Dying to Win? Olympic Gold Medals and Longevity*

Adam Leive

University of Virginia

April 5, 2017

Abstract

This paper investigates how status affects health by comparing mortality between Gold and Silver medalists in Olympic Track and Field. Contrary to conventional wisdom, winners die over two years earlier than losers. Analysis of individual Census records of each U.S. athlete and his parents suggests that income is the key mechanism: losers pursued higher-paying occupations than winners after the Olympics, while parental earnings in childhood were similar. The results suggest that how people respond to pivotal life events can produce long-lasting consequences for health.

JEL codes: I12, I31, J10

*I am grateful to helpful comments from Iwan Barankay, Henry Bergquist, Guy David, Matt Grennan, Scott Harrington, Jonathan Kolstad, Luis Rayo, Amanda Starc, Robert Town, and Nicolas Ziebarth. I thank Steve Holman for his feedback and insights about his experiences in Olympic Track and Field. Funding from the National Institute of Aging through Grant Number T32-AG000186 to the National Bureau of Economic Research and from the Ackoff Fellowship from the Wharton Risk Management and Decision Processes Center is gratefully acknowledged.
Address: L050 Garrett Hall, Batten School of Leadership & Public Policy, University of Virginia, Charlottesville, VA, 22904.
Email: leive@virginia.edu.

I. Introduction

Competition for status is ubiquitous in both professional and social settings. This paper studies how status competition affects long-term health. Disentangling the relationship between status and health is challenging because several channels may operate simultaneously. First, higher status can directly expand income opportunities or other real resources that impact health. Second, higher status may produce psychological effects on health, through changes in stress levels, for example. Third, the very pursuit of status may harm health: time spent working may crowd out labor inputs to health like exercise, or conspicuous consumption may displace inputs purchased in the market like medical care. Fourth, obtaining higher status may affect future motivation and thereby influence real resources and health. Finally, a third variable, such as latent ability, could independently determine both status and health.

The existing literature has struggled to separate these mechanisms. The Whitehall studies of British civil servants provide epidemiological evidence of a positive relationship between status and health in an employment setting (Marmot et al., 1991), but endogenous selection into jobs suggests causality does not run from status to health (Chandra and Vogl, 2010; Case and Paxson, 2011). Other research focusing on well-defined occupations in which status is based on receiving awards—Nobel laureates, Oscar winners, and Major League Baseball Hall of Famers—also tends to find a positive association between status and longevity (Sylvestre, Huszti and Hanley, 2006; Becker, Chay and Swaminathan, 2007; Rablen and Oswald, 2008). However, unobserved heterogeneity between winners and losers and the non-random assignment of status from these contests raises doubts that the results should be interpreted as measuring the effect of status on health.

In this paper, I compare mortality between Gold and Silver medalists in Olympic Track and Field between 1896 and 1948 to overcome these challenges. While the setting is highly specific, its institutional features provide advantages that help to cleanly identify status and distinguish between the channels listed above. Track and Field includes events in running, jumping, and throwing that use only time or distance to objectively measure performance. In each event, the order of finishers creates a clear and undisputed ranking, even though the differences between competitors may be just fractions of a second. The stakes of such competition are high, with an Olympic victory representing the pinnacle of the sport and carrying global recognition. Variation in status is based simply on winning or losing.

Conditional on reaching the Olympic final, randomness plays a larger role in deciding the difference between winners and losers compared to contests judged over a longer time period. The Olympic Gold medalist is determined on a single day every four years. As I later document, the athlete with the best performance in the year prior to the Olympics often fails to win the Olympic final. Additionally, more than half of athletes who set a World Record never win Olympic Gold.[1] Prior success clearly does not guarantee victory in the Olympics.

Another advantage of this setting is that athletes are physically similar in terms of their baseline health by virtue of their participation in the Olympic final. As supporting evidence of this claim, I show that differences in ability between Olympic finalists—which may positively correlate with both health and winning—do not predict mortality by comparing athletes who ever held World Records (the highest ability group) to those who never did. Since athletes are generally young during Olympic competition, there is also less concern that results are biased by reverse causality in which health determines status. However, performance-enhancing drugs (PEDs) complicate this relationship to the extent that PEDs influence both health and the chance of winning. Since it is difficult to determine which athletes use PEDs, I restrict my analysis to the period 1896 to 1948, when there was less suspicion or evidence of PEDs in Olympic competition.[2]

In addition, income directly earned from competition during this period was non-

[1]This statistic excludes athletes who competed during the years when the Olympics was canceled due to WWI and WWII and so is not artificially deflated.

[2]The International Olympic Commission first produced a list of banned substances in 1968. Some athletes experimented with substances to improve performance that also had health effects in the early 1900s, although doping strategies were not yet advanced. For example, George Hicks won the 1904 marathon after consuming raw egg, Strychnine (a poison that also functioned as a stimulant), and brandy. Drugs yielding significant performance benefits like anabolic, androgenic steroids were not used until the 1950s and amphetamines not until the 1960s (Wadler 1998, WADA 2010).

existent due to the prevailing system of amateurism, which prevented athletes from receiving financial compensation tied to their performance. Until the 1980s, regulations prohibited professional athletes from competing in the Olympics and most Olympians held other occupations while training. The Gold medal itself was also worth a modest amount in terms of its metallic content (Economist, 2012).

Matching data on Olympic finishing order with each athlete's date of birth and death, I first document that Gold medalists die 1.5 years earlier than Silver medalists. I estimate survival models that control for observables like height, weight, country, event, and year of birth, which may be correlated with both finishing place and longevity. Using supplementary data on the history of each athlete's performances, there is little evidence the correlation between winning and mortality is driven by selection, in which Gold medalists invest more time or effort training that harms health. Gold medalists did not compete for longer periods of time and pre-Olympic rankings were often similar to Silver medalists.

There is empirical support that income earned later in life is perhaps the key mechanism between winning and a shorter lifespan in this population. Focusing on the sub-sample of U.S. athletes, I collect data on earnings and occupational choices for athletes appearing in the 1940 U.S. Census, which was the first Census to record income. Compared to Gold medalists, Silver medalists earned higher incomes and were more likely to enter professional occupations. Losing is again correlated with a lower hazard of death in this sub-sample, but this effect disappears once income is accounted for. There is a large and statistically significant association between higher income and a longer lifespan. Including income explains 60 percent of the variation in lifespan, roughly twice as much as baseline models with finishing place and other observables alone.

Additional analysis of the Census records of each athlete's family provides suggestive evidence that economic conditions in childhood were likely similar between Gold and Silver medalists. I link athletes appearing in the 1940 Census to their family's earlier records in the 1910, 1920, and 1930 Censuses and collect the occupation of each athlete's parents. Using constructed earnings estimates by occupation from the Integrated Public Use Microdata Series (IPUMS), I test whether parental earnings were equal between Gold and Silver medalists and fail to reject the null of no difference in means. In regression models, the coefficient estimate on the athlete's income remain a statistically significant predictor of lifespan while the estimate on parental earnings is not. Based on a range of specifications, it is unlikely that omitted variable bias explains the relationship between losing and health. The analysis of Census records of each athlete's post-Olympic earnings and his family history is consistent with relative rank influencing motivation. The data does not allow me to distinguish, however, whether losing motivates or winning de-motivates.

This paper's results challenge conventional wisdom and the conclusions from existing studies that being awarded higher status necessarily improves health (Marmot et al. 1978, 1991; Sylvestre, Huszti and Hanley 2006; Becker, Chay and Swaminathan 2007; Rablen and Oswald 2008). Instead, losing can have positive, first-order effects on longevity. While Olympic Track and Field is a highly stylized setting, many people face pivotal life events defined by either success or failure. This study may thus have broader implications for understanding how the binary outcomes of important trials in life can produce long-lasting consequences for health.

II. Status, Health, and the Olympics

Researchers cannot randomize people into groups of high and low status and measure how long they live. Experimental studies among non-human primates, however, provide some insights into how random changes to status affect health. Some biological research shows that higher status can improve psychological and physical health by reducing stress. One study of rhesus macaques pinpointed the molecular mechanisms behind such psychosocial responses, demonstrating that manipulating social status (dominance rank) affects gene regulation tied to immune defense (Tung et al., 2012). By contrast, other biological studies find that under certain conditions, such as when the hierarchy is unstable, the highest-ranking animals experience the greatest stress from psychosocial factors (Sapolsky, 2005). Yet even experimental studies have not tracked the longevity of animals at different ranks of a randomly assigned hierarchy to study long-term health outcomes.

Observational research in humans—mostly from the epidemiology and medical literature—has generally found a positive gradient between health and status. The Whitehall Study of British Civil Servants in the 1960s and its second iteration in the 1980s demonstrate a marked social gradient in health across different ranks of government employees (Marmot et al., 1978, 1991, 2001; Marmot and Feeney, 1997). Conventional risk factors explain only one third of the difference in mortality risk between clerical and administrative grades (Marmot and Brunner, 2005).[3] The later Whitehall research focuses on social support and the organization of the workplace as possible channels between status and health.

While the Whitehall research clearly reveals an important and sizable link between status and health, the likelihood of endogenous selection into Civil Service ranks raises concerns about how to interpret the results. It is difficult to disentangle the extent to which higher status led to better health or whether better unobserved initial health led to or was otherwise correlated with higher status (Chandra and Vogl, 2010). As evidence of selection, Case and Paxson (2011) find that current self-assessed health in the Whitehall II sample predicts future civil service grade, but current civil service grade does not predict future self-assessed health. In addition, some research also disputes the mechanisms between status and health analyzed in the Whitehall research. A prospective cohort study of Finnish industrial employees found that low predictability at work was highly correlated with heart attack risk, but other organizational factors highlighted by Whitehall—such as low decision autonomy at work—were not (Vaananen et al., 2012).

Outside of Whitehall, research has examined major shocks to status from receiving

[3]For heart disease, for example, clerical workers faced a relative risk of dying that was 2.2 times higher than senior administrative staff, and 1.6 times higher than employees in intermediate professional and executive positions.

awards, such as winning the Nobel Prize (Rablen and Oswald, 2008), election to the Major League Baseball (MLB) Hall of Fame (Becker, Chay and Swaminathan, 2007), or receiving an Oscar (Sylvestre, Huszti and Hanley, 2006). The assumption behind these studies has tended to be that status should improve health, and there is support for this for the Nobel Prize and MLB Hall of Fame but inconclusive results for Oscar winners. However, unobserved heterogeneity and the process of choosing winners may limit what can be drawn from the findings. For example, the physical attributes of Oscar nominees differ in ways that affect their health, and bias may stem from correlation with the likelihood of winning an Oscar. People may also undertake different lifestyle decisions, follow different diets, and value their health in unobserved ways. The same might be said for Nobel laureates and their peers. Moreover, actors, baseball players, and academics are all professionals who can be financially compensated for their work. Higher income associated with status may thus confound comparisons of longevity within these populations. Since Track and Field athletes in the early 1900s were all amateurs, the Olympic setting does not face this problem. Another issue is that these other studies judge performance over a longer time frame. For example, baseball players nominated for the Hall of Fame are assessed over their entire career. The long duration of such assessment increases the chance that the factors that lead people to succeed may be correlated with their mortality prospects. It is reasonable to believe that there is less unobserved heterogeneity among Olympic athletes within any given event than among Nobel laureates, Oscar nominees, or MLB players.[4]

Although there is limited economic research on how status affects health, economists have studied the importance of social comparisons both from theoretical and empirical perspectives.[5] Recent models have incorporated peer comparisons and habit formation, rather than assuming utility depends only on absolute consumption levels. In this way, status con-

[4]Not surprisingly, the longevity of Olympians is greater than the general population. Clarke et al. (2012) document that Olympic medalists across all sports live almost 3 years longer than other people of the same age, sex, and country. The research did not explore the reasons for this difference, which might be due to genetics, exercise, diet, income, status, or other factors.

