ESTIMATING AGING EFFECTS IN RUNNING EVENTS

By

Ray Fair and Edward H. Kaplan

August 2017

COWLES FOUNDATION DISCUSSION PAPER NO. 3000

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

Estimating Aging Effects in Running Events

Ray C. Fair[*] and Edward H. Kaplan[†]

August 2017

Abstract

This paper uses world running records by age to estimate a biological frontier of decline rates. Two models are compared: a linear/quadratic (LQ) model and a non-parametric model. Two estimation methods are used: 1) minimizing the squared difference between the observed records and the modeled biological frontier and 2) using extreme value theory to estimate the biological frontier that maximizes the probability of observing the existing world records by age. The results support the LQ model and suggest there is linear percentage decline up to the late 70's and quadratic decline after that. The extreme value estimates suggest that the true biological frontier is on average about 8 percent below the existing world records. The estimated age factors are also compared to the World Master Athletics (WMA) age factors. The two sets of age factors are close except at the old ages, where the WMA factors are noticeably smaller. Also, the WMA age factors do not meet an important biological constraint.

1 Introduction

An important biological question is how fast people's physical abilities decline with age. This paper focuses on running events. World records by age are used to estimate decline rates. Two models are compared. One is the linear/quadratic

[*]Cowles Foundation, Department of Economics, Yale University, New Haven, CT 06520-8281. e-mail: ray.fair@yale.edu; website: *fairmodel.econ.yale.edu*

[†]William N. and Marie A. Beach Professor of Operations Research, Professor of Public Health, and Professor of Engineering, Yale School of Management, Yale University, New Haven, CT 06520-8200. e-mail: edward.kaplan@yale.edu; website: *http://faculty.som.yale.edu/EdKaplan/*

(LQ) model in Fair (1994, 2007), and the other is a nonparametric (NP) model. Comparisons are also made to the World Masters Athletics (WMA) age factors, which are sometimes used for age grading in races. Nearly a hundred years ago Hall (1925) pointed out the potential usefulness of athletic records to study the physiology of muscular exercise. The present paper is in this tradition.[1] World records by age are used to estimate a biological frontier of decline rates.

This paper focuses on two restrictions that seem sensible biologically. The first is that after a certain age (40 is used here) the rate of decline is non decreasing with age. This will be called the "first derivative" restriction. The second is that the change in the rate of decline is non decreasing with age. This will be called the "second derivative" restriction. In short, after decline begins, nothing gets better with age. The LQ model automatically meets these restrictions, and they are imposed on the NP model. In contrast, the WMA decline rates do not always meet the second derivative restriction. As shown below, the WMA second derivatives sometimes decrease with age, and in some cases they not only decrease with age but are negative.

The first set of estimates finds the closest possible lower envelope on the observed data subject to the derivative restrictions mentioned above, where "close" is defined according to the squared difference between the observed records and the modeled biological frontier. These estimates are pessimistic regarding the further evolution of age world records in that it is possible for observed records and values on the biological frontier to be equal for some age/event pairings if the derivative constraints allow this to occur, meaning that no further reductions are possible for such age/event combinations.

The second set of estimates relies on extreme value theory to estimate the biological frontier that maximizes the probability of observing the existing world records by age. This approach requires a strong assumption regarding the distribution of the gaps between the existing records and the biological frontier. These

[1] See Fair (2007) for a review of studies taking this approach.

gaps are assumed to have the same distribution for each age and event. Although this is a very strong assumption, it turns out that the two sets of estimates are similar regarding the estimated rates of decline. An advantage of using extreme value theory is that it permits an estimate of how far the existing world records are from the true frontier on average. In other words, it provides an estimate of how far the current world age records are expected to decline.

As will be seen, the results support the LQ model over the NP model. The latter has results at the very old ages that are not sensible. However, even the NP model dominates the WMA estimates except at the very old ages because it meets the second derivative restrictions, which the WMA age factors do not.

Two interesting features of the LQ results are 1) the decline rate (in percentage terms) is linear up to the late 70's and 2) even at age 90 people are only a little more than twice as slow as they were in their peak years. Assuming that one is not injured or sick, stays in peak shape age corrected, and declines in percentage terms at the same rate as the world records, life is good. The extreme value results estimate that the existing world records are on average about 7.8 percent above the true frontier.

2 The Models

The Linear/Quadratic Model

Assume that one has world records by age for a given running event, where r_k will be used to denote the log of the record time for age k. Using logs means that all decline rates are in percentage terms. For the results below k ranges from 40 to 95 per running event. b_k will be used to denote log of the (unobserved) biological minimum time for age k. By definition,

$$r_k = b_k + \epsilon_k \quad , \tag{1}$$

where ϵ_k is the gap between the record time and the true biological minimum time. It will be close to zero if the record time is close to the biological minimum. Otherwise it is positive. More will be said about this below.

The LQ model postulates that the decline rate (in percentage terms) is linear up to a transition age and then quadratic after that. The transition age is one of the estimated parameters. At the transition age the linear and quadratic segments are constrained to touch and to have the same first derivative. The formula for b_k is

$$b_k = \begin{cases} \beta + \alpha k, & 40 \leq k \leq k^*, \quad \alpha > 0 \\ \gamma + \theta k + \delta k^2, & k > k^*, \quad \delta > 0 \end{cases} \tag{2}$$

with the restrictions

$$\begin{aligned} \gamma &= \beta + \delta k^{*2} \\ \theta &= \alpha - 2\delta k^* \end{aligned} \tag{3}$$

The two restrictions force the linear and quadratic segments to touch and to have the same first derivative at k^*. The unrestricted parameters to estimate are the intercept, β, the slope of the linear segment, α, the age at which the line changes from linear to quadratic, k^*, and the quadratic parameter, δ. The first derivative of b_k with respect to k is α up to the transition age and then increases by a contant amount (2δ) after that. The second derivative is zero up to the transition age and then constant (2δ) after that.

