How Much of Merit Is Due to Luck? Evidence on the Butterfly Effect of Luck

Hideto Koizumi
RIETI
Hitotsubashi University*

December 31, 2024

PRELIMINARY – PLEASE DO NOT CIRCULATE AND CITE.
Comments welcomed.

Abstract

The rise of populism is said to be partly driven by people's anger toward elites and meritocratic societies, while progressive societies aspire to eliminate discrimination and promote equal opportunities and meritocracy. The crucial question remains: does the pursuit of equal opportunities and merit-based evaluation truly lead to a fair society? To test the presumed fairness of meritocracy, this paper quantitatively analyzes the impact of luck in the absence of initial differences in individual characteristics. This study utilizes a unique experimental setting involving Japanese speedboat races. Participants are randomly assigned engines with different capacities in each tournament, ensuring probabilistic fairness across racers but introducing variability in the timing of luck. By identifying racers assigned "lucky" motors during their debut periods as the treatment group, I trace their performance trajectory, examining factors such as the number of first-place finishes and earnings in the short and long run. The findings unveil positive feedback loop effects of initial success induced by luck. Over ten years, the initially modest advantage of the treatment group results in a remarkable 41% lower market exit rates. Despite the presence of downward bias through surviorvalship, I find a widening gap in the number of winning and earnings. These results underscore the pivotal role of early-stage luck in shaping substantial differences in merit, challenging the presumed fairness of meritocracy.

Keywords: feedback loop, butterfly effect, speedboat race, meritocracy, inequality

JEL Codes: D63, D83

*Email addresses: koizumi-hideto@rieti.go.jp. This study is conducted as a part of research at the Research Institute of Economy, Trade and Industry (RIETI). The draft of this paper was presented at the DP seminar of the Research Institute of Economy, Trade and Industry (RIETI). I would like to thank participants of the RIETI DP Seminar for their helpful comments.

1 Introduction

Progressive societies strive to eradicate discrimination and champion equal opportunities and meritocracy. The crucial question remains: does the pursuit of equal opportunities and merit-based evaluation system lead to a fair society? Since an argument that our fate is determined by a pure chance is considered as one of the strongest counter-arguments against meritocracy, to test the presumed fairness of meritocracy, this paper quantitatively analyzes the impact of early-stage luck in the absence of initial differences in individual characteristics. Illustratively, even minor initial disparities can burgeon into significant differences over time, driven by a positive feedback loop—a phenomenon commonly known as the Butterfly Effect. This dynamic is anticipated across various economic spheres, encompassing academic accomplishments among students and the career trajectories of entrepreneurs and CEOs. Enhanced performance not only attracts additional opportunities but also nurtures heightened confidence, fostering greater motivation and diligence, consequently leading to increased success. If these initial disparities arise from chance, it challenges the presumed fairness of meritocracy. To address the question, this research endeavors to conduct a quantitative analysis assessing the extent to which merit is influenced by early-stage luck when individuals, beginning with identical characteristics, including family backgrounds and abilities, engage in competition within a purely meritocratic evaluation framework.

Empirically addressing this question is, however, challenging for two reasons. First, finding a society that operates in the specific way described is difficult. Furthermore, even if such a society exists, estimating the dynamic nature of the Butterfly Effect poses a unique challenge in the social sciences due to the limited availability of long-term, high-frequency performance data in a natural experimental setting. To overcome these challenges, this research leverages a distinctive experimental context involving Japanese speedboat races. In this setting, engines with different capacities are randomly assigned to participants in each tournament. This random allocation ensures fairness among racers on a probabilistic basis but introduces a temporal variation in the actual occurrence of luck. By identifying racers assigned "lucky" engines during their debut periods as the treatment group, I trace their performance trajectory, examining factors such as the number of first-place finishes and earnings.

The findings of this paper reveal a widening performance gap between the treatment and control groups, accompanied by surprisingly substantial Butterfly Effects. Within the first ten years since debut, the modest initial shock fueled by early-stage luck accumulates for racers, resulting in 41% lower market exit rates. Despite the presence of downward bias through surviorvalship, I find a widening gap in the number of winning and earnings.

Furthermore, by examining the risk-taking behavior immediately after the lucky event, I find out the successful experience induced by luck caused racers to take risks almost 40% *less*, while this tendency

immediately vanishes and rather flips to *greater* risk-taking behavior *over time*. Despite the decreased risk-taking behavior immediately after the lucky event, individuals with low initial ability still experiences greater probability of winning immediately after the lucky event, providing suggestive evidence for confidence boost by the success induced by the lucky event. This tendency persists even in the long run.

To derive these results, I construct a dataset of Japanese speedboat racing, consisting of 6,952,990 race-by-racer observations involving 2,717 racers spanning the period from 1996 to 2022. Among these racers, I focus on racers debuted after 1996 for whom I obtained comprehensive, confidential administrative data from the Foundation of Japan Motor Boat Racing Association (FJMBRA) on demographics and pre-debut performance, enabling a balance test. The primary analysis focuses on the period from 1997 to 2011, as this timeframe predates significant institutional changes in engine regulations occurring after 2012 and 2014. The resulting sample comprises 672 racers who debuted during this specified period.

In Japanese speedboat racing, racers are randomly assigned engines with varying capacities at each tournament, including those deemed as "lucky." The engines play a crucial role in determining race outcomes, influencing the odds of races. Information about the engines is made public before races and is included in race programs. Note that these engines are replaced annually at each speedboat racing stadium.

To identify "lucky" engines, I first compute the winning percentage for finishing within the top two places for every engine at each stadium.[1] Then, for each new racer in the debut term, I tally the instances of being assigned "lucky" engines during the debut term, which is the first half-year period since their debut. To designate roughly half of the racers into the treatment group, a cut-off of the 90th percentile for engines is chosen to qualify as "lucky."[2]

With the treatment, I perform simple regression analyses to examine short-run bahavioral change and long-run dynamics. For the short-run analysis, restricting the sample to the debut term, I examine if being assigned lucky engine at this tournament changes behavior and race performance at the next tournament and the tournament following the next. For the long-run analysis, to alleviate identification concerns, I conduct a balance test of racer demographics and pre-debut performance. I ensure the balancedness between the treatment and control groups.

While the identification threat is mitigated by the balancedness between the treatment and control group, note that the difference between the treatment and control group does not correspond to the effects of the treatment on the treatment group. Since increasing performance of one racer entails decreased relative performance of another, the averaged outcome of the control group does not corre-

[1] Winning percentages for finishing within the top two places are used by the organizer to notify the performance of engines for racers and bettors.

[2] The main results are mostly unchanged with different cut-offs of the 85th percentile and the 95th percentile.

spond to the counterfactual in the absence of the treatment. Therefore, the difference captures the sum of the treatment and competition effects. In this sense, the estimates correspond to the effects of luck in the small, local labor market such as a school or a company.