[5]See for example Frank (1985); Easterlin (1995); Clark and Oswald (1996); Falk and Knell (2004); Luttmer (2005); Rayo and Becker (2007b); Heffetz and Frank (2011); Heffetz (2011).

veys hedonic value as economic conditions are compared to a benchmark that may depend on personal history, expectations, and the success of one's peers. For example, Rayo and Becker (2007a,b) develop a model in which agents adjust to a time-varying reference point based on habits and peer comparisons, and make output choices given their preferences. Depending on the underlying parameters, agents may fully adjust to their new reference point quickly or they may habituate only partially and gradually over time. The speed of habituation affects how the agent views his current economic conditions compared to past successes or failures.

Lifestyle decisions and occupational choices represent a potentially important channel between status and health. In analyzing obituaries published in the New York Times, Epstein and Epstein (2013) find that actors, singers, musicians, and athletes die several years earlier than academics, politicians, business executives, and other professionals. The study suggests the earlier death of the former group may result from greater fame and risky behaviors. If success permanently shifts an agent's reference point, risky activities may be rational attempts to attain the utility achieved at the peak of professional success. After Olympic athletes are finished competing, their occupational choices and lifestyle decisions likely affect their health too. Such choices relate to how winning and losing influences motivation. Evidence from field experiments is mixed on whether information about rankings is motivating or discouraging; some research suggests peer comparisons improve future performance (Tran and Zeckhauser 2012), while other studies find informing employees of their relative rank reduces future effort (Barankay 2012a,b).

III. Setting and Data

I focus on the setting of Track and Field because it is the oldest sport where performance is objectively measured.[6] Fewer nations and fewer athletes compete in Swimming or Cycling

[6]The only sport for the first 13 of the ancient Olympic Games that began in 776 BC was a 1-stadium length sprint—called the "stadion"—measuring 192 meters (Perrottet, 2004).

than in Track and Field. I analyze male athletes only since women did not compete in Olympic Track and Field until 1928, with some events being limited to males until the 1990s.

The data includes the order of finish in the Olympic final for each event, the country the athlete competed for, and athlete's birth and death dates, collected from a request to the Olympic Studies Center of the International Olympic Committee and the site Olympedia.org. I focus on comparing Gold to Silver medalists in the spirit of a regression discontinuity design, with some analyses also comparing Gold to other finalists. For most athletes, the data also includes height (measured in centimeters) and weight (measured in kilograms) at the time of the Olympic Games. Appendix Table A.1 lists the number of observations per country to provide a sense of the geographical composition of the sample. I calculate lifespan as the number of days between the athlete's dates of death and birth.

I classify "high ability" athletes as those who ever held multiple World Records, including after the Olympic Games. Ability may be positively correlated with both winning and latent health. I use two or more World Records as the threshold for high ability in case athletic performances from the tails of a distribution are due to random variation. An advantage of this metric for ability is that it is clearly defined.[7]

I impose several sample restrictions to cleanly focus on the relationship between Olympic performance and lifespan. I exclude athletes with recorded deaths due to war because these causes are arguably exogenous and unrelated to behavior. The site Olympedia.org maintains a list of such deaths. Some athletes also may have died from non-biological causes, such as car accidents. It is not possible to determine whether such deaths are random or due to risky behavior. To be conservative, my main sample does not exclude deaths due to accidental causes. Robustness tests in Appendix B exclude athletes who died prior to age 40 to assess the sensitivity of the results to potential outliers. Athletes whose date of death is missing (less than 3 percent of the sample) are dropped since I am unable to verify their

[7]The personal bests of each athlete over their career could be another way to control for ability, but these are highly collinear with year of birth since each event's top performances improve over time.

death. Some of these athletes may still be alive, but excluding them may be a safer strategy since athletes who win Gold may be more likely to have a recorded date of death than other finalists.

Finally, I concentrate analysis on athletes who finish in the top two in a single Olympic Games and single event, which constitutes the majority of Olympians. If athletes compete in multiple Games, it is not clear how an athlete values each performance relative to the other. Since over three-quarters of the sample competes in a single Olympics, focusing on these athletes provides a standard perspective. The results are robust to including athletes who compete in multiple Olympic Games, as shown in Appendix B, using either the best rank from their first Olympic Games or across Olympic Games. After these restrictions, the final sample of Gold and Silver medalists includes 187 athletes with complete dates of birth, death, and finishing place.[8]

Table I presents descriptive statistics of the sample. The average age at death is nearly 74 years, ranging between 23 to over 100. Observable characteristics such as year of birth, height, weight, and ability (as measured by holding multiple World Records) are balanced between winners and losers as shown in Table II.[9] To preview the main results, Figure I first provides non-parametric, unconditional estimates of lifespan. The Kaplan-Meier survival curves plot the share of Gold and Silver medalists alive at each age. On average, Gold medalists die 1.5 years earlier than Silver medalists, which is given by the difference in the areas between the two curves. By age 80, approximately half of Silver medalists remain alive compared to a third of Gold medalists.

---

[8]Including third and fourth place finishers increases the sample size to 395 athletes and including all finalists increases the sample size to 658 athletes.

[9]The proportion of athletes in each event are not equal because some Gold medalists are not matched to a Silver medalist in the same event and year, and vice versa.

IV. Methods

The main identification strategy compares longevity between winners and losers in each event and year of the Olympics. The hazard models outlined below include indicators for losing, event, and year to exploit within-event-year variation in lifespan, as well as the athlete's year of birth. Some models also include an indicator for ability, defined as ever holding multiple World Records as described in Section III. This model does not control for country since there are few cases where pairs of winners and losers in the same event and year compete for the same country. As an alternative identification strategy, I compare the longevity of winners and losers within countries and broad event classes, where similar events are grouped into sprints, middle distance, distance, throws, field, or racewalk as shown in Appendix A. I aggregate individual events in this specification since including fixed effects for both countries and events may create an incidental parameters problem in non-linear models.[10] To capture heterogeneity in body types within event classes, I control for height in some specifications.[11] These models also condition on ability and year of birth as in the main specification. As shown in Section V, the results are similar under both identification strategies, but I focus on models using within-event-year identification since that comparison more directly represents the cutoff between winning and losing.

I model lifespan using parametric and semi-parametric hazard models. The first model is the standard Cox proportional hazards model:

$$\lambda = \lambda_0(t) \exp(x'\beta) \tag{1}$$

where the hazard of death λ depends on an unspecified baseline hazard $\lambda_0(t)$ and an ex-

[10] The estimates are nonetheless similar if I include indicators for individual events rather than the broader event classes.

[11] Medical research links height to an earlier death due to biological factors, such as reduced cell replication and lower cancer incidence (Samaras 2012). The results are not sensitive to including weight, which is highly correlated with height, or body mass index. I include height alone to avoid potential collinearity problems.

ponential function of observables. The explanatory variable of interest is an indicator for whether the athlete lost the Olympic final, defined as finishing as any place other than first. Robust standard errors are clustered by event and year.[12]

To allow for unobserved heterogeneity, I also estimate a mixed proportional hazards (MPH) model that specifies a Gompertz distribution for the baseline hazard[13] and individual-level heterogeneity (frailty) that has a Gamma distribution by specifying the hazard as

$$\lambda = \nu_i \lambda(t) \exp(x'\beta) \tag{2}$$

where ν_i captures individual-level unobserved heterogeneity as a multiplicative effect on the hazard rate.[14] Finally, as a fully parametric approach, I estimate a Gompertz survival model without individual heterogeneity. The estimates are robust to estimating these various parametric and semi-parametric survival models that make different assumptions about unobserved heterogeneity.

V. Results

Table III presents the regression results of the survival models of lifespan described above. The first two columns present Cox proportional hazard models that include event and year fixed effects. Columns 3 and 4 present mixed proportional hazards models and Columns 5 and 6 present Gompertz survival models. Gold medalists represent the omitted finishing place. In all cases, coefficient estimates are exponentiated and so are interpreted as hazard ratios.

[12] The treatment unit is at the event-year level. The statistical significance is not sensitive to clustering at the country-year level or not clustering at all.

[13] The Gompertz distribution has been the workhorse of actuarial science to model mortality since the distribution provides a simple analytic formula for survival based on the observation from many settings that mortality rises exponentially with age (Olshansky and Carnes, 1997).

[14] As additional specifications to explore robustness, I also estimated survival models with shared frailty by country and the increasingly mixed proportional hazards model of Frijters et al. (2011) that models unobserved individual heterogeneity as a random walk, rather than assumed to be constant over time, and obtained similar results.

The hazard estimate of 0.714 in Column 1 indicates 71.4 percent as many Silver medalists are expected to die at any point compared to Gold medalists. The coefficient estimates are similar across models, with the estimate on losing being statistically significant at the 10 percent level. There is very small variance on the unobserved heterogeneity in the MPH models, and so the coefficient estimates from the Gompertz models are nearly identical to those from the MPH models.

Comparing Silver to Gold medalists arguably presents the sharpest cut-off between winning and losing, but the results also hold when including other Olympic finalists who also lost. Table IV presents results that compare the longevity of Gold medalists to that of Silver medalists, Bronze medalists, and 4th place finishers (Columns 1 to 4) or to all other finalists (Columns 5 to 8). Again, losing is associated with a lower hazard of death, and the estimates are similar to the main results in magnitude. The larger sample sizes increase the precision of the estimates.[15]

The estimates are close in magnitude, but sometimes slightly less precise, when using within-country variation rather than within-event-year variation for identification. Table V presents results that include country fixed effects and compare Gold vs. Silver medalists (Columns 1 to 4) and Gold vs. Silver, Bronze, and 4th place finishers. The estimated hazard ratios range between 0.718 and 0.759, in line with the results from Tables III and IV. Without year effects, the athlete's year of birth is now statistically significant in these specifications.

One might expect the association between winning and mortality to be stronger in Olympic Games that were more highly publicized. To investigate this question, I run regressions that split the sample into two halves before and after 1924. Table VI shows the results are driven by later Olympic Games, which were more widely covered through print media, radio, and television. The 1924 Paris Games were the first to be broadcast on radio and the 1936 Berlin Games were the first to be televised, for example.[16] In the later period,

[15]The mixed proportional hazards model failed to converge for regressions including all other finalists, and so Gompertz regression estimates are instead presented in those cases (Columns 7 and 8).