The equation that is estimated is then

$$r_k = \beta + \alpha k + \delta d_k (k^{*2} - 2k^* k + k^2) + \epsilon_k, \tag{4}$$

where $d_k = 0$ if $k \leq k^*$ and $d_k = 1$ if $k > k^*$.

In the data r_k is sometimes "dominated" in that it is greater than r_{k+1}. Under the assumption that b_k never declines with age (after age 40), dominated times must have large gaps associated with them. The dominated observations are "soft." For this reason in the estimation work dominated observations are not used. Since ϵ_k can never be negative, equation (4) is estimated under the restriction that all estimated gaps are non negative.

For the results in Section 6 "age factors," denoted R_k, are presented. They are computed as follows. Let $\hat{b}_k$ denote the predicted value of b_k using the estimated values of β, α, k^*, and δ for $k = 40, \ldots, 95$. Then R_k is

$$R_k = e^{\hat{b}_k}/e^{\hat{b}_{40}}, \;\; k = 40, \ldots, 95 \;. \tag{5}$$

The Nonparametric Model

The LQ model is tightly parameterized, and it is interesting to see how it does against a nonparameteric model constrained only by the first and second derivatives being non-negative and non-decreasing. Letting $\mathcal{A}$ denote the set of ages for which non-dominated observations exist, this is readily achieved as the solution to the following quadratic programming problem:

$$\min_{\{b_{37},\ldots,b_{95}\}} \sum_{k \in \mathcal{A}} (r_k - b_k)^2 \tag{6}$$

subject to

$$b_k \leq r_k, \quad k \in \mathcal{A} \tag{7}$$

$$b_k - b_{k-1} \geq 0, \quad k = 40, ..., 95 \tag{8}$$

$$b_k - 2b_{k-1} + b_{k-2} \geq 0, \quad k = 40, ..., 95 \tag{9}$$

$$b_k - 3b_{k-1} + 3b_{k-2} - b_{k-3} \geq 0, \quad k = 40, ..., 95 \tag{10}$$

The constraints are easily understood: constraint (7) forces b_k to fall at or below the observed records at all ages for which data exist; constraint (8) ensures that the b_k curve can never decrease (equivalent to a non-negative first derivative); constraint (9) enforces convexity (equivalent to a non-negative second derivative); and constraint (10) ensures that the second derivative is non-decreasing. Note that even though only records from age 40 through 95 are used in fitting the model, it is necessary to estimate biological limits at ages 37, 38 and 39 as well to enable constraints (8)-(10). Solution of this model results in linear/quadratic segments that provide a lower envelope of the data that minimizes the sum of squared residuals.

3 The Data

Data for six running events were obtained from the site of the Association of Road Racing Statisticians (AARS): *http://www.arrs.net/SARec.htm*. The data are AARS recognized world records by age. Four of the events are road racing events: 5K, 10K, Half Marathon, and Marathon. Two of the events are outdoor track events: 5,000 meters and 10,000 meters. The road racing data are better than those used in Fair (2007) because they pertain to all runners. The earlier data pertained only to U.S. citizens. In addition, the data ended in 2003 in the earlier study compared to 2016 in the present study. In Fair (2007) the events were 800 meters, 1,500 meters, 5,000 meters, 10,000 meters, 5K, 10K, and Marathon. For the present study the first two have been dropped (no data) and the half marathon has been added.

Age 40 was used as the initial age, and, as noted in Section 2, dominated times were excluded. The number of observations for the 5K, 10K, Half Marathon, Marathon, 5,000 meters, and 10,000 meters were 32, 31, 33, 33, 38, and 33, respectively, for a total of 200 observations. The oldest age for each event was 94, 92, 91, 90, 95, and 93, respectively. The total number of observations for ages 86 and above was 32.

In Fair (2007) the initial age was taken to be 35 instead of 40. The current data, however, suggest that there is not much decline between 35 and 40, certainly less than from age 40 on. The initial age was thus taken to be 40, and no attempt was made to estimate decline rates before 40. As a rough approximation the data suggest that there is about a one percent decline over the five-year period 35–39 (total over the five years).

For comparison purposes WMA data were also collected. The site is *http://www.runscore.com/Alan/AgeGrade.html*. The most recent 2015 tables were used. Data for 6 events were obtained: 5K, 10K, Half Marathon, Marathon, 5,000 meters, and 10,000 meters. The Half Marathon and Marathon have the same age factors, as do the 5,000 meters and 10,000 meters. There are thus 4 different sets of age factors. Observations were obtained for ages 40 through 95. As noted in Section 1, the WMA age factors are sometimes used for age scoring in races.

It may be useful to give a sense of why some observations may be soft. A soft observation can be interpreted as not enough elite runners of the particular age have run the particular distance to have the biological minimum being close to being achieved. Consider, for example, the best age 71 marathon time, which is 3:00:58. The best age 73 time is noticeably smaller: 2:54:48—Ed Whitlock. The best age 83 time is 4:19:07. The best age 85 time is noticeably smaller: 3:56:38—Ed Whitlock. The slower times at the younger ages can be taken to mean that there have not been enough Ed Whitlock's of the particular age running the marathon to be close to the biological minimum. These are the dominated observations that have been excluded from the estimation. For a possibly soft non-dominated observation, take the best age 89 time in the marathon, which is 6:35:38. The best age 90 time is is essentially the same at 6:35:47. But surely b_{90} is noticeably larger than b_{89}, which means at least that the gap for age 89 is large. This problem is also the reason women's data have not been used in this study. More time is needed to build up a sample of old women running road races.