The implications drawn from the results are threefold. First, the immediate policy implication underscores the importance of early-stage intervention. These findings contribute to existing scholarship by elucidating the dynamics of how initial differences in the early stages significantly impact later stages, whether applied to individuals or businesses. Secondly, the results suggest that whether it pertains to personal income tax or corporate tax, a morally justified stance supports a more progressive tax system. Notably, the pronounced Butterfly Effects of early-stage luck are identified even when the initial characteristics are *identical* between the treatment and control groups. This undermines the basis for meritocratic arguments against progressive tax systems, as a substantial portion of merit appears to be attributed not to innate ability or *voluntary* efforts but rather to sheer luck. Early-stage luck results in confidence boost, subsequently leading to increased efforts, leading to even greater confidence, and ultimately elevated ability over time.

Last but not least implication is that equal opportunities with a meritocratic evaluation system will not lead to a fair society with justifiable inequality. The results of this paper unequivocally demonstrate the significance of the role of early-stage luck. It is important to note that the luck studied in this paper is posterior, occurring after birth, rather than the prior luck determined by genes and family backgrounds. This implies that eliminating disparities caused by prior luck does not necessarily eradicate discrepancies induced by everyday luck. Therefore, it becomes challenging to claim that we *deserve* what we are now, as what we establish can be largely influenced by chance. Michael Sandel argues that the prevalent meritocratic claim in contemporary societies significantly diminishes the self-esteem of many non-elites, fueling their resentment towards elites and contributing to recent populist movements worldwide (Sandel (2020)). The findings of this study challenge the validity of the meritocratic claim, providing some empirical support for the *redistribution of esteem* as Sandel (2020) proposes.

This paper contributes to multiple strands of the literature. It adds to the established literature on the enduring effects of early differences in abilities and life conditions. Within the literature that delineates the so-called "relative age" effects of early childhood maturity differences induced by the school start dates and variations in birthdays, closely related studies demonstrate the lasting impact on various outcomes including university enrollment (Bedard and Dhuey (2006)), labor earnings (Kawaguchi (2011)), top management (Du et al. (2012)), and political positions (Muller and Page (2016)).[3] Notice the presence of an ability gap between early-born and late-born children within a class in this context. Another strand of the same literature proves the importance of early life conditions including genes in shaping multiple life skills that culminate in sizable differences in socioeconomic outcomes later in

[3] Yamaguchi et al. (2023) provide empirical evidence on non-cognitive skill formation differences induced by the relative age effects, potentially serving as a mechanism behind some of these long-term effects.

life (see, e.g., Cunha et al. (2005); Heckman and Mosso (2014); Bailey et al. (2021); and Houmark et al. (2024)). This paper complements the literature by (i) providing evidence on the enduring effects of early-stage luck even if ability and other life conditions such as family backgrounds are the same to start with, (ii) demonstrating that adults also experience persistent and escalating effects of small advantages early in their careers, and (iii) portray *how* the gap evolves over time in a highly controlled environment that excludes factors like family backgrounds, neighborhood effects, and nepotism that interact with the evolution of childhood performance in the literature.

Finally, this literature contributes to a strand of papers that find the long-term effects of external factors on performance in general. Almond (2006) investigates the long-term effects of the Spanish flu on economic and socioeconomic outcomes, while Maccini and Yang (2009) examines the impact of childhood rainfall on later-life socioeconomic outcomes. Additionally, Genda et al. (2010) explores the effects of economic downturns on subsequent outcomes in labor markets. In contrast, my paper contributes unique evidence on the dynamic and cumulative effects of early-stage luck on actual *ability*, a dimension not measured by these studies.

Ginsburgh and Van Ours (2003) exploits a setting that is more akin to this paper than these papers. They show the effects of the random ordering of finalists at a world-renowned musical competition on the later-stage market success of the participants. They measure success by the presence of the historical participants' names on the catalog of a Belgian library together with British and French catalogs of classic music players. It is important to note that they were interested in the effects of expert opinions on market recognition and thus market price, unlike this paper that investigates the effects of luck on merit including actual *ability* and earnings. This paper adds to the literature by providing (i) the intermediate dynamics, (ii) analysis on the population of new competitors, and (iii) direct measures for merit. The high-frequency nature of my dataset allows me to delve into the dynamics of luck. Furthermore, unlike the finalists who are at the very top of young pianists, my setting covers the population of the competitors in a sport, allowing for an estimation of distributional effects. Lastly, although their catalog indicator is a great measure for success, it may not perfectly capture ability since, as the focus of their paper, expert opinions may affect the market prices of the musicians through the demand side. They also write, "[e]arnings would, of course, have been a much better choice, but are impossible to collect" (p. 290). The number of first-place finishes and earnings in a sport allows me to directly measure merit, avoiding the influence of expert opinion.

2 Speedboat Racing in Japan

For the description of speedboat racing in Japan, I reference Booth and Yamamura (2018) for certain aspects. Note that for the purpose of this study, I augment this information with greater details on the

engine allocation process, engine structures, and historical changes in institutional regulations. Additionally, I provide fewer details on gender differences, as my study addresses a distinct question than theirs. For the additional details, I rely on Himura (2015), the principal source in Booth and Yamamura (2018), as well as sources such as Fujino (2006), Himura (2017), the official website of the Japanese Speedboat Racing Association (JSRA), and various online articles, including those authored by Himura.

Speedboat racing in Japan is organized in the form of tightly regulated tournaments overseen by the Japanese Speedboat Racing Association. Both male and female racers undergo identical intensive training at the single training institution, the Yamato Boat School.[4] Individuals aged between 15 and 29 aspiring to become professional speedboat racers are required to undergo a year of training at the Yamato Boat School and pass a final examination. There are two timings in a year when they graduate from this school and become professional racers: May 1 and November 1. The wide age range allows individuals from diverse backgrounds to enter the profession, ranging from those with junior high school education to university graduates, some of whom have pursued other careers after their education. Therefore, the time elapsed since graduating from the boat school varies among entrants.

According to Himura (2015), there are approximately 1,600 speedboat racers in Japan, spanning ages 18 to 70, with around 1,400 men and 200 women. While the number of female racers has been increasing due to increasing demand for female racers, in my sample period, roughly speaking, 25 % of tournaments are mixed-sex, 5% are female-only, and 70% are male-only. Racers are randomly assigned to mixed-sex or single-sex tournaments. The country hosts 24 speedboat racing stadiums, with races taking place about four days a week in each stadium. There are two terms for evaluation in a year: the first term ranges from May 1 to October 31, while the second term ranges from November 1 to April 30.[5] The racing circuit consists of a 600-meter artificial pond or sectioned-off water body, around which competitors race three times, totaling a distance of 1,800 meters.

As depicted in Figure 1, speedboat racing employs the premature start system, requiring boats to cross the starting line within one second after the starting clock hits zero. Initially, racers are assigned their pits by the association committee. Nevertheless, racers have the strategic option to alter their starting lanes during the pit-out period as described in Figure 1, potentially resulting in a different starting position for the race. Being positioned directly behind another boat is considered a violation. In general, for safety reasons, new racers are assigned the most outward pits (6th lane). This makes it hard for new racers to win a race unless they are assigned a good engine.

[4]Women not only train alongside men in similar conditions but also engage and compete in races under the same circumstances.

[5]The JSRA applies the results of evaluation from the first term to January 1 till June 30 next year, and those of the second term to July 1 till December 31.