[16]The Olympics received more news coverage in later years. Performing a search for articles with the word "Olympics" on the New York Times site during the entire year of an Olympic Games reveals the following

47.1 percent of Silver medalists are expected to die relative to Gold medalists at any given time (Column 3). The estimates on losing from the earlier period are still below 1 but not statistically significant.

VI. Tests of Selection

This section provides evidence that the correlation between lifespan and Olympic finish is likely not due to selection. Specifically, one might be concerned that Gold medalists spend more time or effort training than Silver medalists and other finalists. Such additional training could perhaps directly harm health or could crowd out other activities that might benefit health. For example, perhaps Gold medalists have longer athletic careers, and therefore delay or avoid the pursuit of high-paying professional careers that require advanced training or time investments. Such behavior is not perfectly observed, but using data on historical performances of each athlete can be used to assess the possibility of selection. I compile data from the top 100 performances globally in each event and year, as obtained from the site Track and Field Statistics. The data is complete for events going back to 1911, with the exception of the racewalk. Using this comprehensive list of historical performances, I test whether (1) Gold medalists competed for the same number of years as Silver medalists, and (2) the best performances of Gold medalists was equal to that of Silver medalists prior to the Olympics.

The length of each athlete's career is measured as the number of years in which they appear in a top 100 list. Though it is possible an athlete may have competed for longer and been ranked outside the top 100 in a given year, this metric captures the number of years when their performances ranked at a reasonably high level. Gold medalists compete for an average of 6.0 years while Silver medalists compete for an average of 6.3 years, and this difference is not statistically significant. There are also not statistically significant differences

counts: 1896: 81, 1900: 36, 1904: 201, 1908: 204, 1912: 533, 1920: 323, 1924: 1,170, 1928: 1,190, 1932: 1,490, 1936: 1,450, 1948: 695. It is not clear why the number of articles drops off in 1948, but one possibility is greater coverage on television and radio. There is a similar pattern in coverage using nationwide results from the website newspaperarchive.com.

between the career lengths of Gold medalists and 3rd or 4th place finishers. Winners do not appear to be competing for longer periods of time than losers.

To compare whether whether the performances of winners and losers were similar prior to the Olympics, I construct a ranking of the top performers in the 24 months prior to the date of the opening ceremonies of that particular Olympics. A 24 month window provides a long enough window to rank all athletes while still capturing performances relatively close in time to the Olympic Games. Using an 18 month or 36 month window yields similar results.[17] I rank unique athletes, not performances, so that only the best performance of an athlete counts towards the ranking.[18] The assumption is that an athlete's expected finish is based on him running, jumping, or throwing his best in recent years and all other competitors doing the same.[19]

Pre-Olympic performances were similar between Gold and Silver medalists matched to the same event and year. In 52 percent of cases, Gold medalists were ranked higher than Silver medalists leading up to the Olympics. The other 48 percent of the time, Silver medalists ranked higher than Gold medalists. This 50-50 split provides support to Gold and Silver medalists being comparable in terms of prior athletic training and fitness, on average, and to the role of chance in assigning Olympic victory. Table VII presents statistics on pre-Olympic rankings for Gold and Silver medalists, with matched pairs in the same event and year in Columns 1 and 2 and all athletes in Columns 3 and 4. Gold medalists tended to post slightly better performances prior to the Olympics. The median pre-Olympic ranking of Gold medalists was 2nd compared to 3rd for Silver medalists. Twenty percent of both Gold and Silver medalists were ranked first prior to the Olympics, and half of each group is

[17]All Gold and Silver medalists appear in this data source, and all but one have recorded performances ranking in the top 100 prior to the Olympic Games.

[18]In calculating the pre-Olympic rankings for the 100 meter and 1500 meter runs, I also consider times posted in the 100 yard and mile runs, respectively, since the distances are extremely close. I subtract 18 seconds from mile times to convert to 1500 meter times and multiply 100 yard times by 1.1 to convert to 100 meter times. These conversions are consistent with the scoring metrics of the International Association of Athletics Federations.

[19]I only observe the top 100 performances by event and year, rather than the full history of each athlete's performances. With the full history, another approach would be to construct distributions of expected finish.

ranked in the top 3. But more Gold medalists were previously ranked in the top 5 and top 10 than were Silver medalists. More Silver medalists came from further down the ranking distribution. The average ranking among Gold medalists was 7th prior to the Olympics versus 11th for Silver medalists. These average differences are larger when also including unmatched athletes—Gold medalists without a matched Silver medalist and vice versa—as shown in Columns 3 and 4.

Although there is some evidence that Silver medalists had worse performances prior to the Olympics, these differences in pre-Olympic rank do not explain the variation between winning and lifespan. Figure II plots lifespan against pre-Olympic rank for all Gold and Silver medalists. There is a clustering of athletes ranking within the top 10 and with wide variation in lifespan. The minority of athletes ranked outside the top 25 prior to the Olympics are more likely to earn Silver than Gold, and also tend to die at older ages—possible evidence of selection. Table VII presents Cox regressions that include pre-Olympic rank as an additional control variable. Column 1 reports the specification without pre-Olympic rank for reference.[20] The magnitude of the coefficient estimate on losing changes little and remains statistically significant upon adding the athlete's pre-Olympic rank in Column 2. Consistent with the scatterplot, the estimated hazard of death is smaller for lower-ranked athletes, driven by those who were ranked outside the top 25. If the sample is restricted to the majority of athletes ranked within the top 25, pre-Olympic rank does not predict lifespan and the coefficient estimate on losing retains its magnitude and statistical significance. The athlete's pre-Olympic rank also adds modest explanatory power. The third row from the bottom of Table VII presents the share of explained variation, similar to an R^2 from a linear regression, as developed by Royston (2006).[21] In the baseline model, 19.7 percent of the variation in lifespan is explained by finishing place and other observables. Including the pre-Olympic rank increases this share to 24.2 percent. If pre-Olympic ranking represents a

[20]This is a subsample of that presented in Table II because rankings are not available for all events in all years.

[21]As Royston (2006) describes, this statistic is a modification of that proposed by Nagelkerke (1991) based on the likelihood ratio statistic.

measure of physical health or effort invested in training before the Olympics, then winners and losers largely appear similar along this dimension, with the exception of a minority of Silver medalists.

Another interpretation to the patterns between pre-Olympic ranking and lifespan relates to how performance compares to expectations. An athlete's pre-Olympic rank may serve as a reference point in this setting given the objective nature of competition. Studies in psychology have examined the facial expressions of Olympic medalists as shown on television to study their reaction soon after the event, arguing that an athlete's (ex ante) expectations affect their perception of their actual performance ex post (Medvec et al., 1995; McGraw et al., 2005). If pre-Olympic rank is taken as a measure of expected finish, then some Silver medalists from the lower-end of the distribution of rankings greatly out-perform expectations, whereas Gold medalists were ranked higher, on average, before the Olympics. How performance compares to a reference point is central to economic models of utility based on success relative to a performance benchmark (Rayo and Becker 2007a,b) and expectations-based reference dependence (Koszegi and Rabin, 2006). Yet in using data on pre-Olympic performance, there is no apparent way to distinguish expectations from effort spent training. Regardless, the correlation between losing and a longer lifespan does not appear to be explained by (1) how Olympic performance compares to expectations, or by (2) selection in which winners invest more time and effort in training.

VII. Mechanisms

This section investigates potential mechanisms driving the correlation between Olympic finishing place and mortality. The key empirical challenge is that many important life decisions that occur after the Olympics are unobserved. I focus on studying income and occupational choices of U.S. athletes who can be observed in Census records. I find suggestive evidence that income may explain the earlier death of winners. Losers earned more money than winners and, within the sample, income is highly correlated with lifespan. Differences in income explain the variation between winning and mortality: after controlling for income, the estimated mortality hazard does not significantly differ between winners and losers. Including income also explains roughly twice as much variation in lifespan as regressions without it. Taken together, these tests suggest that real resources may explain the earlier death of winners in this population.

A. Income and Occupational Choices

Real resources like income represent a potentially important mechanism between status and health, either through earnings from winning or by influencing future motivation. The institutional features of this setting make income earned as a direct result of competition limited. Amateurism prevailed until the 1980s and these regulations were strictly enforced, as evidenced by Jim Thorpe—the legendary multi-sport athlete—being stripped of his 1912 Olympic Gold medals for earning money to play minor league baseball in 1909 and 1910 (Flatters, 2000). Most athletes held other occupations while training between Olympic Games.[22] An illuminating account of what could be expected financially after the Olympics comes from the autobiography of Mel Sheppard, a Gold medalist in the 1908 Games. Sheppard describes the parting words he and his Track and Field teammates received from President Theodore Roosevelt after returning from the Olympics during a visit to the White House: "I'm going to give you lads the same friendly bit of advice I gave to my Rough Riders. Remember you're heroes for ten days—when that time's up, drop the hero business and go to work" (Sheppard, 1924, p52). The Gold medal itself was worth a modest amount in terms of its metallic content.[23]

[22] For example, Hannes Koheleman—a Gold medalist distance runner—laid bricks in construction (see The New York Times, "Hannes Kolehmainen, Marathon Champion, is Now U.S. Citizen," January 15, 1921) and Charlie Paddock—a Silver medalist sprinter—worked for a newspaper (see Dallas Morning News, "Obituary: Paddock, Charles William." July 23, 1943).

[23] Before 1912, the gold in the winner's medal was worth about $350 adjusting for inflation and the commodity prices of the year it was awarded (The Economist, 2012). After 1912, gold was no longer used

While financial rewards from competition were limited, athletes may have pursued various occupations after their athletic career ended, and income from these life decisions may be important to health. Occupational choices after the Olympics help shed light on whether losing serves to motivate. To study this channel, I collect data from the 1940 Census for the sub-sample of U.S. athletes competing between 1920 and 1936. The 1940 Census was the first to record income. Specifically, the survey records annual wage income in 1939 as well as whether the respondent received any supplemental income from other sources. In addition, the 1940 Census also records information on occupation, home ownership, labor supply, race, marital status, and education. The individual records of each Census respondent become publicly available 72 years after the survey, enabling me to observe these variables for my sample of U.S. athletes. I use the geneology site Ancestry.com to track athletes and report the details of the procedure to retrieve individual records in Appendix C.

Data on occupational choices and average annual earnings by occupation is presented in Table IX. For comparison, the first two columns list the percentage of each occupational category and the average earnings for all U.S. males aged 20 and older in the labor force, respectively. These statistics are tabulated from the 100 percent sample of the 1940 Census made available through the Integrated Public Use Microdata Series (IPUMS). The IPUMS sample, which excludes any identifying information but is useful to gauge broader trends, constructs a measure of average earnings by 3-digit occupation code. Table IX then aggregates these codes up to broader occupational categories following the IPUMS census classification. At the high end of the earnings distribution, professional workers earned the most money, followed by proprietors, managers, and officials. At the bottom of the income distribution, farm laborers and domestic services earned the least.