Regarding possible changes over time, it may be that the b_k curve is shifting down over time as nutrition, running knowledge, and the like improve. For this paper it is assumed that the curve does not shift over time. The world record data are primarily since 1990, and changes in b_k between 1990 and the present are likely to be modest relative to the fluctuations in r_k due to the soft-times problem. Of the 200 non dominated observations used in the estimation, only 26 occurred before 1990. The earliest was 1975, and there were a total of four in the 1970's.

4 Pooling

As just discussed, a problem with the data is that there are not many observations at the very old ages and some of these may be soft. To lessen the sensitivity of the results to this problem, the six events have been pooled. The six events were pooled under the assumption that the curve for each event is the same except for the intercept. For LQ this means that 9 coefficients are estimated in equation (4): α, δ, k^*, and the six β's. As noted above, the total number of observations is 200.

Figure 1 provides some support for the pooling, where the non-dominated times are plotted. It is clear that the points per event reveal very similar patterns across the ages.

Pooling for the NP model can be achieved by postulating a fixed effects model. Letting r_{ik} and b_{ik} denote the observed records and biological minima at age k for the i^{th} event, it is postulated that

$$b_{ik} = b_{1k} + \beta_i, \quad i = 2, ..., 6 \tag{11}$$

where the β_i's represent event-specific fixed effects. Let $\mathcal{A}_i$ denote the set of non-dominated ages for which there are observations in the i^{th} event. The quadratic program objective described above is then modified to be

$$\min_{\{b_{1,37},...,b_{1,95},\beta_2,...,\beta_6\}} \sum_{i=1}^{6} \sum_{k \in \mathcal{A}_i} (r_{ik} - b_{ik})^2 \tag{12}$$

and the constraints are modified to be

$$b_{ik} = b_{1k} + \beta_i, \quad i = 2, ..., 6 \tag{13}$$

$$b_{ik} \leq r_{ik}, \quad i = 1, ..., 6; \quad k \in \mathcal{A}_i \tag{14}$$

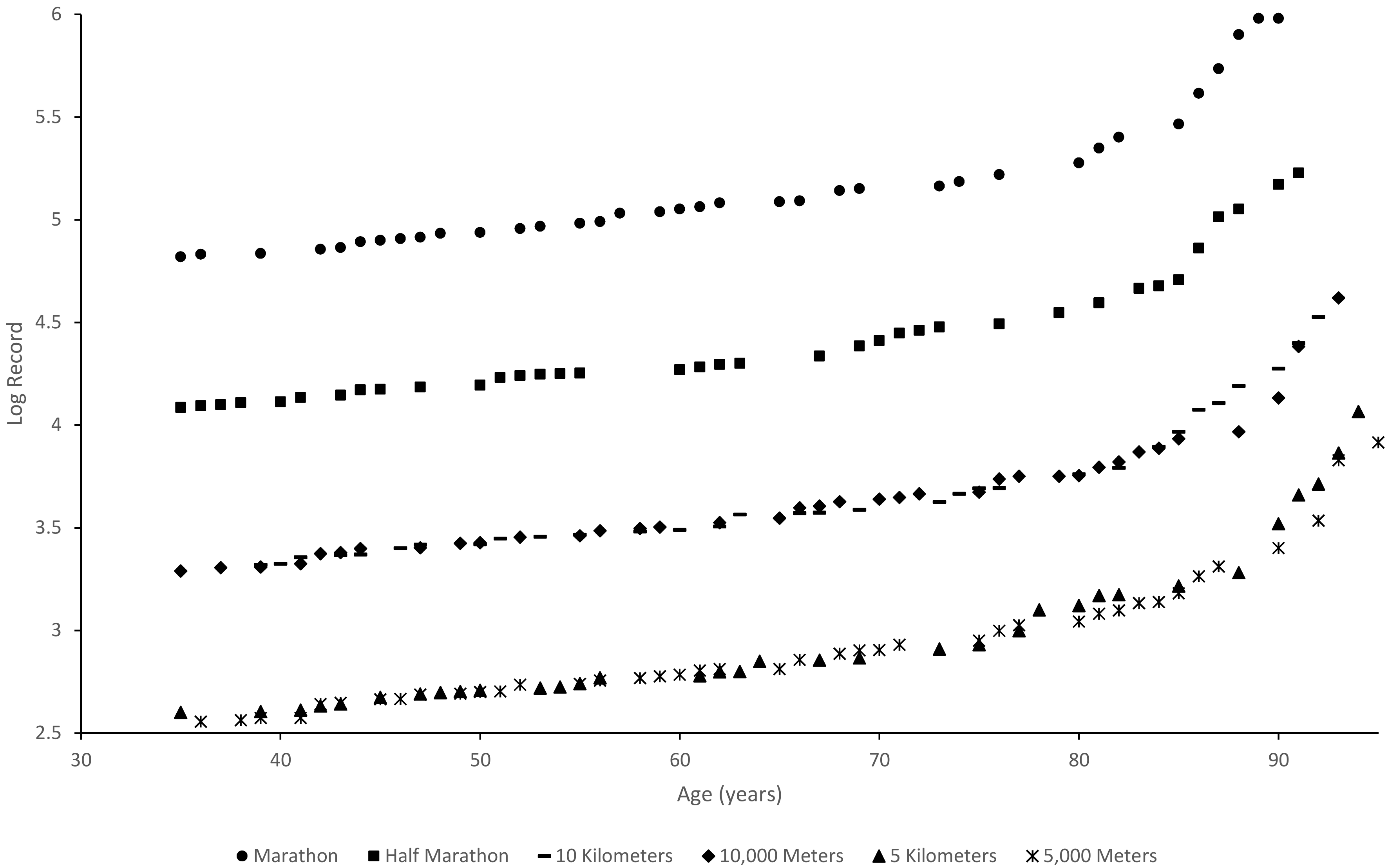
$$b_{1k} - b_{1,k-1} \geq 0, \quad k = 40, ..., 95 \tag{15}$$