2.1 Tournament Grade, Racer's Grade, Prizes, and Penalties

Tournaments are categorized into either graded or non-graded (so-called "usual races") tournaments. Graded tournaments are further classified into four sub-classes, ranging from the lowest rank grade III (GIII) to grade II (GII), grade I (GI), and Special Grade (SG). In general, non-graded, "usual race" tournaments are bottom-graded and SG tournaments are the top-graded. Each tournament is held at a stadium over three to seven days (usually four to six days) with a racing meet on each day. Each racing meet comprises twelve races, each with six racers competing. These races progress through various stages, from preliminaries and semi-finals to the finals.

A racer earns points based on their finishing positions in races.In non-graded and GIII tournaments, points are awarded as follows: 10 for first place, 8 for second, 6 for third, 4 for fourth, 2 for fifth, and 1 for sixth. GI and GII tournaments add 1 point to these values(Himura (2015)), while SG tournaments add 3 points prior to the first term of 2005 and 2 points thereafter.[6] Penalties, such as those for poor navigation or rule violations, can result in a loss of 7 points.

Not only tournaments are graded, but also racers. There are four racer grades: A1 (the top racer grade, around 20% of all racers), A2 (around 20% of all racers), B1 (around 50% of all racers), and B2 (the bottom racer grade, around 10% of all racers). Individual racers' grades change twice a year (the first period is from January 1 to June 30, while the second period is from July 1 to December 31) and the change is based on the number of points and races in the most recent term.[7] For example, for a racer grade from January 1 to June 30, 2023, the performance during the first term of 2022 (May 1 to October 31, 2022) is used.

The prizes offered are substantial but vary significantly across tournaments and the stages of tournaments explained in the Appendix. The amount of prizes on average is ordered from SG to G3 tournaments in a descending manner. In general, graded tournaments have higher average prizes than non-graded tournaments. The amounts of SG tournament prizes range from around 18 million to 110 million yen depending on tournaments, while those of non-graded tournaments tend to be slightly more than 0.74 million yen. Notice that the discrepancies of prizes between tournaments are much larger than those of the aforementioned points. Therefore, points are far from precise measures of earnings.

2.2 Engines

On a day before a tournament, boats and engines are randomly allocated to racers participating in the tournament.[8] The same allocated boats and engines are used throughout the tournament. For random

[6] From https://sp.macour.jp/columns/himura/154720/ accessed on November 2, 2023.

[7] Booth and Yamamura (2018) write that these grades are determined based on points aggregated for three years, but it must have been a typography for a half year. Note that each racer's license renewal is conducted ever three years.

[8] See an actual lottery for motor and boat allocation at https://www.youtube.com/watch?v=r1OTsnCeJGc accessed on November 2, 2023.

allocation, lotteries are publicly conducted. In general, each stadium provides two separate lottery wheels for boats and engines (one for boats, and the other for engines). Each racer draws a ball from each lottery wheel in front of the other racers. Racers are not permitted to refuse the boat or engine randomly assigned to them. These boats and engines are made by the same company for all stadiums and are annually renewed, but the renewal month of a year varies across stadiums. As Himura (2017) explains, boats do not affect the results of races as much as engines do. He recommends not to care about the past performance of boats such as the winning percentage for finishing within the top two places unless it is terribly poor, while he emphasizes the significance of engines.[9]

This is because of a reason caused by an assembly of many parts similar to horsepower differences among automobiles of the same brand due to individual differences among the capacities of the same engines. Speedboat engines (motors) are made of around 400 different parts. While the production process is standardized, some margins of errors are permitted. When rotational speeds of motors attain high that causes heat, these small errors get amplified, due to so-called engine "bloating." This will lead to slower piston moves in the engine and thus shaft. According to Himura (2022), there are significant capacity differences among new engines even in the very first race of these engines. Racing programs provide data on the performance of each engine in the past races such as the winning percentage for finishing within the top two places, and bettors use this information.

Racers are allowed to fine-tune engines, which requires some experiences. While there are around 400 parts in an engine, a racer is allowed to replace only nine kinds of them: piston, ring, carburetor, gear box, electrical system parts, crankshaft, cylinder box, screw propeller, and carrier body. To replace these parts, they have to consult with a mechanic and obtain his or her permission. The replacement costs are free in general.

Before April 1, 2012, racers were allowed to bring their own propellers. This system was abolished mainly because this system made the betting hard. After 2012 April 1, racers have to use propellers that accompany assigned engines. The rule change changed skill sets needed to win a race.

After the change, they are allowed to tune up the propellers provided by the association using wooden hammers to adjust the propellers. These tuning skills take some experience and make it difficult for new racers to adjust them for their racing styles.

Overall, this system change makes it hard to measure the capacity of an engine at one point in time since what is "good" depends on whether or not the previous user happened to have adjusted the propellers in a way that it matches the racer's style, especially for racers with less experience. Section 4.2 empirically prove this. Since I cannot tell if the adjustment done by the previous user matches a particular new racer's racing style, I focus on those racers who debuted before October 31, 2011.

Note that the JSRA introduced a new model of engine (Yamato 331) in December 2014. The new

[9]Bettors usually ignore boat performance but carefully check engine performance. See https://wsobv.com/boatrace/7567 and https://kyoutei-navi.com/beginner/motor/ accessed on November 2, 2023.

model has much less horsepower,[10] and this makes it even harder for new racers to win a race since they are assigned the most outward pit and lane. I empirically show this in section 4.2. By the end of 2015, all stadiums replaced entire engines with the new model. Additionally, the new model introduces a larger variance in capacity than the old model.[11]

Right after random draws of boats and engines, racers are provided with racing programs for the most recent past tournaments that have information on the winning percentages for finishing within the top two places of engines and which racers were assigned these engines. Given the information, racers determine strategies during the actual races.