The corresponding statistics for Gold and Silver medalists are presented in Columns 3 to 6 of Table IX. While both groups entered occupations that earned substantially more than the U.S. average, Silver medalists chose occupations that paid more than occupations chosen

and the winner's medal was made mostly of silver and copper, making it worth even less.

by Gold medalists. The large majority of Silver medalists were classified as Professional Workers and entered occupations, including physicians, with particularly high earnings. Gold medalists were more likely to be classified as Proprietors, Managers, and Officials. A common occupation among Gold medalists was athletic coach, which the Census classifies as a Semi-Professional Worker. In my sample, the average earnings of Silver medalists were 16 percent higher than Gold medalists based on differences in occupational choices.

Income differences also largely explain the variation between lifespan and losing. Table X reports results of Cox regressions that include variables for income, labor supply, home ownership, and demographics, as collected from the Census for each athlete in the U.S. sub-sample. Without controlling for income, losing is again correlated with a lower mortality hazard as shown in Column 1. This specification also includes year and event class effects, height, year of birth, and indicators for White and married. The coefficient estimate on losing is large in magnitude and statistically significant, similar to the results from the full sample. Including income in Column 2 drives the estimate on losing closer to 1, and it is no longer statistically significant. By contrast, higher income correlates with a lower mortality hazard. The coefficient estimates on both income variables are large and statistically significant. Including income also explains 57.8 percent of the variation in lifespan as reported by the modified R^2 statistic, compared to just 30.8 percent with finishing place and other observables in Column 1. As shown in Column 3, which includes the number of weeks worked and excludes income, labor supply alone explains far less of the variation in lifespan. Column 4 includes both income and labor supply and demonstrates that income is a strong predictor of mortality. In this sub-sample, the association between winning and an earlier death is explained by higher income among Silver medalists. [24] Including ranking data does not alter these

[24]By comparison, other estimates of the role of income on mortality vary widely based on age and the type of data, ranging from zero to roughly twice as large as my estimates (Smith, 1999; Deaton and Paxson, 2001; Deaton, 2003; Cutler et al., 2011)

Income and occupations of parents

While losers earned more than winners after the Olympics, it is possible that an athlete's occupational choices and earnings were influenced by that of his parents. This section tests whether parental earnings differ systematically across Silver and Gold medalists, which would constitute a source of omitted variable bias if living standards in childhood are correlated with mortality. To study each athlete's family history, I collect their parent's occupations recorded in the 1910, 1920, and 1930 Censuses, when the athletes were in childhood.[25] I impute earnings for their parents' occupations based on the IPUMS occupation codes from the 1940 Census. In the few cases in which a parent held multiple occupations over different waves of the Census, I calculate the average of the two earnings estimates. As a second measure of parental earnings, I also assign the industry-specific average earnings in 1925, which is closer to the time period of Census surveys, collected from Margo (2006). In both cases, parental earnings are not statistically different between Gold and Silver medalists based on a simple t-test (with the parents of Gold medalists having slightly higher earnings). Failing to reject the null hypothesis that parental earnings are equal between winners and losers can be interpreted as another test of balance between the two groups, now on childhood economic conditions.

Table XI presents results from Cox regressions that include parental earnings along with the individual athlete's income and other Census variables. Since hazard models of mortality cannot include individual fixed effects, these specifications use information about parental earnings to control for unobservables at the family level. Columns 1 to 4 use imputed parental earnings based on 1925 industry occupations and Columns 5 to 8 use imputed parental earnings based on 1940 occupational codes. Column 1 includes the athlete's income and that of his parents in logs and Column 2 in levels. The coefficient estimates on parental earnings are not statistically significant while those on the athlete's income once again are. The specifications in Columns 3 and 4 includes the difference between athlete's income and

[25]More specifically, I record the occupation of the household head, which was the father in most cases.

parental income. This measure of the within-family difference in income is positively related to the athlete's lifespan and highly significant. The results are similar when parental income is imputed using 1940 occupation codes in Columns 5 to 8. Taken together, the analysis of Census records are consistent with idea that relative rank influences motivation. The data does not allow me to distinguish whether losing motivates or winning de-motivates, however.

B. Competing explanations

News coverage

Even though amateurism prevented athletes being directly compensated for their performance, it is possible that athletes received non-monetary rewards, like housing or job opportunities, that could have first-order effects on longevity. To study this mechanism, I collect text-based data on newspaper coverage of each athlete from the website newspaperarchive.com.[26] I focus on U.S. athletes among the Census sub-sample because the site mainly includes U.S. newspapers and to compare the importance of news coverage against income income as a mechanism. For each athlete, I search for stories containing their first and last name, the word "Olympics", and the year and event they participated in.[27] I record the number of news stories within two decades of the Olympic Games the athlete competed in. The rationale for restricting coverage to this period is that any changes to living standards as a result of Olympic performance are likely to be reflected in coverage closer to the competition. For example, there was very little newspaper coverage of athletes competing in the first few Olympic Games, but much more coverage in the 1960s and later after most were deceased.[28]

Not surprsingly, winners receive substantially more news coverage than losing ath-

[26]Other research on news coverage has also used data from this source (Gentzkow et al., 2011)

[27]In case the athlete is known primarily by his nickname, I also include searches that replace the athlete's first name with the nickname reported on sportsreference.com. In addition, since the long jump was historically called the "broad jump" during my sample, I search for this term in that event.

[28]The post-1950s coverage of athletes competing in the first modern Olympic Games tends to recount the experience of these early athletes to establish the history of the Games.

lete—between 2 and 3 times as much. Yet news coverage is not correlated with lifespan. Table XII reports hazard regressions that include newspaper coverage variables estimated on the U.S. sub-sample with Census records. Columns 1 and 2 present Cox regressions with variables for news coverage along with finishing place, with Column 1 including the count of stories and Column 2 including an indicator for over 50 stories to allow for non-linearity. Losing has a strong and statistically significant association with lifespan, while variables for news coverage do not. Columns 3 through 4 include variables for news coverage and income, omitting finishing place (which is positively correlated with coverage). As before, higher income is associated with lower hazards of death, and the coefficient estimates on news coverage are not statistically significant. Columns 5 and 6 include variables for finishing place, news coverage, and income. These regressions do not suggest that media exposure influences mortality.

Mean reversion

The correlation between finishing place and mortality is unlikely to be driven by regression to the mean. Due to variability in performance, those ranked below the mean before the Olympics are likely to rank closer to the mean in the Olympic final. The winner in the Olympics may have previously been ranked closer to the mean, finding himself the fortunate recipient of good luck on the day that counts. As discussed earlier, such variation in performance is key to the identification strategy. For mean reversion to explain the relationship between losing and lifespan, however, performance would have to correlate contemporaneously with mortality risk. It seems possible that performance could correlate with other measures of current health, such as resting heart rate or VO_2max, but unlikely it would correlate with long-term health outcomes like mortality.

VIII. Conclusion

This paper has compared the longevity of Olympic Track and Field athletes to investigate how competition for status influences health. Counterintuitively, Silver medalists live over two years longer than Gold medalists, on average. The institutional features of the Olympic setting—though highly stylized—allow status to be cleanly identified. The sharp cutoff between winning and losing in the Olympic final helps to reduce possible unobserved heterogeneity between athletes. I also demonstrate that winners and losers are balanced in terms of observables like height, ability, and age, and that awarding status involves a large degree of randomness, likely because it is a physical contest held on a single day every four years. Specific features of the setting are also instrumental in isolating different channels between status and health. In particular, there is less concern of reverse causality here than other contests, at least during the period before the rise of performance-enhancing drugs that I study. The prevailing system of amateurism also prevented any compensation to be earned directly from competition. There also does not appear to be evidence of selection: Gold medalists do not have longer athletic careers and pre-Olympic rankings were often similar to Silver medalists.

Using individual Census records of each athlete and his family, I find empirical support for occupational choices and income earned after Olympic competition as a potential channel between status and health. Silver medalists pursued occupations that paid more money than those chosen by Gold medalists. Income, reported in the 1940 Census, is positively correlated with lifespan in the sample and fully accounts for the relationship between losing and mortality. To test whether childhood economic conditions were similar between Gold and Silver medalists, I link athletes to their family's earlier records in 1910, 1920, and 1930 Censuses and impute parental incomes based on their occupations. The failure to reject the null hypothesis that parental incomes were equal between Gold and Silver medalists provides another test of balance. Based on a range of specifications and estimation methods, it is unlikely that omitted variable bias explains these patterns. Since amateurism limited any income earned from competition, the analysis of Census records is consistent with relative rank influencing motivation. Data limitations prevent me from determining whether losing motivates or winning de-motivates, however. It is important to note that if Gold medalists enjoyed more income-related opportunities from winning than Silver medalists, such benefits should reduce Gold medalists' mortality risks, not increase them.

There are several limitations of the study. In terms of data, many decisions and shocks after the Olympics are unobserved. The Census records add key information about occupational choices, income, and marital status, but lifestyle factors are likely to matter as well. For example, decisions about smoking and alcohol consumption would be informative to study behavioral responses to winning and losing. Second, data on each athlete's income after the Olympics is measured only once in 1940. One would ideally observe multiple years of earnings, although lifetime earnings are clearly endogenous. For each athlete's parents, income must also be imputed in the pre-1940 Censuses based on occupation since it is not collected in the survey. Third, the sample size is small, especially on the U.S. sub-sample with Census records. A larger sample could help to improve precision of the estimates and increase confidence in the results. Finally, the setting of Olympic Track and Field raises questions about external validity. This study's findings may be applicable more broadly insofar as the pivotal events in people's lives resemble such competition.

Despite these limitations, this paper's findings challenge conventional wisdom and the conclusions from existing studies that being awarded higher status necessarily improves health (Marmot et al. 1978, 1991; Sylvestre, Huszti and Hanley 2006; Becker, Chay and Swaminathan 2007; Rablen and Oswald 2008). Instead, losing can have positive, first-order effects on longevity. The most important trials in our lives often involve a binary outcome, like victory or defeat. This paper's results suggest how people respond to such successes or failures can produce long-lasting consequences for health.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson, "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration," *American Economic Review*, 2012, *102* (5), 1832–1856.

_ , _ , and _ , "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 2014, *122* (3), 467–506.

Barankay, Iwan, "Rank Incentives: Evidence from a Randomized Workplace Experiment," 2012.

_ , "Rankings and Social Tournaments: Evidence from a Crowd-Sourcing Experiment.," 2012.

Becker, David, Kenneth Chay, and Shailender Swaminathan, "Mortality and the Baseball Hall of Fame: An Investigation into the Role of Status in Life Expectancy," *iHEA World Conference*, 2007.

Bleakley, Hoyt and Joe Ferrie, "Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations," *Quarterly Journal of Economics*, 2016.

Case, Anne and Christina Paxson, "The Long Reach of Childhood Health and Circumstance: Evidence from the Whitehall II Study," *Economic Journal*, 2011, *121* (554), F183–F204.