$$b_{1k} - 2b_{1,k-1} + b_{1,k-2} \geq 0, \quad k = 40, ..., 95 \tag{16}$$

$$b_{1k} - 3b_{1,k-1} + 3b_{1,k-2} - b_{1,k-3} \geq 0, \quad k = 40, ..., 95 \tag{17}$$

In addition to defining the biological limits for events 2 through 6 via a constant translation of the limits for event 1 at each age, the constraints enforce the non-decreasing convex lower envelope on the observed records for each event. Note that via the fixed effects, the constraints on events 2 through 6 also impact the estimation of the biological limits for event 1.

Figure 1
Log Age Records, Six Events
Non-Dominated Observations



5 Use of Extreme Value Theory

The estimation of the two models described above finds the closest possible lower envelope on the observed data subject to the derivative restrictions, where "close" is defined according to the squared difference between observed records and the modeled biological frontier. This section discusses the use of extreme value theory to estimate the true b_{ik} curve statistically.

Extreme value theory (see Chapter 9 in David (1970)) is that branch of probability that focuses on the distribution of extreme values— maxima and minima— sampled from populations of interest. This theory applies directly to the current problem since by definition a world running record is the minimum recorded time of all attempts at the distance in question. This definition carries over to age-dependent records, each of which can be recognized as the minimum time observed to date over all such prior attempts.

The key fact that will be employed is generally attributed to Gnedenko (1943), which loosely states that for a non-negative random variable, as the sample size $n \to \infty$, if the probability distribution of the observed minimum from a sample of size n exists, then it must be in the form of the Weibull distribution. In the present application this means for a given event i and age k that the prior probability distribution of the random variable R_{ik}, the minimum running time, should be given by

$$\Pr\{R_{ik} \leq r_{ik}\} = 1 - e^{-\eta_{ik}(r_{ik}-b_{ik})^{\lambda_{ik}}}; r_{ik} > b_{ik}; \quad \eta_{ik} > 0; \lambda_{ik} > 0. \tag{18}$$

In this model, η_{ik} is referred to as the *scale* parameter, λ_{ik} is called the *shape* parameter, and b_{ik} is known as the *shift* parameter. In the present application the shift parameter b_{ik} corresponds precisely to the true biological minimum b_{ik} and is the object of the estimation.

Einmahl and Magnus (2008) previously applied extreme value theory to estimate the fastest times or largest distances possible, but their approach is different from the current approach in that they did not attempt to estimate biological limits as a function of age. They instead based their estimates on reports of the *m* fastest times recorded to date for the largest recorded number *m* available from data. Their results suggest that observed records for many running events are extremely close to the minimum achievable. For example, they estimate that only 20 seconds could be shaved off of the world record for the marathon, while only another 3 to 4 seconds could fall from the existing world record in the 1,500 meters.

Ideally, one would want to identify the scale and shape parameters for all event and age combinations, but recall that the data provide only one observation—the world age record time—for 200 age/event combinations. To identify the parameters it is assumed that the scale and shape parameters are the same for all age/event pairs. b_{ik} is then modeled using either the LQ model or the NP model, where the six events are pooled as discussed in Section 4. While clearly a simplification, assuming that $\eta_{ik} = \eta$ and $\lambda_{ik} = \lambda$ does afford a very simple interpretation of the results—on average the increase on the log scale from the biological minimum to the modeled event record, call this $\overline{\Delta}_{\text{log}}$, is given by the relatively simple formula

$$\overline{\Delta}_{\text{log}} = \left(\frac{1}{\eta}\right)^{\frac{1}{\lambda}} \Gamma(1 + \frac{1}{\lambda}) \tag{19}$$

where $\Gamma(\cdot)$ is the well-known gamma function. Equation (19) is just the expected value of the zero-shift Weibull random variable. This implies that the percentage decrease from the observed record to the biological minimum is approximately

$$\text{Percent Decrease} = 1 - e^{-\overline{\Delta}_{\text{log}}}. \tag{20}$$

To estimate extreme value models, instead of minimizing the sum of squared deviations subject to constraints, maximum likelihood estimation is employed. The probability density corresponding to a given observed record r_{ik} is given by

$$f_{R_{ik}}(r_{ik}) = \eta\lambda(r_{ik} - b_{ik})^{\lambda-1} e^{-\eta(r_{ik}-b_{ik})^{\lambda}}, \quad r_{ik} > b_{ik}. \tag{21}$$

Maximum likelihood estimates are most conveniently obtained by maximizing the sum of the *log* of the probability densities over all observations. Thus, for the NP model, the maximum likelihood estimates are the solutions to

$$\max_{\{b_{1,37},\ldots,b_{1,95},\beta_2,\ldots,\beta_6,\eta,\lambda\}} \sum_{i=1}^{6} \sum_{k\in\mathcal{A}_i} (\log(\eta\lambda)+(\lambda-1)\log(r_{ik}-b_{ik})-\eta(r_{ik}-b_{ik})^{\lambda}) \quad (22)$$

subject to constraints (13)–(17). Similarly, the maximum likelihood estimates for LQ are the solutions to the problem shown above, but with b_{ik} specified according to the LQ model in equations (2)–(3).