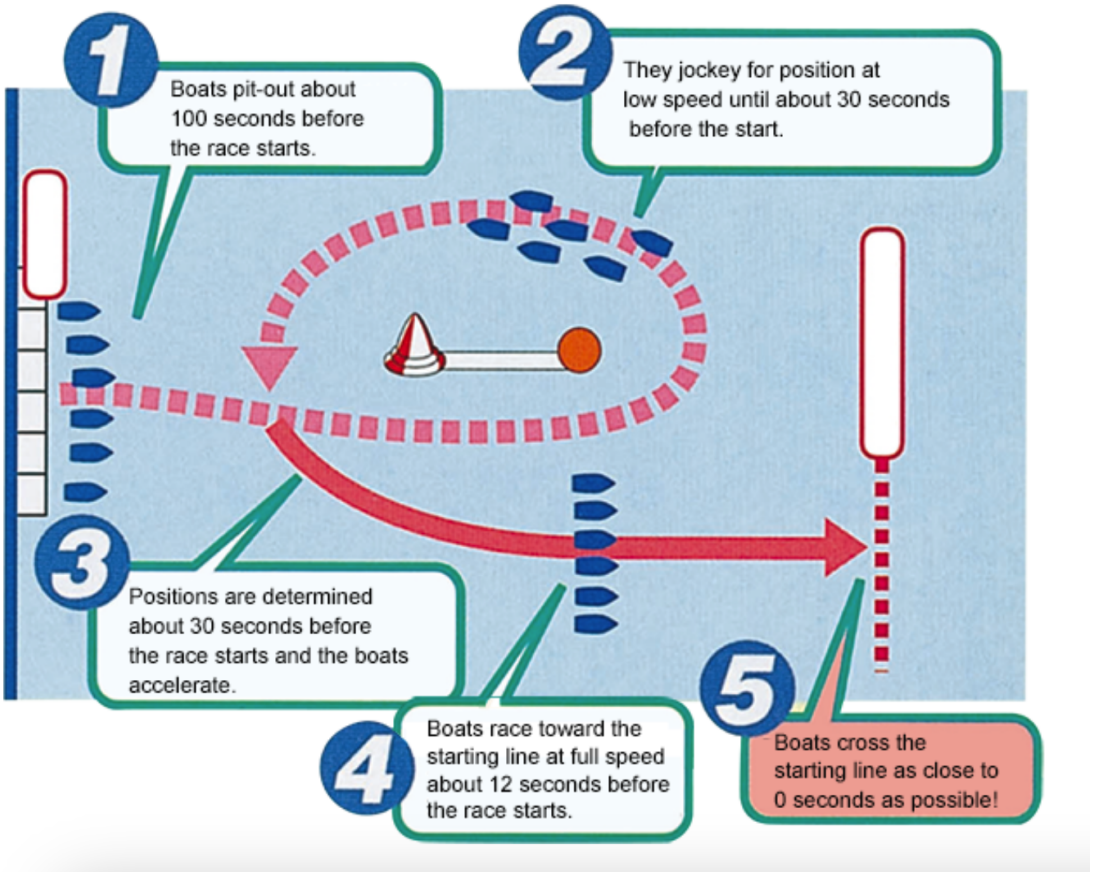
2.3 Strategies: lane changes

As explained above, while there are initial lanes assigned to racers, racers are allowed to change lanes during the premature start. However, while lane changing can bring benefits, it can also bring costs for two reasons. One comes from the rule that prohibits racers to stop their engines and break before starts. Lane changes toward more inward lanes require racers to accelerate more than the other racers, and this makes it more likely for racers to position too close to the starting line. This results in insufficient approach distances, which leads to insufficient acceleration before starts. The second reason comes from rule violation risks. The race rules are strict, and breaking them leads to serious penalties. Therefore, lane changing can be considered a risk-taking behavior and thus is used as a measurement of risk-taking as used by Booth and Yamamura (2018). The details of the penalties are summarized in the next section.

[10] See an article about this at the JSRA website, https://www.boatrace.jp/owpc/pc/extra/race/24_20150526/monthly/himuken.html.

[11] See the same article, https://www.boatrace.jp/owpc/pc/extra/race/24_20150526/monthly/himuken.html.

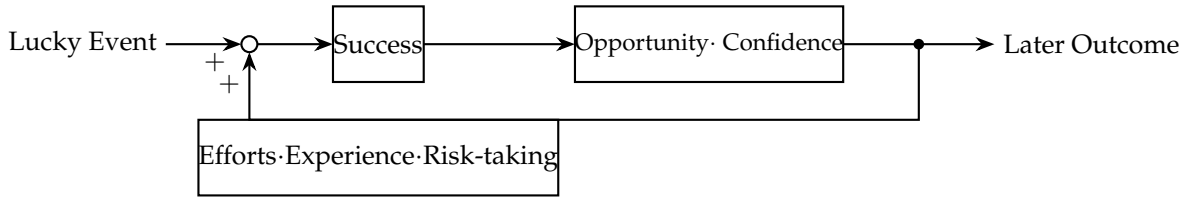
Figure 1: The Premature Start System



Source: Japan Boat Race Association, https://www.boatrace.jp/owpc/pc/extra/en/index.html(accessed November 2, 2023).

3 Econometric Framework

Figure 2: Graphical Illustration of Positive Feedback Loop Effects of Luck



Notes: the structure of the graph follows the standard positive feedback loop. It starts with an lucky event that positively affects debutants. The lucky event leads to a successful outcome—i.e., winning higher placements in a race—-, which is expected to lead to heightened confidence and more opportunities. Greater confidence is expected to induce racers to take more risks, while more opportunities provide them more experiences and potentially induce more efforts.

First, I introduce an econometric framework for each component of the feedback loop after the initial lucky event, illustrated in Figure 2. In particular, restricting the sample to the debut terms of racers who debuted between November 1, 1997 and October 31, 2011, I first analyze the short-term effects of being assigned lucky engines. In particular, I examine how the random assignment of a lucky engine at the current tournament affects the behavior and performance of racers at the next tournament. Second, to trace the trajectory of racer performance, using the debut term as the reference period, I trace the differences between the treatment and control groups in the first 10 years of racers' careers.

Note that in both short-term and long-term analysis, despite random allocation of lucky engines, the "control" group does not correspond to the counterfactual of the "treatment" group in the standard sense, due to competition. In other words, since racers compete against each other, an increase in the performance of lucky racers generally entails a decrease in the performance of non-lucky racers. Furthermore, better performance of rivals may psychologically affect racers, and thus the mean of the outcome from the control group in any analysis does not correspond to the counterfactual mean of the treatment group in the absence of treatment. Rather, the difference between the two groups indicates the relative performance or behavioral change between non-lucky racers and their lucky rivals.

Nevertheless, the difference between the two groups is exactly what this paper tries to estimate. In a real world, any market involves competition, and someone's performance usually affects another in a relative sense if one looks at a market close enough. For example, a higher level of competition from your classmates affects your ranking in terms of scholarship. Since this paper is interested in how the *gap* between lucky and non-lucky agents evolve over time, the difference precisely estimates what this paper wants.

3.1 Short-term Effects of Luck: Risk and Confidence Channels

First, I examine if lucky racers take risks at the next tournament or experiences greater success through confidence boost, restricting the sample to the debut term (i.e., the first six months since debut). Below,

I avoid the use of excessively many subscripts as much as possible.

$$y_{irnp} = \alpha_i + \eta_{m(i,n)} + \kappa_p + \xi_{s(n)} + LuckyMotor_{in\tau} + EngineCapacity_{in} + X_r + \mu_{irnp}, \tag{1}$$

where y_{irnpc} is an outcome of racer i with assigned pit p (initial lane) at race r of tournament n (unique tournament ID) that is the mth tournament—m can be uniquely determined by i and n—held at stadium s—each tournament is uniquely held at stadium s—since the debut tournament of i, $LuckyMotor_{in\tau}$ equals 1 if racer i was assigned a lucky engine at tournament n that is the τth tournament of racer i for $\tau \in \{m-1, m-2\}$, $EngineCapacity_{in}$ is the winning percentage of the engine (within the first two places) assigned to racer i at the mth tournament, and all the known race-level environmental factors, X_r, that affect race outcomes including the height of waves, wind velocity with wind direction, weather, and race distance. Fixed effects are used for wind velocity with wind direction, weather, and race distance. Note that 99.9% of tournaments are non-graded. Thus, we do not need to control for the kind of tournament. While the awards of these non-graded races may vary slightly depending on the sales of stadiums that hold the tournament, this fluctuation is quite small. Then, the fixed effects for tournaments based on unique tournament identifiers are not included because all the race-level factors that potentially affect race outcomes are included as controls. Furthermore, since in general only one debutant attends each tournament, taking out fixed effects for unique tournament identifiers is equivalent to taking out the average outcome values of a racer within a tournament, which largely includes the treatment effect. Therefore, I do not take out fixed effects for tournaments.