Chandra, Amitabh and Tom Vogl, "Rising up with shoe leather? A comment on Fair Society, Healthy Lives (The Marmot Review)," *Social Science & Medicine*, 2010, *71*, 1227–1230.

Clark, Andrew and Andrew Oswald, "Satisfaction and Comparison Income," *Journal of Public Economics*, 1996, *61* (3), 359–381.

Clarke, Philip, Simon Walter, Andrew Hayen, William Mallon, Jeroen Heijmans, and David Studdert, "Survival of the fittest: retrospective cohort study of the longevity of Olympic medallists in the modern era," *BMJ*, December 2012, *345*, e8308–e8308.

Cutler, David, Adriana Lleras-Muney, and Tom Vogl, "Socioeconomic Status and Health: Dimensions and Mechanisms," in Sherry Glied and Smith, eds., *The Oxford Handbook of Health Economics*, Oxford University Press, 2011.

Deaton, Angus, "Health, Inequality, and Economic Development," *Journal of Economic Literature*, 2003, *41*, 113–158.

_ and Christina Paxson, "Mortality, Education, Income, and Inequality among American Cohorts," in David Wise, ed., *Themes in the Economics of Aging*, University of Chicago Press, 2001, pp. 129–170.

Easterlin, Richard, "Will Raising the Incomes of All Increase the Happiness of All?," *Journal of Economic Behavior and Organization*, 1995, *27*, 35–47.

Economist, "All that glisters: How much is an Olympic gold medal worth?," *The Economist*, April 2012.

Epstein, C.R. and R.J. Epstein, "Death in The New York Times: The Price of Fame is a Faster Flame," *QJM*, 2013, *106* (6), 517–521.

Falk, Armin and Marcus Knell, "Choosing the Joneses: Endogenous Goals and Reference Standards," *Scandinavian Journal of Economics*, 2004, *106* (3), 417–435.

Flatters, Ron, *Thorpe Preceded Deion, Bo*, ESPN, Sports Century, 2000.

Frank, Robert, *Choosing the Right Pond*, New York: Oxford University Press, 1985.

Frijters, Paul, John Haisken-DeNew, and Michael Shields, "The Increasingly Mixed Proportaional Hazards Model: An Application to Socioeconomic Status, Health Shocks, and Mortality," *Journal of Business and Economic Statistics*, 2011, *29* (2), 271–281.

Gentzkow, Matthew, Matthew Shapiro, and Michael Sinkinson, "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 2011, *101*, 2980–3018.

Heffetz, Ori, "A Test of Conspicuous Consumption: Visibility and Income Elasticities," *Review of Economics and Statistics*, 2011, *93* (4), 1101–1117.

— **and Robert Frank**, "Preferences for Status: Evidence and Economic Implications," in Jess Benhabib, Alberto Bisin, and Matthew Jackson, eds., *Handbook of Social Economics*, Vol. 1A, The Netherlands: Elsevier, 2011, pp. 69–91.

Koszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

Luttmer, Erzo, "Neighbors as Negatives: Relative Earnings and Well-Being," *Quarterly Journal of Economics*, 2005, *120* (3), 963–1002.

Margo, Robert, "Wages and Wage Inequality," in Susan Carter, Sigmund Gartner, Michael Haines, Alan Olmstead, Richard Sutch, and Gavin Wright, eds., *Historical Statistics of the United States*, New York: Cambridge University Press, 2006.

Marmot, M and A Feeney, "General explanations for social inequalities in health," *IARC Scientific Publications*, 1997, (138), 207–228.

Marmot, M G, G D Smith, S Stansfeld, C Patel, F North, J Head, I White, E Brunner, and A Feeney, "Health inequalities among British civil servants: the Whitehall II study," *Lancet*, June 1991, *337* (8754), 1387–1393.

— **, G Rose, M Shipley, and P J Hamilton**, "Employment grade and coronary heart disease in British civil servants.," *Journal of Epidemiology and Community Health*, December 1978, *32* (4), 244–249.

Marmot, M, M Shipley, E Brunner, and H Hemingway, "Relative contribution of early life and adult socioeconomic factors to adult morbidity in the Whitehall II study," *Journal of epidemiology and community health*, May 2001, *55* (5), 301–307.

Marmot, Michael and Eric Brunner, "Cohort Profile: The Whitehall II Study," *International Journal of Epidemiology*, 2005, *34*, 251–256.

McGraw, A. Peter, Barbara Mellers, and Philip Tetlock, "Expectations and Emotions of Olympic Athletes," *Journal of Experimental Social Psychology*, 2005, *41*, 438–446.

Medvec, Victoria, Scott Madey, and Thomas Gilovich, "When Less is More: Counterfactual Thinking and Satisfaction among Olympic Medalists," *Journal of Personality and Social Psychology*, 1995, *69* (4), 603–10.

Nagelkerke, Nico J.D., "A Note on a General Definition of the Coefficient of Determination," *Biometrika*, 1991, *78* (3), 691–692.

Olshansky, S. Jay and Bruce Carnes, "Ever Since Gompertz," *Demography*, 1997, *34* (1), 1–15.

Perrottet, Tony, *The Naked Olympics*, New York: Random House, 2004.

Rablen, Matthew and Andrew Oswald, "Mortality and immortality: The Nobel Prize as an experiment into the effect of status upon longevity," *Journal of Health Economics*, 2008, *27* (6), 1462–1471.

Rayo, Luis and Gary Becker, "Evolutionary Efficiency and Happiness," *Journal of Political Economy*, 2007, *115* (2), 302–336.

— **and** — , "Habits, Peers, and Happiness: An Evolutionary Perspective," *American Economic Review Papers & Proceedings*, 2007, *97* (2), 487–491.

Royston, Patrick, "Explained Variation for Survival Models," *Stata Journal*, 2006, *6* (1), 83–96.

Samaras, Thomas, "How height is related to our health and longevity: A review," *Nutrition and Health*, 2012, *21* (4), 247–261.

Sapolsky, Robert, "The Influence of Social Hierarchy on Primate Health," *Science*, 2005, *308*, 648–652.

Sheppard, Mel, "Spiked Shoes and Cinder Paths: An Athlete's Story," in "Sport Story Magazine" 1924.

Smith, James, "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status," *Journal of Economic Perspectives*, 1999, *13* (2), 145–166.

Sylvestre, Marie-Pierre, Ella Huszti, and James A Hanley, "Do OSCAR winners live longer than less successful peers? A reanalysis of the evidence," *Annals of internal medicine*, 2006, *145* (5), 361–363.

Tran, Anh and Richard Zeckhauser, "Rank as an inherent incentive: Evidence from a field experiment," *Journal of Public Economics*, 2012, *96* (9), 645–650.

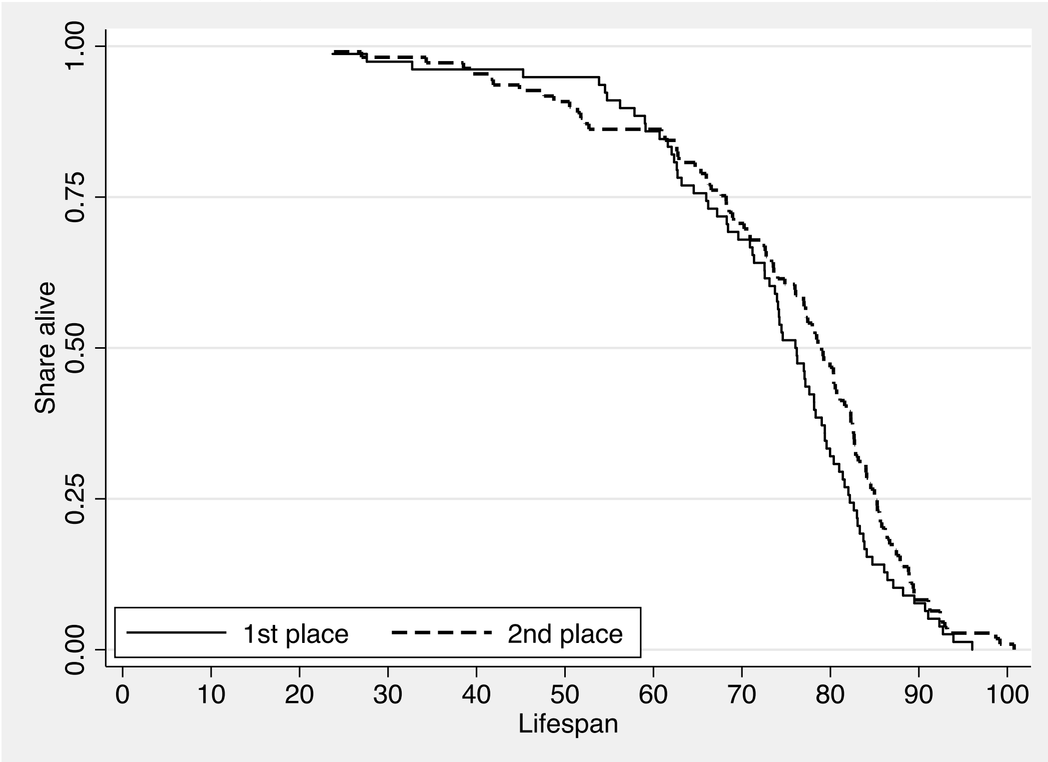
Tung, Jenny, Luis Barreiro, Zachary Johnson, Kasper Hansen, Vasiliki Michopoulos, Donna Toufexis, Katelyn Michelini, Mark Wilson, and Yoav Gilad, "Social Environment is Associated with Gene Regulatory Variation in the Rhesus Macaque Immune System," *Proceedings of the National Academy of Sciences*, 2012, *109* (17), 6490–6495.

Vaananen, Ari, Aki Koskinen, Matti Joensuu, Mika Kivimaki, Jussi Vahtera, Anne Kouvonen, and Paavo Jappinen, "Lack of Predictability at Work and Risk of Acute Myocardial Infarction: An 18-Year Prospective Study of Industrial Employees," *American Journal of Public Health*, 2012, *98* (12), 2264–2271.

WADA, World Anti-Doping Agency, *A Brief History of Anti-Doping* 2010.