6 The Results

There are two models to estimate—LQ and NP—and two estimation methods—minimizing the sum of squared deviations subject to the constraints and maximum likelihood. The four sets of estimates will be denoted LQmin, LQext, NPmin, NPext, respectively. The results from these estmates will be compared to the WMA results and to the earlier results for the LQ model in Fair (2007). The results are presented in Tables 1 and 2 and Figures 2 and 3.[2]

Table 1 first presents the implied age factors for ages 50, 60, 70, 80, 90, and 95, where the age factors are given by equation (5). It then presents the 10-year rates of decline. Table 2 presents the first and second derivatives of b_{ik} with respect to k for LQmin, NPmin, WMA 5K, and WMA MA. Figure 2 plots the predicted values of b_{ik} for the 5K for LQmin, NPmin, WMA. Figure 3 does the same for the marathon. The actual values in Figures 2 and 3 are the non-dominated times.

The results are similar in Table 1 for the three LQ rows. The estimates of α, k^*, and δ are slightly higher for LQext versus LQmin, which leads to the age factors being slightly higher. Both are slightly higher than the LQ estimates in Fair (2007).

[2]Since the WMA results are the same for the half marathon and marathon, only the marathon results are presented. Similarly for the 5,000 meters and 10,000 meters; only the 5,000 meter results are presented. In Table 2 only the 5K and marathon results are presented for WMA. The WMA age factors were normalized to have them equal 1.0 at age 40 to make them comparable to the other age factors. The same was true for the age factors in Fair (2007).

Table 1
Coefficient Estimates and Implied Age Factors

	Estimates			Age Factors						No.	Max
	$\hat{\alpha}$	$\hat{k}^*$	$\hat{\delta}$	R_{50}	R_{60}	R_{70}	R_{80}	R_{90}	R_{95}	Obs.	Age
LQ: Fair(2007)	0.0080	75.1	0.00164	1.08	1.17	1.27	1.43	2.15	2.99	267	96
LQmin	0.0088	77.7	0.00251	1.09	1.19	1.30	1.44	2.27	3.43	200	95
LQext	0.0101	79.9	0.00307	1.11	1.22	1.35	1.50	2.27	3.51	200	95
NPmin				1.05	1.14	1.28	1.52	2.17	3.84	200	95
NPext				1.07	1.15	1.27	1.50	2.19	3.83	200	95
WMA 5K				1.08	1.16	1.27	1.51	2.12	2.84		
WMA 10K				1.08	1.18	1.30	1.55	2.19	2.97		
WMA MA				1.08	1.18	1.30	1.55	2.19	2.97		
WMA 5,000M				1.08	1.18	1.29	1.54	2.19	2.98		

10-year Rates of Decline (percentage points)

	41-50	51-60	61-70	71-80	81-90	86-95
LQ: Fair (2007)	8.4	8.4	8.4	12.8	50.1	76.9
LQmin	9.2	9.2	9.2	10.6	57.3	102.1
LQext	10.6	10.6	10.6	10.6	51.4	105.8
NPmin	4.8	8.6	12.6	18.5	42.7	122.9
NPext	7.0	7.4	10.1	18.7	45.8	122.5
WMA 5K	7.6	8.2	9.4	18.7	40.1	63.0
WMA 10K	8.4	9.2	10.2	18.7	41.4	66.4
WMA MA	8.5	9.5	10.5	18.0	41.1	66.5
WMA 5,000M	8.0	8.8	9.9	19.2	42.1	67.3

In general, however, the estimates are all fairly close. The estimates are also close for NPmin versus NPext, especially at the older ages.

Comparing the LQ rows with the NP rows, NP has slightly lower age factors through age 70 and higher age factors at age 95. The larger age factors at age 95 reveals a problem with the NP estimates, which can be seen in Table 2 for NPmin. At age 88 there is a huge jump in the second derivative, which means a large change in the decline rate after that. This large jump does not seem sensible biologically (there is nothing magic about age 88), and so the NP estimates after age 88 are not trustworthy. NPext has a similar jump at age 88.

Table 2
First and Second Derivatives)