As for y_{irnp}, I examine three outcomes. One is a binary variable that equals 1 if racer i finishes within the first two places and 0 otherwise. Rather than the first place, I use the placement within the first two since few debutants win the first place, which leaves insufficient variation to detect the effects. The second one is a binary variable that equals 1 if racer i takes risk and change lanes inward after pits-out and 0 otherwise. For this outcome, I exclude those racer-by-race observations in which racer i is assigned pit 1, the most inward lane. The last one is the difference between the pit (initial lane) assigned at the first race of the $m(i,n)$th tournament and the pit assigned at the first race of the $m(i,n)-1$th tournament for racer i. In general, more inward pits are assigned to higher graded racers and better performing racers. Pits at semi-finals and finals are assigned according to the race outcomes of preliminaries within the same tournament. Usually, debutants are assigned the most outward pit in preliminaries during the debut term for safety reasons. Since the association may assign pits based on the outcomes of races after the first one of preliminaries and since some racers might be absent in these races due to accidents or penalties, I use the first race that racer i attends at the $m(i,n)$th tournament. This outcome becomes positive if racer i is assigned a better pit at the $m(i,n)$th tournament compared to the previous tournament, which captures a favorable treatment given by the association based on

the previous tournament outcome, and negative otherwise. Following Bertrand et al. (2004), standard errors are clustered at the racer-by-tournament level since randomization is conducted at the level and the serial correlation is anticipated within racer-by-tournament observations.

3.2 Long-term

For the long-term analysis, I use the time frame of 10 years after the debut term given the last cohort debuted on May 1, 2011. I analyze both extensive and intensive margin measures. For the intensive-margin anslysis, the expected surviviorship bias through market exits will increase the mean of the control group and thus result in a downward bias. Then, the estimates for the intensive margin measures are a lower bound of the true effects. Nevertheless, I believe that the lower bound is still meaningful to estimate. Since there is no pre-treatment period, we do not have a valid reference period that has no treatment effect. Then, for each (half-year race-season) term $t' = \{0, ..., 19\}$, I estimate the following equation:

$$y_i = \gamma_0 + \gamma_1 D_i + \xi_i, \tag{2}$$

where y_i includes termwise performance measures including the (non-cumulative) number of first-place finishes, earnings, the rate of inner-lane cut-ins, and the number of races under GII or above tournaments. The last measure is included to gauge if racer i is aiming to the top. This is because some racers strategically decline race offers to top-graded tournaments and instead attend lower-graded tournaments with less competition. Note that racers are eligible for the top-graded tournaments only when their racer grades reach A2, which takes place some time after their debut. Therefore, for this measure in particular, I estimate the above equation for $t' \in \{5, ..., 19\}$. Notice that taking out fixed effects for cohorts will result in bias since while racers of the same cohort do not compete in the same tournament during the debut term, they do compete against each other especially after the first two race-season terms. Therefore, I do not take out cohort fixed effects in this regression.

4 Data, Treatment Definition, Stylized Facts

The main dataset comes from publicly available data provided by the JSRA. The institution provides all the race results of all stadiums that took place after July 17, 1996. The race results include the race dates, stadium names, tournament names, race distance, weather, wind strength, wave height, placement (arrival order and any rule violation), initial lane position, racer names, racer ID, race categories (e.g., semi-finals), engine ID, boat ID, exhibition time records for test runs, lanes after lane changing, start time, and race time.

The FJMBRA generously provided the author the confidential demographics of racers and engine

replacement dates. The demographics include pre-debut performance measures (winning rate at races at the training school), birth dates, sex, height, weight, and blood type. While I will explain the definition of the treatment in the next section in more detail, to avoid reverse causality, when I compute engines' capacity by their top-two-place percentages, I exclude the race outcomes of those racers who debuted in the term of the race. Given that every motor gets replaced on an annual basis, I do not use observations before November 1, 1997.

This procedure leaves me 6,952,990 race-by-racer observations with 2717 racers. Among these racers, I focus on racers who debuted during the period between November 1, 1997 and October 31, 2011, which leaves 672 racers who debuted in this period. For the long-term analysis, I will label those racers who were assigned lucky engines at least once during their debut term as the treatment group. I then trace the effect of initial shocks over time. Some of these racers exit from the market. With the variation in the timing of debut and market exits, I use the unbalanced panel of 672 racers.

As for the prize amounts, they are not included in the race results, and earnings are not recorded in the racer demographics. I obtained the administrative data on the annual earnings of each racer from the Boat Racing Association, while the annual earnings data are based on total earnings from January 1 through December 31 of calendar years. To conduct analysis based on terms that are finer units than yearly units, I manually collect prize amounts from various sources as much as possible. The details of the data sources are explained in the Appendix. The correlation between the estimated earnings and the actual earnings from the administrative data is approximately 0.98 on an annual basis, and therefore, I use the estimated earnings unless noted.

The outcomes of interest throughout this paper are (a) the number of first places, (b) earnings, (c) the number of GII or above tournament races, and (d) the proportion of races in which a racer changes his or her initial lane toward the first lane. Together with exit rates, outcome (a) and (b) are meant to capture the merit or success of individual racers, (c) is to measure whether racers aim higher because because some racers strategically decline offers to attend higher-graded tournaments to ace races at lower-graded tournaments, and (d) is to quantify risk-taking behavior.

4.1 Definition of Treatment for the Long-run Analysis

The definition of the treatment is as follows. Every year, some new graduates from the Yamato school debut on May 1, while others debut on November 1. Note that average racers are assigned eight tournaments during their debut term. To define lucky engines, I first compute the winning percentage for finishing within the top two places of every motor at each stadium excluding those new racers outcomes in their debut terms. Then, since the random allocation of engines is at the racer-by-tournament level, for each new racer in the debut term, I count the number of assignment of lucky engines during the debut term. I label those racers who are assigned lucky engines at least once in their debut term.

Table 1: Balance Test

	(1) Control	(2) Treat	(3) (1) vs. (2)
Male	0.833	0.855	-0.021
	(0.020)	(0.019)	(0.028)
Age	21.155	21.000	0.155
	(0.128)	(0.124)	(0.178)
Height	162.837	163.466	-0.630
	(0.298)	(0.262)	(0.398)
Weight	51.645	51.784	-0.139
	(0.153)	(0.148)	(0.213)
Blood type B	0.199	0.191	0.008
	(0.022)	(0.022)	(0.031)
Born in Apr or May	0.158	0.148	0.009
	(0.020)	(0.020)	(0.028)
Pre-debut Winning Rate	4.843	4.945	-0.102
	(0.059)	(0.059)	(0.084)
N	342	330	672

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As for the definition of lucky engines, I restrict winning percentages for finishing within the top two places to be at or above the 90th percentile. The results are similar and robust with different cutoffs of 85th and 95th percentiles, while the five percentile difference results in large change in the proportion of the treatment and control group.

