interval	26%	43%	21%	21%

Panel B: *Females*

	Total Sample	Optimal	Non-Optimal	Diff
Number of Observations	300,470	89,658	210,812	
Average Pace Ratio over each Interval				
The first 10km	1.00	1.00	1.00	-
10km - Halfway	0.98	1.00	0.96	0.04
Halfway - 30km	0.93	0.98	0.91	0.07
30km-40km	0.87	0.95	0.83	0.12
40km-End	0.94	0.99	0.91	0.08
Percentage of Runners that:				
Speed up over the 10km-Halfway interval	35%	51%	29%	22%
Speed up over the final interval	40%	52%	35%	18%

See Table 1 for sample construction, and the appendix for all variable definitions. A pacing ratio of 1.00 indicates that the runner held a constant pace over the interval examined **relative to the first 10km interval**. For the first 10km all runners have a pace ratio of 1.00. A pace ratio less than 1.00 indicates the runner is slowing down and over 1.00 indicates the runner is speeding up. All within gender pace ratio differences, and percentage that speed up, are significantly different at the 1% level.

Table 6: An analysis of people that had a poor race outcome: "bonked" or "hit the wall"

Panel A: Males	**Total Sample**	**Bonkers**	**Non-Bonkers**	**Difference**
Percent of Sample	100%	**36%**	64%	
Average Finish Time	4:28:20	4:48:38	4:17:04	0:31:34
Implied Finish time based on pace to 30km	4:16:05	4:24:50	4:11:14	0:13:36
Change in anticipated time	0:12:15	0:23:48	0:05:51	0:17:58
Pace for first 30km	0:06:04	0:06:17	0:05:57	0:00:19
Pace from 30 to 40km	0:07:07	0:08:19	0:06:27	0:01:52
Percent Change in Pace from 30 to 40 km	17%	33%	8%	24%
Number of Finishers	463,433	165,445	297,988	

Panel B: Females	**Total Sample**	**Bonkers**	**Non-Bonkers**	**Difference**
Percent of Sample	100%	**21%**	79%	
Average Finish Time	4:59:05	5:24:11	4:52:29	0:31:42
Implied Finish time based on pace to 30km	4:49:30	5:00:10	4:46:42	0:13:28
Change in anticipated time	0:09:35	0:24:01	0:05:47	0:18:14
Pace for first 30km	0:06:52	0:07:07	0:06:48	0:00:19
Pace from 30 to 40km	0:07:43	0:09:13	0:07:19	0:01:53
Percent Change in Pace from 30 to 40 km	12%	30%	8%	22%
Number of Finishers	300,470	62,588	237,882	

See Table 1 for sample construction, and the appendix for variable definitions. A runner is classified as having "bonked" when he or she slows down by more than 20% in the third ten kilometer interval of the race (i.e., between kilometers 30 to 40). All within gender differences between bonkers and non-bonkers are significantly different at the 1% level.

Table 7: The likelihood of running an optimal race and bonking based on the runners' performance in the first 10 kilometers of the race.

Panel A: ***Male Finishers***

Group	Number of Finishers	Average 10k Split	Average Finish time	Percent that run an Optimal Race	Percent that Bonk
Elites	40,415	0:41:37	3:08:47	36%	26%
Runners	150,377	0:50:37	3:51:26	30%	35%
Joggers	268,017	1:04:40	4:58:14	19%	38%
Walkers	4,624	1:41:00	7:11:52	16%	11%

All between group differences (e.g. Elites vs. Joggers) are significantly different from zero at the 1% level using a two-tailed t-test.

Panel B: ***Female Finishers***

Group	Number of Finishers	Average 10k Split	Average Finish time	Percent that run an Optimal Race	Percent that Bonk
Elites	27,720	0:48:24	3:37:03	46%	15%
Runners	92,886	0:57:25	4:19:44	38%	21%
Joggers	175,941	1:12:17	5:29:36	23%	22%
Walkers	3,923	1:47:14	7:21:27	15%	4%

All between group differences (e.g. Elites vs. Joggers) are significantly different from zero at the 1% level using a two-tailed t-test.

Panel C: ***Difference between male and female finishers.***

Group	Percent that run an Optimal Race	Percent that Bonk
Elite	-10%*	11%*
Runners	-8%*	13%*
Joggers	-4%*	16%*
Walkers	1%	6%*

*Difference is significantly different at the 1% level based on a two-tailed t-test.

See Table 1 for sample construction, and the appendix for all variable definitions. We place runners into the following groups based their 10K split times. Because the average female is 7 minutes slower than the average male over the first 10K, we add 7 minutes to the male cut-off time for each group. To come up with the cut-off times for each group we form quintiles of each gender based on the 10k split times: 45 minutes represents the average 10k pace for the top quintile of male runners over than interval. 55 minutes represents the slowest 10k split for the second quintile; 93 Minutes represents the fastest walking pace over the first 10k – 15 minutes per mile or 9.3 minutes per Km.

Group	**Male Runner 10-k Time**	**Female Runner 10k-Time**
Elite	Less than or equal to 45 minutes	Less than or equal to 52 minutes
Runners	Between 45 and 55 minutes	Between 52 and 62 minutes
Joggers	Between 55 and 93 minutes	Between 62 and 100 minutes
Walkers	Above 93 minutes	Above 100 minutes