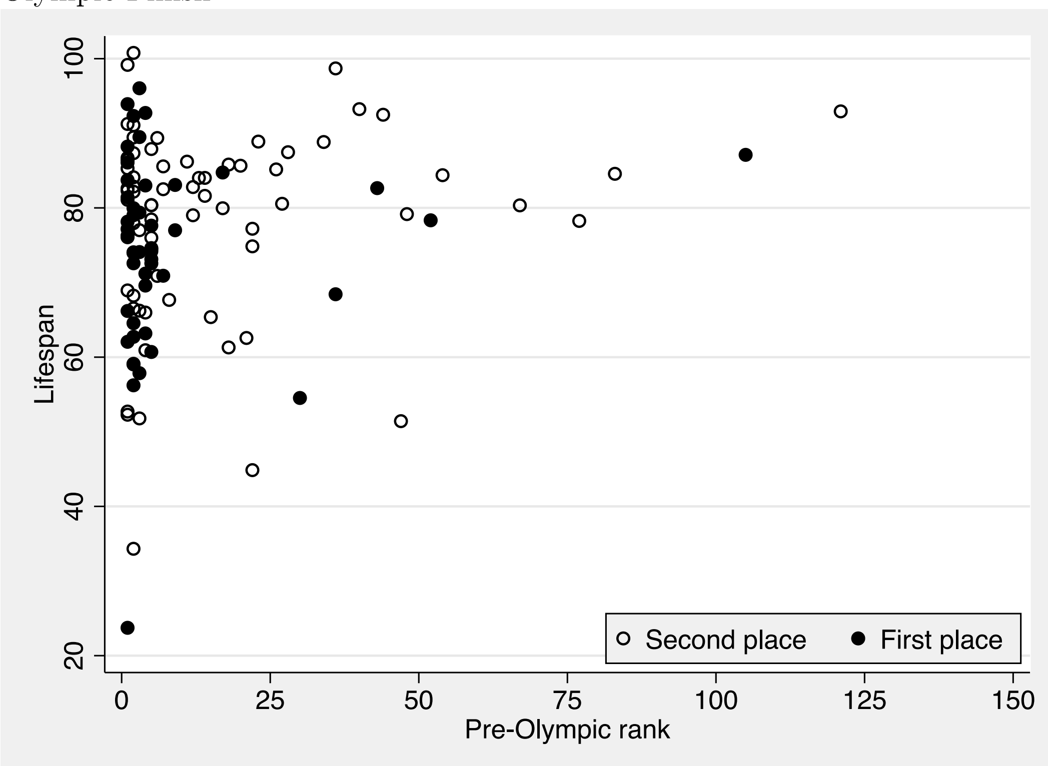
Wadler, Gary, *Doping in Sport: From Strychnine to Genetic Enhancement, It's a Moving Target* 1998.

Figure I: Kaplan-Meier Survival Curves by Finishing Place



Note: This figure plots the proportion of athletes still alive at each age by Gold or Silver medal status. The difference in area between the two curves is equivalent to the difference in lifespan. On average, Silver medalists live 1.2 years longer than Gold medalists based on the raw data plotted here.

Figure II: Scatterplot of Lifespan Against Difference Between pre-Olympic Ranking and Olympic Finish



Note: This scatterplot displays the lifespan for each observation in the sample on the vertical axis. The horizontal difference plots the athlete's ranking based on pre-Olympic performance. Silver medalists are indicated by hollow circles and Gold medalists are indicated by solid circles. While most Gold and Silver medalists were ranked within the top 5, there are several Silver medalists who were ranked well outside the top 25, and lifespan appears to be higher for such athletes.

Table I: Lifespan Regressions with Event and Year Fixed Effects

Variable	Mean	s.d.	Min	Max	N
Lifespan	73.99	14.91	23.73	100.77	187
Year of birth	1898.43	14.84	1869	1926	187
World Record holder	0.03	0.18	0	1	187
Height (cm)	180.51	7.13	160	195	146
Weight (kg)	74.57	11.87	51	110	143
Distance event	0.16	0.37	0	1	187
Middle-distance event	0.10	0.30	0	1	187
Sprints event	0.22	0.42	0	1	187
Field event	0.32	0.47	0	1	187
Throwing event	0.16	0.36	0	1	187

Note: This table displays statistics on lifespan and various observables for Gold and Silver medalists.

Table II: Balance Tests

Variable	1st place mean (N=78)	2nd place mean (N=109)	Difference	p-value of difference
Year of birth	1898.0	1898.7	-0.69	0.755
World Record holder	0.03	0.04	-0.01	0.674
Height (cm)	181.28	179.89	1.39	0.244
Weight (kg)	75.32	73.95	1.38	0.492
Distance event	0.17	0.16	0.01	0.844
Middle-distance event	0.09	0.11	-0.02	0.650
Sprints event	0.24	0.21	0.03	0.599
Field event	0.35	0.29	0.05	0.446
Throwing event	0.13	0.17	-0.05	0.390

Note: This table displays means of year of birth, height, weight, the types of events, and the fraction of World Record holders for Gold and Silver medalists. The final column presents the p-value from the t-test that the means of the corresponding variable are equal between Gold and Silver medalists. There are not statistically significant differences in these observables between the two groups of athletes. Height data is available for 146 athletes. Weight data is available for 143 athletes. The other variables are available for 187 athletes.

Table III: Lifespan Regressions with Event and Year Fixed Effects

	(1) Cox	(2) Cox	(3) MPH	(4) MPH	(5) Gompertz	(6) Gompertz
Lose (1=yes, 0=no)	0.714* (-1.87)	0.719* (-1.84)	0.730* (-1.76)	0.736* (-1.71)	0.730* (-1.76)	0.736* (-1.71)
Year of birth	0.989 (-0.34)	0.990 (-0.30)	0.996 (-0.13)	0.997 (-0.09)	0.996 (-0.13)	0.997 (-0.09)
Ever set World Record (1=yes, 0=no)		0.802 (-0.42)		0.804 (-0.41)		0.804 (-0.41)
Year effects	Yes	Yes			Yes	Yes
Individual event effects	Yes	Yes			Yes	Yes
Country effects	No	No	No	No	No	No
Event class effects	No	No	No	No	No	No
Frailty	None	None	Individual	Individual	None	None
Observations	187	187	187	187	187	187
Log likelihood	-763.06	-762.98	82.77	82.86	82.77	82.86

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival model regressions. Losing is defined as finishing in second place. Columns 1 and 2 estimate Cox models. Columns 3 and 4 estimate Mixed Proportional Hazards (MPH) models that assume the hazard follows a Gompertz distribution and allows for individual heterogeneity. Columns 5 and 6 estimates survival models that assume a Gompertz distribution for the hazard without allowing for individual heterogeneity. The coefficient estimate below 1 on losing indicates the hazard of death is lower among Silver medalists than Gold medalists. Robust t-statistics clustered by event-year in parentheses, except in models with individual heterogeneity (frailty). *p<0.1, **p<0.05, ***p<0.01.

Table IV: Survival Regressions Including Other Olympic Finalists

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cox	Cox	MPH	MPH	Cox	Cox	Gompertz	Gompertz
	Sample: 1st vs. 2nd - 4th places				Sample: 1st vs. 2nd - 8th places			
Lose (1=yes, 0=no)	0.740**	0.740**	0.763*	0.763*	0.754**	0.754**	0.779**	0.778**
	(-2.40)	(-2.40)	(-1.94)	(-1.93)	(-2.48)	(-2.48)	(-2.21)	(-2.21)
Year of birth	1.007	1.008	1.009	1.009	1.013	1.013	1.014	1.014
	(0.37)	(0.41)	(0.47)	(0.48)	(1.01)	(1.02)	(1.12)	(1.10)
Ever set World Record		0.870		0.960		0.967		1.138
(1=yes, 0=no)		(-0.25)		(-0.10)		(-0.06)		(0.31)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual event effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country effects	No	No	No	No	No	No	No	No
Event class effects	No	No	No	No	No	No	No	No
Frailty	None	None	Individual	Individual	None	None	None	None
Observations	395	395	395	395	658	658	658	658
Log likelihood	-1938.4	-1938.4	131.0	131.0	-3567.8	-3567.8	185.5	185.5

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival regressions that include other finalists. Columns 1 to 4 present results that compare Gold medalists to places 2 through 4 and Columns 5 through 8 compare Gold medalists to places 2 through 8. MPH models failed to converge when including all finalists, and so columns 7 and 8 present results from Gompertz regressions for comparison. In all models, the coefficient estimates are similar to the main results presented in Tables 3. Robust t-statistics clustered by event-year in parentheses, except in MPH models with individual heterogeneity. *p<0.1, **p<0.05, ***p<0.01.

Table V: Lifespan Regressions with Country Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: 1st vs. 2nd places				Sample: 1st vs. 2nd - 4th places			
	Cox	Cox	MPH	MPH	Cox	Cox	Gompertz	Gompertz
Lose (1=yes, 0=no)	0.760	0.720*	0.759	0.718*	0.741**	0.712**	0.759**	0.729**
	(-1.60)	(-1.69)	(-1.62)	(-1.73)	(-2.43)	(-2.42)	(-2.34)	(-2.39)
Year of birth	0.977***	0.979**	0.978***	0.979**	0.983***	0.985***	0.983***	0.986***
	(-2.94)	(-2.42)	(-3.25)	(-2.40)	(-4.07)	(-2.75)	(-4.15)	(-2.78)
Height (cm)		1.016		1.012		1.012		1.010
		(1.00)		(0.83)		(1.23)		(1.13)
Year effects	No	No	No	No	No	No	No	No
Individual event effects	No	No	No	No	No	No	No	No
Country effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event class effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Frailty	None	None	Individual	Individual	None	None	None	None
Observations	187	146	187	146	395	294	395	294
Log likelihood	-773.9	-566.4	72.3	55.8	-1940.5	-1363.1	130.1	96.4

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival model regressions with country fixed effects. All regressions also include indicators for event classes (sprints, middle distance, distance, throws, field, and racewalk), rather than individual events as in Table 2. Appendix Table A.2 lists which individual events are grouped into one of the six event classes. Similar to Table 2, the coefficient estimate below 1 on losing indicates the hazard of death is lower among Silver medalists than Gold medalists. Robust *t*-statistics clustered by event-year in parentheses, except in shared frailty models. *p<0.1, **p<0.05, ***p<0.01.

Table VI: Cox Regressions by Year of Olympic Games

	(1)	(2)	(3)	(4)
	Cox	MPH	Cox	Gompertz
	Years: 1896-1924		Years: 1928-1948	
Lose (1=yes, 0=no)	0.932	0.916	0.471***	0.447***
	(-0.22)	(-0.30)	(-2.95)	(-3.12)
Year of birth	1.026	1.032	0.909**	0.922*
	(0.35)	(0.50)	(-2.25)	(-1.67)
Ever set World Record (1=yes, 0=no)				
Year effects	Yes	Yes	Yes	Yes
Individual event effects	Yes	Yes	Yes	Yes
Country effects	No	No	No	No
Event class effects	No	No	No	No
Observations	86	86	101	101
Log likelihood	-270.4	41.4	-343.6	69.6

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions that split the sample by years 1896 to 1924 (columns 1 and 2) and years 1928 to 1948 (columns 3 and 4). Robust t-statistics clustered by country-year in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table VII: Statistics of Prior Rankings by Finishing Place

	Gold medalists	Silver medalists	Gold medalists	Silver medalists
	Matched pairs only		All athletes	
Median Ranking before Olympics	2	3	3	5.5
Mean Ranking before Olympics	7.1	11.5	7.0	15.9
Top 1 before Olympics (percent)	20.6	20.6	24.4	15.6
Top 3 before Olympics (percent)	55.8	52.9	57.8	39.1
Top 5 before Olympics (percent)	82.3	58.9	82.2	50.0
Top 10 before Olympics (percent)	91.1	64.7	88.9	57.8
Top 25 before Olympics (percent)	91.1	91.1	91.1	82.8

Note: This table displays the average rank and percentage of the sample by their pre-Olympic ranking for Gold and Silver medalists. Columns 1 and 2 include athletes matched to the same event and year ("matched pairs") and Columns 3 and 4 including all Gold and Silver medalists. Rankings are calculated based on the 24 months prior to the opening ceremony of each Olympics.