Table 1 shows the results of the standard balance tests for the sample using the values prior to racers' debut terms. In Table 1, the first row is the proportion of male racers, the second row racer age, the third row racer height, the fourth row racer weight, the fifth row racer blood type being type B, the sixth row whether or not racers are born in April or May, and the last row, pre-debut winning rate, corresponds to the pre-debut performance (the sum of race points explained above divided by the number of races) at the Yamato school. At the Yamato school, there are seven tournaments throughout training periods, and those who have high winning rates are specially treated and assigned the number of races equivalent to B1-graded racers.[12] The balancedness of being born in April or May is tested since school starts in April in the Japanese education system, and it is known that early-born children have advantages. Table 1 demonstrates the balancedness in each of these variables between the treatment and control group.

4.2 Engine and race outcomes

Before the main results, I shall check if good engines actually affect the race outcomes of new racers. Using the debut terms of new racers, I conduct an OLS regression of the indicator of the first or second-place finish on engine capacity measures.

[12]The special debut cutoff is three-folded: (i) the winning rate above 6.50, (ii) the winning percentage for finishing within the top two places higher than 30%, and (iii) the penalty rate below 0.7.

Table 2 presents how luck and risk-taking behavior affect the outcomes of races today. The first column shows the effects of the continuous measure of engine capacity and the effects of risk-taking behavior through changing lanes toward the first lane on the binary outcome of whether a racer finishes within the top two places. The second column demonstrates how the effects of the continuous measure of engine capacity differ by the initial ability measured by racers' performance at the training school. The third column includes the differential effects of risk-taking by the initial ability. The last column show the binary measure of engine capacity whose cutoff is the top ten percentile. Due in part to the non-linearity and in part to the loss of variation, while the binary measure capture engine capacity effects, there is no statistically significantly different effects by the initial ability.

Table 2: How Luck And Risk-taking Affect Races Today

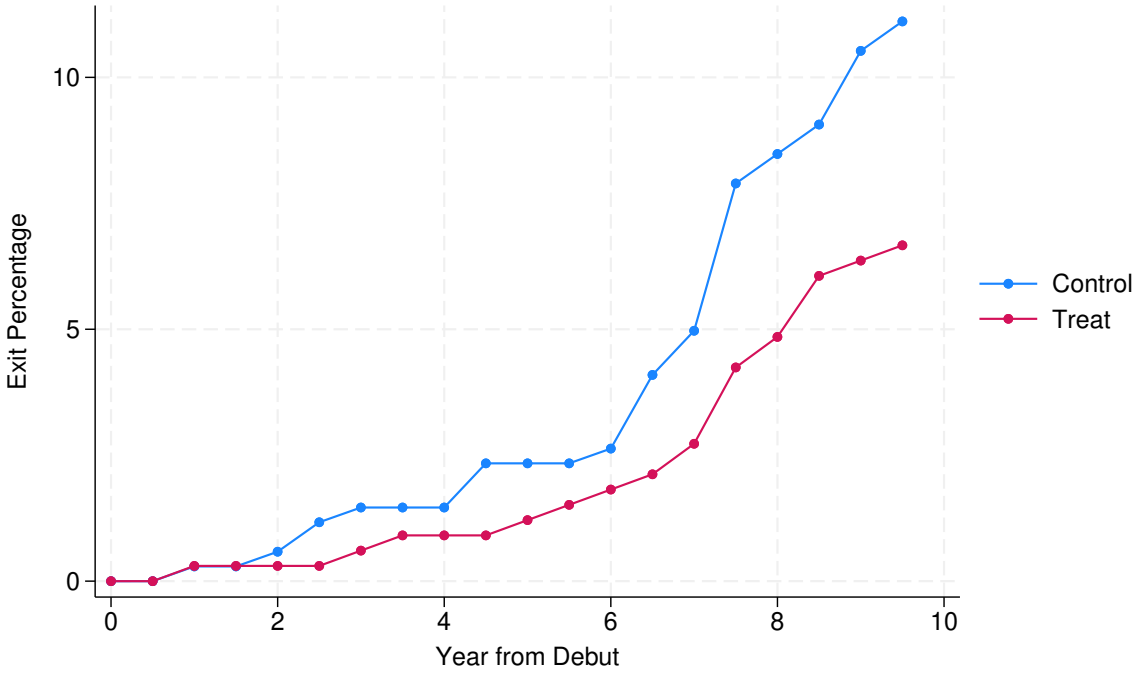
	(1)	(2)	(3)	(4)
	1st or 2nd	1st or 2nd	1st or 2nd	1st or 2nd
EngineCapacity	0.179***	0.111***	0.104***	
	(0.026)	(0.029)	(0.028)	
Changed Lane toward the 1st	0.089***		0.045**	
	(0.018)		(0.019)	
EngineCapacity x High Ability		0.154**	0.170**	
		(0.053)	(0.054)	
Changed Lane toward the 1st x High Ability			0.097**	
			(0.037)	
Lucky This Tournament				0.013**
				(0.006)
Lucky This Tournament x High Ability				0.018
				(0.012)
Observations	36038	39658	36006	39658
Control Mean	0.078	0.078	0.078	0.078
Control	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the racer-by-tournament level in parentheses. The first column shows the effects of the continuous measure of engine capacity and the effects of risk-taking behavior through changing lanes toward the first lane on the binary outcome of whether a racer finishes within the top two places. The second column demonstrates how the effects of the continuous measure of engine capacity differ by the initial ability measured by racers' performance at the training school. The third column includes the differential effects of risk-taking by the initial ability. The last column show the binary measure of engine capacity whose cutoff is the top 10 percentile. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.3 Exit Trend: Extensive-margin Results

Using the treatment definition above, Figure 3 shows the exit rate trend over the first ten years since debut for the treatment and control groups. At the end of the tenth year, the gap amounts to approximately 4.7 percentage point difference (11.5% for the control group and 6.8% for the treatment group) that is 41% decrease in the exit rate based on the control group exit rate. The treatment effect is statistically significant using the t-test. This shows the enormous butterfly effects already. Also, since these retiring racers are in general those racers who perform poorly, my intensive-margin measure inevitably suffer from downward bias. Therefore, it is important to note that my intensive-margin measure results are generally lower bounds of the treatment plus competition effects.

Figure 3: Exit Trend of Treatment And Control



5 Results and discussions

5.1 Short-term Effects

5.2 Effects on Risk-taking at the Next Tournament

As Table 3 demonstrates, the lucky event induced racers to take *less* risks at the next tournament and win more compared to those who were not assigned lucky engines at the previous tournament. The first column shows the effects of being assigned lucky engine at the previous tournament on the binary measure of whether a racer changed his or her lane toward the first lane. The effects are negative

and reach more than 40% decrease in the risk-taking behavior. The second column demonstrates no differential effects by the initial ability. The third column displays no effects on risk-taking behavior at the following tournament after the next one. The last column shows that this effect is not driven by the initial lane assignment difference measured by the difference between the initial lane of the first race within the previous tournament and that of this tournament.

The results demonstrate how success makes racers less risk-taking immediately after the success, which does not last long.