	LQmin		NPmin		WMA 5K		WMA MA	
	100	100	100	100	100	100	100	100
Age	Δb_k	$\Delta^2 b_k$	Δb_k	$\Delta^2 b_k$	Δb_k	$\Delta^2 b_k$	Δb_k	$\Delta^2 b_k$
42	0.8834	0.0000	0.3466	0.0357	0.7150	0.0051	0.7692	0.0778
43	0.8834	0.0000	0.3824	0.0357	0.7202	0.0051	0.8172	0.0480
44	0.8834	0.0000	0.4180	0.0357	0.7254	0.0052	0.8239	0.0067
45	0.8834	0.0000	0.4536	0.0357	0.7307	0.0053	0.8308	0.0068
46	0.8834	0.0000	0.4893	0.0357	0.7361	0.0054	0.8269	-0.0038
47	0.8834	0.0000	0.5250	0.0357	0.7415	0.0055	0.8447	0.0178
48	0.8834	0.0000	0.5606	0.0357	0.7471	0.0055	0.8519	0.0072
49	0.8834	0.0000	0.5963	0.0357	0.7527	0.0056	0.8592	0.0073
50	0.8834	0.0000	0.6319	0.0357	0.7584	0.0057	0.8667	0.0074
51	0.8834	0.0000	0.6676	0.0357	0.7642	0.0058	0.8630	-0.0037
52	0.8834	0.0000	0.7032	0.0357	0.7701	0.0059	0.8819	0.0189
53	0.8834	0.0000	0.7390	0.0357	0.7761	0.0060	0.8897	0.0078
54	0.8834	0.0000	0.7745	0.0357	0.7821	0.0061	0.8977	0.0080
55	0.8834	0.0000	0.8103	0.0357	0.7883	0.0062	0.9058	0.0081
56	0.8834	0.0000	0.8459	0.0357	0.7945	0.0063	0.9023	-0.0035
57	0.8834	0.0000	0.8815	0.0357	0.8009	0.0064	0.9224	0.0201
58	0.8834	0.0000	0.9172	0.0357	0.8074	0.0065	0.9310	0.0086
59	0.8834	0.0000	0.9529	0.0357	0.8140	0.0066	0.9398	0.0087
60	0.8834	0.0000	0.9885	0.0357	0.8206	0.0067	0.9487	0.0089
61	0.8834	0.0000	1.0242	0.0357	0.8274	0.0068	0.9454	-0.0033
62	0.8834	0.0000	1.0598	0.0357	0.8343	0.0069	0.9669	0.0215
63	0.8834	0.0000	1.0955	0.0357	0.8413	0.0070	0.9764	0.0094
64	0.8834	0.0000	1.1312	0.0357	0.8485	0.0071	0.9860	0.0096
65	0.8834	0.0000	1.1668	0.0357	0.8557	0.0073	0.9958	0.0098
66	0.8834	0.0000	1.2025	0.0357	0.8631	0.0074	0.9929	-0.0029
67	0.8834	0.0000	1.2381	0.0357	0.8706	0.0075	1.0159	0.0230
68	0.8834	0.0000	1.2738	0.0357	0.9178	0.0472	1.0263	0.0104
69	0.8834	0.0000	1.3095	0.0357	1.0194	0.1016	1.0370	0.0106
70	0.8834	0.0000	1.3451	0.0357	1.0971	0.0777	1.0478	0.0109
71	0.8834	0.0000	1.3808	0.0357	1.2046	0.1074	1.0999	0.0520
72	0.8834	0.0000	1.4164	0.0357	1.3020	0.0974	1.2088	0.1089
73	0.8834	0.0000	1.4521	0.0357	1.4171	0.1151	1.3356	0.1268
74	0.8834	0.0000	1.4877	0.0357	1.5085	0.0914	1.4389	0.1033
75	0.8834	0.0000	1.5998	0.1120	1.6326	0.1241	1.5609	0.1220
76	0.8834	0.0000	1.7117	0.1120	1.7478	0.1152	1.6883	0.1274
77	0.8834	0.0000	1.8237	0.1120	1.8687	0.1209	1.8368	0.1485
78	0.9004	0.0170	1.9357	0.1120	2.0111	0.1424	1.9626	0.1258
79	1.2649	0.3645	2.0476	0.1120	2.1460	0.1349	2.1109	0.1482
80	1.7668	0.5020	2.1597	0.1120	2.2888	0.1428	2.2837	0.1729

Table 2 (continued)

Age	LQmin 100 Δb_k	100 $\Delta^2 b_k$	NPmin 100 Δb_k	100 $\Delta^2 b_k$	WMA 5K 100 Δb_k	100 $\Delta^2 b_k$	WMA MA 100 Δb_k	100 $\Delta^2 b_k$
81	2.2688	0.5020	2.2716	0.1120	2.4405	0.1517	2.4350	0.1513
82	2.7708	0.5020	2.3836	0.1120	2.6023	0.1618	2.6130	0.1780
83	3.2727	0.5020	2.4956	0.1120	2.7926	0.1903	2.8209	0.2079
84	3.7747	0.5020	2.6076	0.1120	2.9795	0.1869	3.0093	0.1884
85	4.2766	0.5020	2.7196	0.1120	3.1812	0.2016	3.2310	0.2217
86	4.7786	0.5020	2.8315	0.1120	3.4187	0.2375	3.4718	0.2408
87	5.2805	0.5020	2.9435	0.1120	3.6382	0.2195	3.7543	0.2825
88	5.7825	0.5020	4.3559	1.4124	3.9189	0.2807	4.0239	0.2696
89	6.2845	0.5020	5.7683	1.4124	4.2070	0.2880	4.3426	0.3187
90	6.7864	0.5020	7.1807	1.4124	4.5482	0.3413	4.7194	0.3768
91	7.2884	0.5020	8.5931	1.4124	4.8824	0.3341	5.0950	0.3756
92	7.7903	0.5020	10.0055	1.4124	5.3063	0.4239	5.5435	0.4484
93	8.2923	0.5020	11.4179	1.4124	5.7610	0.4547	6.0812	0.5377
94	8.7942	0.5020	12.8302	1.4124	6.3088	0.5477	6.6453	0.5641
95	9.2962	0.5020	14.2427	1.4124	6.8833	0.5746	7.3322	0.6869

The WMA age factors are close to those for LQ Fair (2007). This is not, however, completely independent information because the LQ Fair (2007) results have probably influenced the WMA results in various (subjective) ways. The main difference between the current LQ age factors and the WMA age factors is at the very old ages, where the WMA age factors are smaller. Table 2 shows what is problematic about the WMA results. The second derivatives are not always non decreasing, and in five cases for the marathon they are negative. So the biology is probably not quite right.

Looking at Table 2 more closely, for LQ the second derivative is zero up to age 77 and then after a transition it is constant at 0.005020 from age 80 on. For NP the second derivative is 0.000357 up to age 74, then increases to 0.001120 until age 87, and then has the large jump to 0.014124. There are thus three linear/quadratic segments for NP, compared to one, of course, for LQ, where the last segment for NP is not sensible.