Table VIII: Cox Regressions with pre-Olympic Rankings

	(1) Subsample: pre-Olympic ranking data available	(2) Subsample: pre-Olympic ranking data available	(3) Subsample: ranked in the top 25 before Olympics
Lose (1=yes, 0=no)	0.515***	0.512***	0.588*
	(-3.00)	(-2.97)	(-1.92)
Pre-Olympic ranking		0.989***	1.025
		(-2.62)	(1.04)
Year of birth	0.953	0.916	0.894*
	(-0.77)	(-1.35)	(-1.68)
Year effects	Yes	Yes	Yes
Individual event effects	Yes	Yes	Yes
Country effects	No	No	No
Event class effects	No	No	No
Modified R^2 based on Royston (2006)	0.197	0.242	0.239
Observations	117	117	98
Log Likelihood	-421.0	-418.7	-334.1

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions that include variables measuring the athlete's pre-Olympic ranking. The first column replicates the main results from Table 2 on the sub-sample with available ranking data, estimating a hazard ratio of death of 0.515 for Silver medalists vs. Gold medalists. Column 2 adds the athlete's pre-Olympic rank. A higher pre-Olympic rank (worse performance) is positively correlated with lifespan, but the effect of losing remains statistically significant and is of a similar magnitude to the result in Column 1. Column 3 restricts the sample to those ranked in the top-25 before the Olympics, and shows no correlation between pre-Olympic ranking and lifespan for this sub-sample . Robust t-statistics clustered by country-year in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table IX: Distribution of Occupations and Average Earnings (1950 US Dollars)

	All U.S. Males aged 20+ in Labor Force		Gold medalists		Silver medalists	
Occupation Category	Percent	Mean Earnings	Percent	Mean Earnings	Percent	Mean Earnings
Professional Workers	4.0	4,340	25.0	4,375	68.0	4,994
Proprietors, Managers, Officials (Except Farm)	8.8	4,090	31.3	4,080	8.0	3,950
Semiprofessional Workers	0.8	3,167	12.5	3,200	4.0	3,200
Craftsmen, Foremen, and Kindred Workers	14.8	3,001	0	-	0	-
Clerical Workers	7.2	2,651	6.3	3,600	8.0	2,500
Salesmen & Saleswomen	5.8	2,547	18.8	2,800	4.0	2,400
Operatives and Kindred Workers	17.0	2,474	0	-	0	-
Protective Service Workers	1.8	2,294	0	-	0	-
Laborers (Except Farm)	13.5	1,968	6.3	2,000	4.0	2,000
Service Workers (Except Domestic and Protective)	4.2	1,696	0	-	0	-
Farmers and Farm Managers	14.1	1,406	0	-	4.0	1,400
Farm Laborers and Foremen	7.5	813	0	-	0	-
Domestic Service Workers	0.5	599	0	-	0	-
Total	100	2,404	100	3,644	100	4,272

Note: This table presents the percent of Gold medalists, Silver medalists, and the male U.S. labor force aged 20 and older by occupational category in 1940 from the U.S. Census. The average earnings by category are constructed from finer 3-digit occupation codes and reported in the IPUMS Census data file. Appendix B provides additional details about the collection and analysis of the Census data. Columns 1 and 2 are tabulated using the 100 percent IPUMS 1940 Census file, which excludes identifying information but provides 3-digit occupation codes and corresponding earnings. Silver medalists pursued higher-paying occupations than Gold medalists after the Olympics.

Table X: Cox Regressions with 1940 U.S. Census Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	U.S. sub-sample: Complete Census data					U.S. sub-sample: Complete Census and pre-Olympics rank data			
Lose (1=yes, 0=no)	0.325*	0.543	0.491	0.643	0.591	0.266**	0.456	0.480	0.445
	(-1.92)	(-0.98)	(-1.09)	(-0.77)	(-0.88)	(-2.12)	(-1.25)	(-1.12)	(-1.27)
Log wage income ($100s), annual		0.646**		0.656**			0.664**	0.684*	
		(-2.50)		(-2.34)			(-2.19)	(-1.89)	
Wage income ($100s), annual					0.962***				0.964**
					(-2.68)				(-2.44)
Income from other sources (1=yes, 0=no)		0.067**		0.085**	0.245**		0.057**	0.074**	0.187**
		(-2.57)		(-2.42)	(-2.07)		(-2.53)	(-2.43)	(-2.18)
Owned home (1=yes, 0=no)		1.864	1.312	1.820	1.723		1.784	1.812	1.705
		(0.91)	(0.51)	(0.90)	(0.82)		(0.81)	(0.81)	(0.78)
Number of weeks worked in 1939			1.066	1.061	1.106*			1.069	1.126**
			(1.25)	(1.31)	(1.84)			(1.27)	(2.27)
White (1=yes, 0=no)	0.818	1.401	1.223	1.896	0.847	0.655	1.760	2.692	1.098
	(-0.39)	(0.76)	(0.22)	(1.10)	(-0.20)	(-0.66)	(0.91)	(1.47)	(0.11)
Married (1=yes, 0=no)	0.349*	0.962	0.570	1.472	2.600	0.683	1.012	2.525	6.342
	(-1.85)	(-0.06)	(-0.59)	(0.57)	(0.92)	(-0.55)	(0.01)	(0.69)	(1.13)
pre-Olympic ranking							0.990	0.981	0.985
							(-0.43)	(-0.61)	(-0.49)
Height (cm)	1.155***	1.201***	1.144***	1.200***	1.186***	1.160***	1.195***	1.189***	1.185***
	(2.79)	(3.40)	(2.75)	(3.32)	(3.40)	(2.69)	(3.21)	(3.06)	(3.15)
Modified R^2 based on Royston (2006)	0.308	0.580	0.351	0.595	0.541	0.284	0.556	0.580	0.544
Observations	39	39	38	38	38	36	36	35	35
Log likelihood	-95.90	-83.52	-90.86	-79.63	-82.47	-86.7	-75.6	-71.4	-73.1

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions for the U.S. sub-sample with data from the 1940 Census. All regressions include year and event class effects as well as year of birth (which is not significant). As in Tables 2 and 3, the coefficient estimate on losing is below 1 and statistically significant in Column 1. Once income and labor supply are included in Columns 2 through 5, the estimate on losing increases closer to 1 and is no longer statistically significant while the variables for income enter significantly. The estimates are robust to including pre-Olympic rank in Columns 6 to 9. Robust t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table XI: Cox Regressions with Parental Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Imputation of parental income based on: 1925 Industry classification				Imputation of parental income based on: 1940 Occupational classification			
	Income measured in:				Income measured in:			
	Logs	Levels	Levels	Levels	Logs	Levels	Levels	Levels
Lose (1=yes, 0=no)	0.520	0.524	0.443	0.456	0.546	0.492	0.468	0.471
	(-1.09)	(-0.80)	(-1.14)	(-1.17)	(-0.99)	(-1.21)	(-0.94)	(-1.00)
Athlete's wage income ($100s), annual	0.597***	0.958*			0.583**	0.952***		
	(-2.63)	(-1.93)			(-2.27)	(-2.97)		
Parent's wage income ($100s), annual	0.861	0.972			0.484	0.943		
	(-0.14)	(-0.28)			(-0.46)	(-1.22)		
Athlete's income minus parent's income ($100s)			0.964***	0.966***			0.970**	0.974*
			(-3.50)	(-3.00)			(-2.57)	(-1.79)
Athlete's income from other sources (1=yes, 0=no)	0.085**	0.277*		0.332	0.063*	0.134*		0.447
	(-2.17)	(-1.68)		(-1.57)	(-1.79)	(-1.78)		(-1.16)
Number of weeks worked in 1939	1.054	1.105	1.105	1.089	1.068	1.155*	1.075	1.064
	(1.10)	(1.54)	(1.60)	(1.36)	(1.27)	(1.73)	(1.21)	(1.04)
White (1=yes, 0=no)	1.650	0.722	0.553	0.626	2.498	1.905	0.399	0.465
	(0.65)	(-0.33)	(-0.53)	(-0.44)	(0.73)	(0.47)	(-0.78)	(-0.60)
Height (cm)	1.245***	1.212***	1.217***	1.199***	1.239***	1.222***	1.217***	1.196***
	(3.64)	(3.41)	(3.11)	(3.47)	(3.97)	(3.92)	(3.13)	(2.99)
Modified R^2 based on Royston (2006)	0.642	0.557	0.509	0.547	0.645	0.582	0.499	0.520
Observations	37	37	37	37	37	37	37	37
Log likelihood	-73.94	-78.59	-80.92	-79.02	-73.71	-77.31	-81.39	-80.37

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions for the U.S. sub-sample with data from the 1940 Census along with parental income based on occupations collected in the 1910, 1920, and 1930 U.S. Census. The imputation for parental income is described in detail in Appendix B. All regressions include year and event class effects, as well as year of birth, and indicators for married and home ownership (none of which are significant as in Table IX). Parental income does not enter significantly while the variables for athlete income is large and statistically significant. The results are similar whether parental income is based on 1925 industry classifications (columns 1-4) or 1940 occupational codes (columns 5-8). Robust t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table XII: Cox Regressions with News Coverage Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	U.S. sub-sample: Complete Census data					
Lose (1=yes, 0=no)	0.271**	0.304**			0.687	0.590
	(-2.13)	(-2.16)			(-0.62)	(-0.93)
Count of news stories	0.997		0.998		0.997	
	(-1.14)		(-0.66)		(-0.78)	
Count of news stories $>$ 50 (1=yes, 0=no)		0.431		1.607		1.261
		(-1.01)		(0.57)		(0.25)
Log wage income ($100s), annual			0.655**	0.654**	0.657**	0.664**
			(-2.49)	(-2.47)	(-2.36)	(-2.26)
Income from other sources (1=yes, 0=no)			0.078**	0.067**	0.084**	0.071**
			(-2.55)	(-2.43)	(-2.44)	(-2.36)
Owned home (1=yes, 0=no)			1.814	1.554	1.856	1.610
			(0.92)	(0.66)	(0.91)	(0.66)
Number of weeks worked in 1939			1.079	1.069	1.067	1.059
			(1.34)	(1.54)	(1.12)	(1.30)
White (1=yes, 0=no)	0.681	0.914	2.258	2.320	2.018	2.130
	(-0.63)	(-0.15)	(1.40)	(1.36)	(1.10)	(1.24)
Married (1=yes, 0=no)	0.344*	0.289*	1.649	1.305	1.542	1.160
	(-1.87)	(-1.92)	(0.76)	(0.38)	(0.63)	(0.20)
Year of birth	0.926	0.938	1.056	1.065	1.039	1.042
	(-0.71)	(-0.50)	(0.41)	(0.51)	(0.28)	(0.32)
Height (cm)	1.168***	1.155***	1.183***	1.203***	1.194***	1.210***
	(3.03)	(2.98)	(3.10)	(2.87)	(3.12)	(3.05)
Modified R^2 based on Royston (2006)	0.348	0.361	0.591	0.593	0.594	0.603
Observations	39	39	38	38	38	38
Log likelihood	-95.27	-95.24	-79.83	-79.65	-79.59	-79.10

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions for the U.S. sub-sample with data from the 1940 Census, along with variables for news coverage. All regressions include year and event class effects. The variables for news coverage do not enter significantly and the estimates for income remain large and statistically significant. Robust t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Appendix A: Additional Information on Sample Composition

This appendix presents more information on the full list of events included in the analysis and on the composition of athletes by country. Table A1 lists each event, grouped into one of six mutually exclusive classes of events. Indicators for these six event classes are included in models that include country effects rather than event and year effects. Several individual events were discontinued or replaced in later years (e.g. 200m hurdles, 80m hurdles, 3200m and 4000m steeplechases, pentathlon). The Olympic program in Men's Track and Field has remained largely fixed since 1928. Events that are part of the current Olympic program in Track and Field are denoted with an asterisk in Table A1.