Table 3: How the Luck Affect Risk-taking Behavior at the Next Tournament

	(1) LaneChange	(2) LaneChange	(3) LaneChange	(4) diff
Lucky Previous (m-1) Tournament	-0.006**	-0.007**		0.052
	(0.002)	(0.003)		(0.099)
Lucky Previous x High Ability		0.001		0.032
		(0.005)		(0.166)
Lucky m-2 Tournament			-0.005	
			(0.004)	
Lucky m-2 x High Ability			0.008	
			(0.006)	
Observations	30209	30179	25747	4572
Control Mean	0.014	0.014	0.014	-0.084

Notes: Standard errors clustered at the racer-by-tournament level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2.1 Confidence Channel

Now, I examine how luck affects the performance of racers at the next tournament and the following tournament after the next. Table 4 presents the results. The first column shows the overall effects of the assignment of a lucky engine at the previous tournament based on the binary measure on whether a racer finishes the top two places. The second column shows the differential effects by the initial ability. While there is no overall effect of the luck at the previous tournament on the performance of the current tournament, there is weak evidence of such effects for the low-ability individuals at the 10 percent significance level. The third column shows if there is an effect of the lucky assignment today on the performance at the tournament following the next one, while the last column presents no differential effects by the initial ability.

In sum, while it is weak evidence, it is remarkable that there is a positive effect on the race performance at the next tournament for the low-ability individuals, despite the fact that they take less risk. Then, there must have been some additional effects of luck at the previous tournament to drive up the performance of low-ability racers. I attribute this to confidence boost, assuming that low-ability individuals have less confidence than high-ability individuals. The confidence boost appears to last even after the next tournament.

Table 4: How the Luck Affect the Next Tournament

	(1) 1st or 2nd	(2) 1st or 2nd	(3) 1st or 2nd	(4) 1st or 2nd
Lucky Previous (m-1) Tournament	0.003	0.012*		
	(0.005)	(0.006)		
Lucky Previous x High Ability		-0.021*		
		(0.011)		
Lucky m-2 Tournament			0.014**	0.014*
			(0.006)	(0.008)
Lucky m-2 x High Ability				0.001
				(0.012)
Observations	34679	34650	29649	29619
Control Mean	0.081	0.081	0.081	0.081
Sample	All	All	All	All

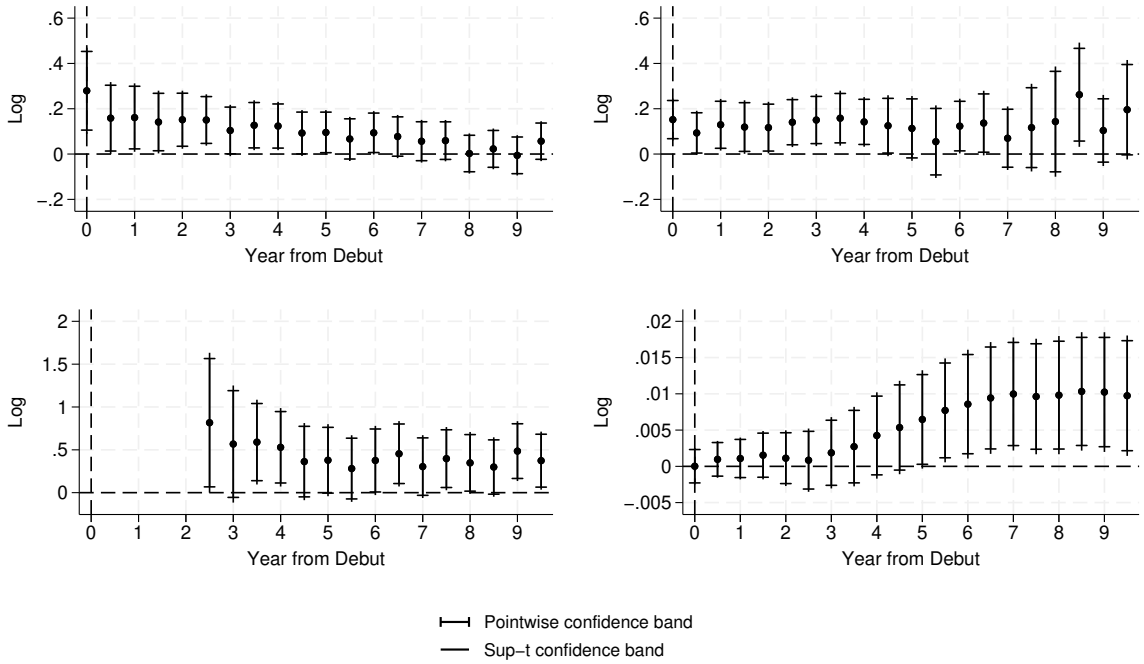
Notes: Standard errors clustered at the racer-by-tournament level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Long-term Intensive Margins

In Figure 4, coefficients with 95% confidence intervals and uniform sup-t confidence intervals from equation (1). Panel (a) of Figure 4 corresponds to the total number of the first place, (b) the total earnings, (c) the total number of GII or above tournament races, and (d) the proportion of the number of races in which racer i changed lanes toward the first lane, all at the term level.

From panel (a), we can see persistent effects on the term-wise total number of first-place finishes, meaning that the cumulative difference widens over time. The disappearnce of the effects after the first seven years are most likely driven by the downward bias through market exits of poorly performing control group racers. A similar pattern can be seen in the other variables including the total earnings, the number of G2 or above races, and the proportion of lane change toward the first lane.

Figure 4: Dynamics of Initial Luck
(a) Number of Total 1st Places (top-left), (b) Total Earnings (top-right), (c) Number of G2 Races (bottom-left), (d) Proportion of Races with Lane Change (bottom-right)



6 Conclusion

This paper demonstrated both short-term and long-term effects of luck. In particular, I showed that luck yesterday does affect risk-taking behavior and success tomorrow. Then, I turn to the long-term analysis and demonstrated how initial luck that is small have persistent effects over time leading to a widening gap in cumulative outcomes. These results prove how fragile the basis of meritocracy is and shed light on how much of our merit is induced by random luck even if we start at the same starting point.

References

ALMOND, D. (2006): "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," *Journal of political Economy*, 114, 672–712.

BAILEY, M. J., S. SUN, AND B. TIMPE (2021): "Prep School for poor kids: The long-run impacts of Head Start on Human capital and economic self-sufficiency," *American Economic Review*, 111, 3963–4001.

BEDARD, K. AND E. DHUEY (2006): "The persistence of early childhood maturity: International evidence of long-run age effects," *The Quarterly Journal of Economics*, 121, 1437–1472.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics*, 119, 249–275.

BOOTH, A. AND E. YAMAMURA (2018): "Performance in mixed-sex and single-sex competitions: What we can learn from speedboat races in Japan," *Review of Economics and Statistics*, 100, 581–593.

CUNHA, F., J. HECKMAN, AND S. NAVARRO (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *oxford Economic papers*, 57, 191–261.

DU, Q., H. GAO, AND M. D. LEVI (2012): "The relative-age effect and career success: Evidence from corporate CEOs," *Economics Letters*, 117, 660–662.

FUJINO, T.-I. (2006): *Yokuwakaru Kyotei no Subete*, Tokyo: Saneki Books.

GENDA, Y., A. KONDO, AND S. OHTA (2010): "Long-term effects of a recession at labor market entry in Japan and the United States," *Journal of Human resources*, 45, 157–196.

GINSBURGH, V. A. AND J. C. VAN OURS (2003): "Expert opinion and compensation: Evidence from a musical competition," *American Economic Review*, 93, 289–296.