Figures 2 and 3 show the closeness of LQmin and NPmin except at the very old ages, where NPmin increases faster. For the WMA lines, they decrease slower at the very old ages. The two figures give a good sense of what is being estimated.

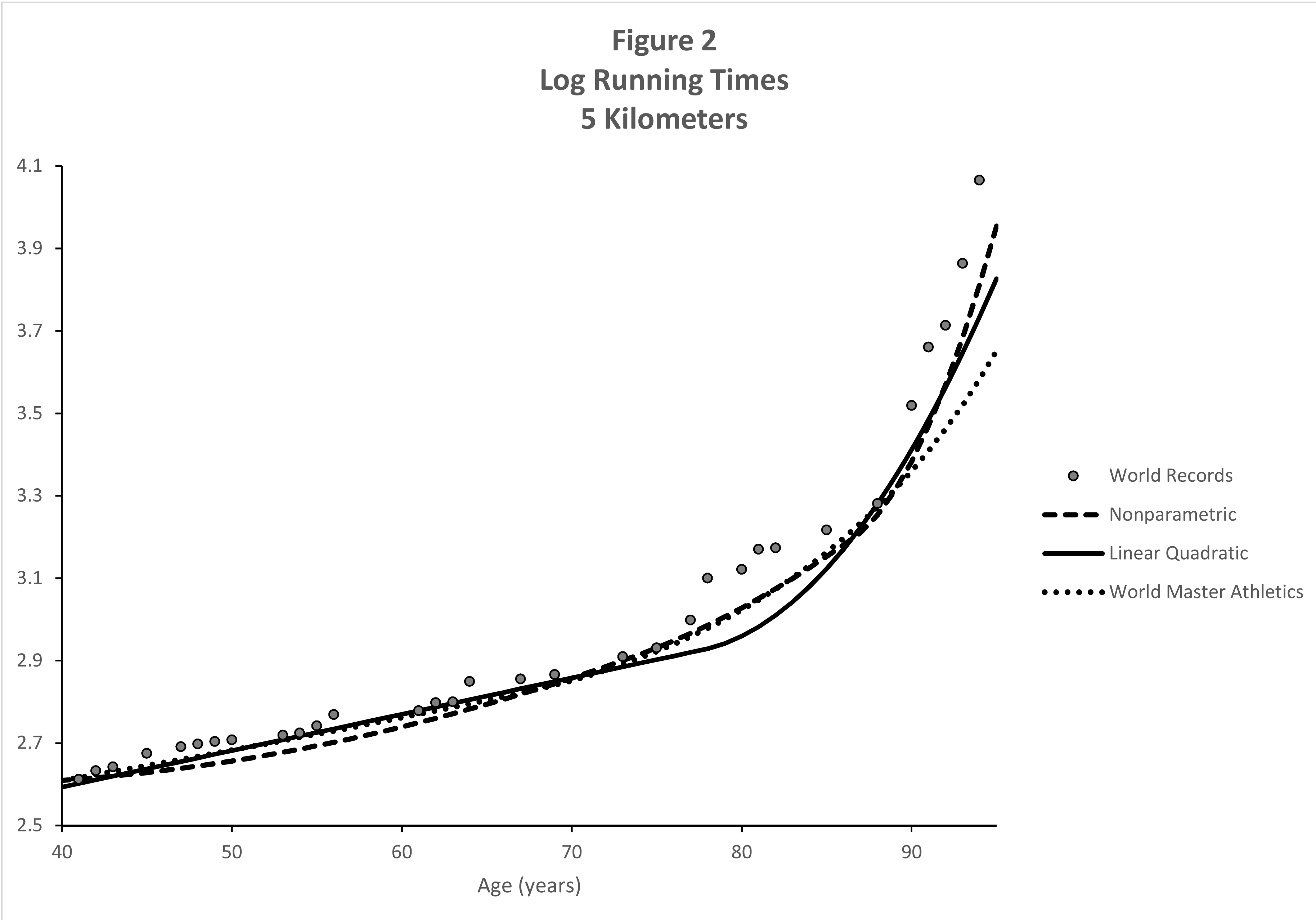


Figure 2
Log Running Times
5 Kilometers
4.1
3.9
3.7
3.5
3.3
3.1
2.9
2.7
2.5
40
50
60
70
80
90
Age (years)
World Records
Nonparametric
Linear Quadratic
World Master Athletics