The number of Gold and Silver medalists by country is presented in Table A.2. The first column presents total counts by country and the second column presents counts of athletes with complete data on height, as collected from the site Olympedia.org. The baseline sample includes 170 athletes and 130 of these have recorded data on height and weight. In both cases, the U.S. accounts for over half of the sample, with the remaining composed of athletes primarily from Western Europe, Scandinavia, and Canada.

Table A.1: Categorization of Individual Events into Event Classes

Sprints	Middle-distance	Distance	Throws	Field	Racewalk
100m*	800m*	3000m	56lb weight	Decathlon*	3000m walk
100m hurdles	1500m*	3000m steeplechase*	Discuss*	Heptathlon	3500m walk
110m hurdles*		3200m steeplechase	Discuss, ancient style	Pentathlon	10km walk
200m*		4000m steeplechase	Discuss, both hands	Triathlon (long jump, shot, 100y)	10 mile walk
200m hurdles		5000m*	Hammer*	High jump*	20km walk*
400m*		5 miles	Javelin*	High jump, standing	50km walk*
400m hurdles*		10000m*	Javelin, freestyle	Long jump*	
60m		Marathon*	Shot put*	Long jump, standing	
80m hurdles			Shot put, both hands	Pole Vault*	
				Triple jump*	
				Triple jump, standing	

Note: * denotes event is part of current Olympic program in Track and Field.

Table A.2: Number of Observations by Country

Country	N (total)	N (with height and weight data)
Australasia (Australia and New Zealand)	5	2
Argentina	2	2
Belgium	1	1
Canada	5	5
Estonia	1	1
Finland	11	10
France	6	4
Great Britain	21	12
Germany	3	3
Greece	2	0
Haiti	1	1
Hungary	4	2
Italy	3	2
Japan	2	2
Latvia	1	0
Norway	2	1
South Africa	3	2
Sri Lanka	1	0
Switzerland	2	1
Sweden	10	7
Czechoslovakia	1	0
USA	98	84
Yugoslavia	1	1
Total	187	143

Appendix B: Robustness to Sample Restrictions

The correlation between lifespan and losing is robust to relaxing the sample restrictions described in the main text. Appendix B Table 1 presents regression results that exclude any athletes who die before age 40, in addition to dying in war. In 1900, life expectancy at age 20 or 25 (when many athletes were competing) was generally another 40 years among countries in the sample. Life expectancy at selected ages and time periods are presented in life tables from the Australian Institute of Health and Welfare (2017) and the U.S. National Center for Health Statistics (2016). So having already survived to compete in the Olympics, one could expect these athletes to live to roughly age 65. Deaths before age 40 may represent noise rather than a response to winning or losing in the Olympic Games. The magnitudes of the coefficient estimates are larger (smaller hazard ratios) than the main results in Table III and are more precisely estimated. Similar results are obtained in using other age cutoffs, such as age 30 or 50. The main findings are also robust to including athletes who compete in multiple Olympic Games, using either the best finish from their first Olympics (when "treatment" is first assigned) or across all Olympic Games (Appendix B Table 2).

Table B.1: Robustness Tests: Survival Regressions Excluding Deaths Before Age 40

	(1)	(2)	(3)	(4)	(5)	(6)
	Cox	Cox	MPH	MPH	Gompertz	Gompertz
Lose (1=yes, 0=no)	0.662**	0.667**	0.660**	0.667**	0.661**	0.667**
	(-2.23)	(-2.19)	(-2.24)	(-2.18)	(-2.24)	(-2.18)
Year of birth	0.979	0.981	0.983	0.984	0.983	0.984
	(-0.64)	(-0.60)	(-0.54)	(-0.49)	(-0.54)	(-0.49)
Ever set World Record		0.802		0.778		0.779
(1=yes, 0=no)		(-0.40)		(-0.46)		(-0.46)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual event effects	Yes	Yes	Yes	Yes	Yes	Yes
Country effects	No	No	No	No	No	No
Event class effects	No	No	No	No	No	No
Frailty	None	None	Individual	Individual	None	None
Observations	180	180	180	180	180	180
Log likelihood	-729.6	-729.5	108.4	108.5	108.4	108.5

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival model regressions. As a check that the main results are not driven by outliers, athletes who die before age 40 are excluded. Losing is defined as finishing in second place. Columns 1 and 2 estimate Cox models. Columns 3 and 4 estimate Mixed Proportional Hazards (MPH) models that assume the hazard follows a Gompertz distribution and allows for individual heterogeneity. Columns 5 and 6 estimates survival models that assume a Gompertz distribution for the hazard without allowing for individual heterogeneity. The coefficient estimate below 1 on losing indicates the hazard of death is lower among Silver medalists than Gold medalists. Robust t-statistics clustered by event-year in parentheses, except in models with individual heterogeneity (frailty). *p<0.1, **p<0.05, ***p<0.01.

Table B.2: Robustness Tests: Survival Regressions Including Multiple Olympic Games

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1st vs. 2nd places				1st vs. 2nd-4th places			
	Rank using best finish from: first Olympic Games		Rank using best finish from: any Olympic Games		Rank using best finish from: first Olympic Games		Rank using best finish from: any Olympic Games	
	Cox	MPH	Cox	Gompertz	Cox	MPH	Cox	Gompertz
Lose (1=yes, 0=no)	0.802*	0.810*	0.776*	0.784*	0.786**	0.804**	0.817**	0.832**
	(-1.66)	(-1.67)	(-1.94)	(-1.93)	(-2.43)	(-2.28)	(-2.17)	(-2.05)
Year of birth	1.022	1.022	1.022	1.023	1.021	1.022	1.020	1.021
	(1.05)	(1.02)	(1.07)	(1.04)	(1.45)	(1.57)	(1.38)	(1.50)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual event effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country effects	No	No	No	No	No	No	No	No
Event class effects	No	No	No	No	No	No	No	No
Frailty	None	Individual	None	None	None	Individual	None	None
Observations	315	315	315	315	597	597	597	597
Log likelihood	-1462.6	124.7	-1462.2	125.2	-3183.7	190.8	-3184.4	190.3

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival model regressions. These samples include athletes competing in multiple Olympic Games. The coefficient estimate below 1 on losing indicates the hazard of death is lower among Silver medalists than Gold medalists. Robust t-statistics clustered by event-year in parentheses, except in models with individual heterogeneity (frailty). *p<0.1, **p<0.05, ***p<0.01.

Appendix C: Data Collection and Analysis of U.S. Census Surveys

This appendix describes the procedure for collecting Census records for each athlete. Individual Census records are made publicly available 72 years after each survey. The geneology website Ancestry.com provides digitized Census records from each Census from 1850 through 1940, which can be used to identify specific people based on information recorded in the surveys. To retrieve the records for each U.S. athlete, I first searched using the athlete's name, year of birth, and state of birth. I also followed the "Suggested Hints" provided by Ancestry, which link to other Census records as well as other documents like birth, marriage, and death certificates and army registration cards. These hints are created through a machine learning process and through the family trees built by geneological research that link historical records together. In some cases, the names on the original hand-written Census records are imprecise, leading to the digitized records to be misspelled and requiring additional strategies to search for athletes. For example, Ancestry's digitized records mistakenly list Edward Gourdin as Edward Gonodin, Robert Van Osdel as Robert Van Vadel, Leo Sexton as Leo Septon, and Raymond Barbuti as Raymond Barbutte. To locate athletes whose names do not appear on any of the search returns, I conduct a geographical search that starts with recent known street addresses from either 1930 Census records or army registration cards. The army registration cards also include the date of birth, rather than simply the year of birth as recorded in the Census records, which increases the likelihood of a match along with the athlete's name and place of birth. I then work backwards, manually combing through the the list of Census records from a specific geographical location to retrieve the records of athletes whose names have been misspelled. When street addresses are not available, I begin with all males born in the athlete's state of birth during a 1-year window (older and younger) around the athlete's year of birth.

This process retrieves 80 percent of U.S. athletes in the 1940 Census who competed between 1920 and 1936. This high rate is achieved by a detailed inspection for each athlete based not only on searching by name, but also on geographic and demographic information collected from other biographical sources to narrow the search process. Other studies in economic history that merge individual records across surveys by surname, year of birth, and place of birth, tend to have substantially lower match rates because they use much larger samples of Census data (see e.g. Abramitzky et al. 2012, 2014; Bleakley and Ferrie 2016).

The 1940 Census includes variables for wage income earned in 1939 and whether any income was earned from supplemental sources (yes or no). The survey also collects the number of weeks worked in 1939, number of hours worked the prior week, whether the person owns their home or rents, and their stated occupation and industry of employment. The digitized records available on Ancestry include the name of the occupation but not the 3-digit occupation code used for classification, which are instead only listed on the hand-written sheets. I collect the 3-digit occupation codes for each athlete from their original Census records. The 1910, 1920, and 1930 Censuses include occupation and industry but not income. I link athletes to these earlier surveys to record the occupations of their parents.

Earnings by occupation are imputed following two approaches. The first approach uses the average earnings by 3-digit occupation code as reported in the 100% 1940 IPUMS Census file. There are 235 different occupation codes in the 1940 Census classification. The IPUMS mean earnings by occupation include both salary and other income, using data from the 1950 Census, and combine earnings for both males and females. The second approach is based on 15 different industry classifications in 1925 as reported in Margo (2006): gas and electricity; farming; manufacturing; mining; construction; railroad; telephone; wholesale and retail trade; finance, insurance, and real estate; domestic services; medical services; public school teachers; nonprofit services; personal services; and government. This approach relies on assigning the alphabetic occupation and industry fields recorded on the Census records to one of these 15 industries, and so is less precise than using the 1940 numeric occupational codes. While the first approach using 3-digit occupation codes is the preferred method since it is more detailed, the results are nevertheless qualitatively similar in both cases.