HECKMAN, J. J. AND S. MOSSO (2014): "The economics of human development and social mobility," *Annu. Rev. Econ.*, 6, 689–733.

HIMURA, K.-I. (2015): *Yokuwakaru Boat Race no Subete (Basic Knowledge about Boat Race)*, Tokyo: Saneki Books.

——— (2017): *Boat Race no Subetegawakaru (from Introduction to Betting Techniques)*, Tokyo: Saneki Books.

——— (2022): "Motor Seinousa wa Kaishodekinai," [Online article published by a website operated by Makuuru Magazine at https://sp.macour.jp/columns/himura/160519/; posted 6-June-2022].

HOUMARK, M. A., V. RONDA, AND M. ROSHOLM (2024): "The Nurture of Nature and the Nature of Nurture: How Genes and Investments Interact in the Formation of Skills," *American Economic Review*, 114, 385–425.

KAWAGUCHI, D. (2011): "Actual age at school entry, educational outcomes, and earnings," *Journal of the Japanese and International Economies*, 25, 64–80.

MACCINI, S. AND D. YANG (2009): "Under the weather: Health, schooling, and economic consequences of early-life rainfall," *American Economic Review*, 99, 1006–1026.

MULLER, D. AND L. PAGE (2016): "Born leaders: political selection and the relative age effect in the US Congress," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 179, 809–829.

SANDEL, M. J. (2020): *The tyranny of merit: What's become of the common good?*, Penguin UK.

YAMAGUCHI, S., H. ITO, AND M. NAKAMURO (2023): "Month-of-birth effects on skills and skill formation," *Labour Economics*, 102392.

Appendices

A Earnings Data Details

In this section, I note the details of how I computed earnings. Unfortunately, the JSRA does not list the prize amounts and Himura (2015) and Himura (2017) do not list them other than the winner prize amount of the winner of graded tournaments. The main source of prizes for graded tournaments and a lower bound for non-graded tournaments come from two online articles written by two speedboat racing predictors.[13] I fill in the incomplete list described in the data section.

[13] The main source is https://kyoutei-navi.com/sg/grade/ supplemented by https://boat-race.biz/4788 since the first source lacks prize amounts for some within-tournament races such as "dream match" races. Both are accessed on November 13, 2023.

B Institutional Details And Rules

B.1 How Races Stage Within A Tournament

Consider a GIII tournament spanning six days. Racers participate in multiple preliminaries from day 1 to day 4, with the top 18 moving on to semi-finals. Racers not making the cut participate in "standard competition" races on day 5 and 6. Most of the tournaments employ the three semi-finals system, while few employ other systems such as two semi-finals and no semi-final. Generally speaking, those racers with higher points will be assigned more inward and thus better initial pits and lanes than the other racers on day 5. Day 6 features finals, "specially selected A group" races for the third and fourth-place racers, "specially selected B group" races for the fifth and sixth-place racers, and another round of "standard competition" races. The number of races within a tournament remains the same for racers and engines in general.

B.2 Racer Grade Details

To be graded as an A1 racer, a racer has to achieve the following three criteria. One is the winning percentage for finishing within the top two places higher than 30% and the winning percentage for finishing within the top three places higher than 40%. Another is a boat accident/penalty rate below 0.7, which will be explained shortly. The last criterion is the number of races attended. This last criterion was modified from 70 races to 90 races and became effective as of the second term of 2010.[14] Among those who satisfied these criteria, some top racers are awarded A1 grade, so that the top 20% of all racers are graded A1. These A1 racers are eligible for most of the graded tournaments, and among these A1 racers, some will be selected to participate in SG tournaments. A2 grade has the same criteria, except for the minimum number of races to remain at 70 even after the second term of 2010. Those racers who satisfy these criteria but are not screened into A1 grade will be awarded A2 grade in a way that the proportion of A2 racers becomes roughly 20%. A2 grade racers are also eligible for most of the graded tournaments except SG tournaments, while an A2 racer is eligible for an SG tournament if he or she is the winner of the SG tournament in the previous year. The vast majority of non-A1 and non-A2 racers will be categorized into B1 grade, except new racers and those racers who caused boat accidents in general. The criteria for B1 grade changed in the first term of 2016. Prior to this term, the criteria were (a) the winning percentage for finishing within the top two places above 10%, (b) the winning percentage for finishing within the top three places above 20%, and (c) the minimum number of races above 50, so that 50% of all racers are classified into B1. These criteria, however, became too strict to achieve 50% proportion toward 2016. The new criteria are criterion (c) plus (d) total points in the previous term divided by the total number of races in the previous term is above 2, without

[14] From http://www.kenkyu.co.jp/station/archives/2010_10_4_106.html accessed on November 2, 2023.

criteria (a) and (b).[15] To prevent racers from earnings points by attending lower-graded or non-graded tournaments only, the Association assigns racers to tournaments to equalize opportunities as much as possible for non-graded tournaments. In general, racer grade A1 and A2 racers are allocated races of approximately 15 days, B1 around 12 days, and B2 around 8 days within each month, and each non-graded tournament consists of racers proportional to the distribution of these racer grades. A racer can decline tournament participation offers from the JSRA for reasons like family issues, but a racer cannot decline an offer to attend one tournament and choose to participate in another tournament that takes place during the same period.

B.3 Prize Component

There are two components in the prize amounts. One is compensation for participating in races. Then, even the sixth (last) place in each race offers some money to participating racers. The other part is the actual prize/award for placements. Note that every placement generally offers some prizes that are the highest for the first place and lowest for the sixth place. The amounts of these prizes are higher in the semi-finals and even higher in the finals than the standard competition races within a tournament.

B.4 Propellers And Mentorship

Speedboat racing society adopts the apprenticeship system, and mentors sometimes give mentees propellers. The random assignment must equalize the probability of having a mentor with good propellers.

B.5 Penalties and Retirement Admonishment

The JSRA sets penalty points for the following acts: (a) flying or late start at the finals of any tournament (30 points), (b) flying or late start at the non-finals of any tournament (20 points), (c) disqualification for interference (15 points), (d) disqualification or absence in a race for a racer's lapse (10 points), and (e) (minor) bad boat navigation and violation of standby behavior (2 points). If a racer's penalty rate (total penalty points divided by the number of races) within one term surpasses 0.7, or if the racer's total number of races attended within one term does not reach 60, this racer will be downgraded to racer grade B2. Furthermore, every flying or late start will result in the cancellation of tournament assignments for 30 days following the last tournament that has already been assigned and scheduled for the penalized racer.

If a racer poorly performs throughout the most recent four consecutive terms, he or she is "recommended" to retire. While the recommendation is not enforceable, most racers do quit following the recommendation. The JSRA sets the standard for recommendation as either (i) the winning rate (the

[15] From https://www.boatrace.jp/owpc/pc/site/news/2016/03/1236/ accessed on November 2, 2023.

sum of race points explained above divided by the number of races) of the most recent four consecutive terms is lower than 3.80, (ii) penalty rates of the most recent four consecutive terms is above 0.7, or (iii) the total number of races attended throughout the four consecutive terms is below 60. Note that this rule does not apply to new racers for the first three years. Then, the first retirement recommendation can take place at the end of the fifth year of racers since their debut.