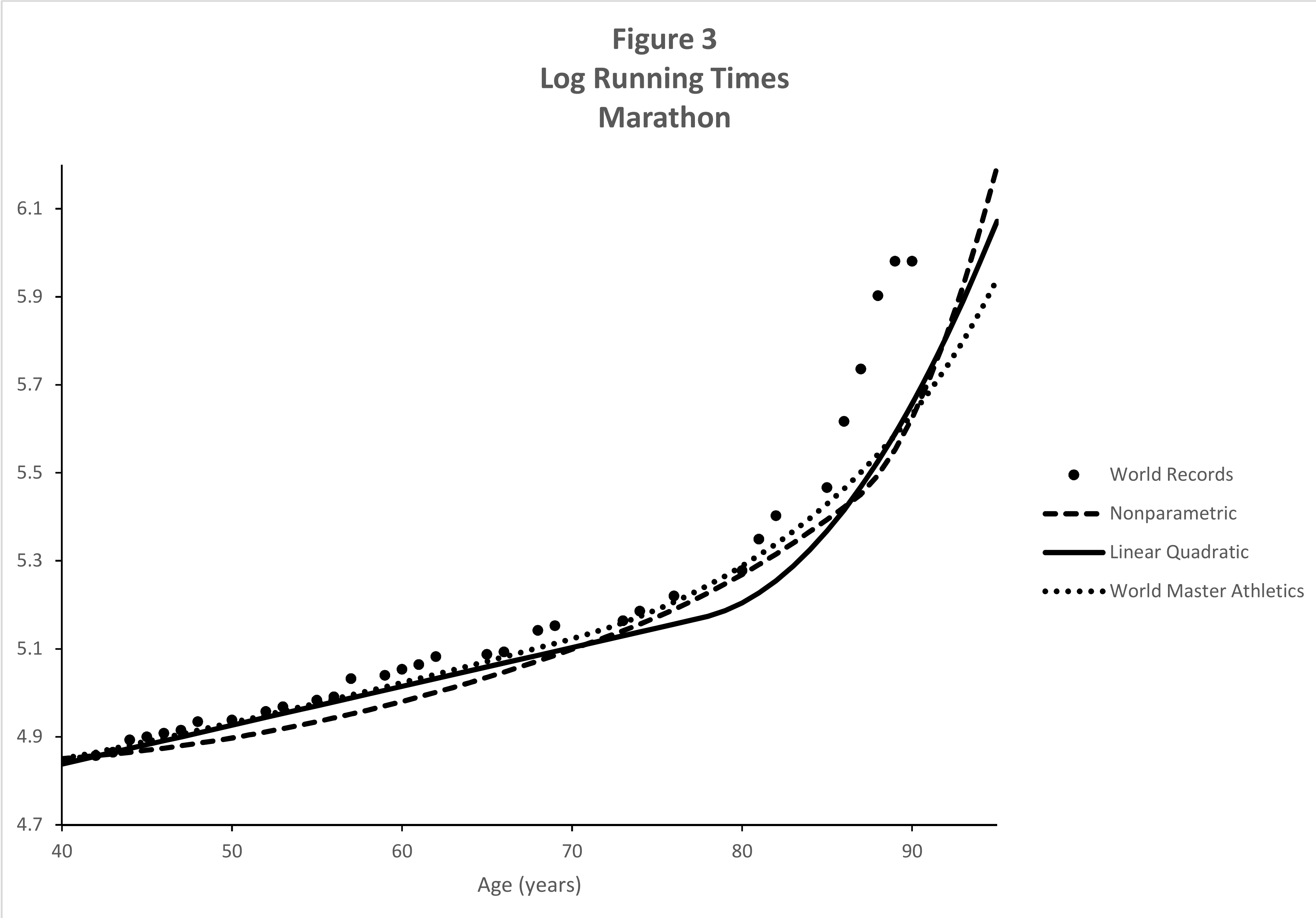


Figure 3
Log Running Times
Marathon
6.1
5.9
5.7
5.5
5.3
5.1
4.9
4.7
40
50
60
70
80
90
Age (years)
World Records
Nonparametric
Linear Quadratic
World Master Athletics

Remember that because of pooling the 5K and marathon lines for LQmin and for NPmin are the same except for the intercept. The plots for LQext and NPext are very close to those for LQmin and NPmin, and these have not been presented for space reasons.

Consider now more detailed results behind the extreme value estimates LQext and NPext. Of great interest is the fact that the estimated Weibull parameters η and λ are quite similar for both models. For the NP model, $\widehat{\eta} = 22.3$ and $\widehat{\lambda} = 1.27$, while for the LQ model $\widehat{\eta} = 18.5$ and $\widehat{\lambda} = 1.15$. Via equation (19) the average log increase of existing records over the biological minimum is equal to 0.0803 and 0.0747 for the NP and LQ models respectively. From equation (20), these estimates imply that the average percentage decrease from observed records to the biological minimum approximately equals 0.0772 and 0.0720 respectively. Clearly these estimates qualitatively convey the same information and suggest that one should not be surprised to see existing records fall by about 7.5 percent on average. This result is quite different from that reported by Einmahl and Magnus (2008). As mentioned earlier, they estimated that the men's marathon record could only be expected to fall by 20 seconds; by contrast, our analysis suggests that the marathon record could be reduced by nine minutes or more.

7 Comparing Fits

Recall that the gaps ϵ_k reflect distance from the (upper bounded) biological minimum b_k to the extant record r_k, and as such are not measurement errors in the usual sense. Nonetheless it is worth discussing briefly the closeness of the convex hulls for the LQ and NP models. That the NP model "fits" better is obvious from a cursory examination of Figures 2 and 3. The sum of squared deviations produced by the NP model equals 2.12, a reduction of 10 percent from the LQ sum of squared deviations of 2.36. However, this reduction comes at considerable cost. While the NP model can indeed be thought of as nonparametric subject only to the constraints that the resulting b_k's are non-decreasing and convex with respect to age, this model actually requires the estimation of 64 parameters: one for each age from 37 to 95 inclusive, plus an additional five fixed effects to enable pooling over all six running events.

To put this in perspective, with 200 actual non-dominated running records used to fit the models in this paper, the LQ model with its nine free parameters averages 22.2 observations per parameter, while the NP model averages $200/64 = 3.1$ observations per parameter estimated. One might argue that since the NP model results in a piecewise quadratic model with only three segments connecting four change points and an initial level (at age 37) for a total of 13 parameters estimated for the 5K model in addition to the five fixed effects, the ratio of data to estimated parameters for the NP should be increased to $200/18 = 11.1$. This sounds better, but the argument for counting only 18 as opposed to 64 parameters is flawed—*a priori* it is not known how many segments are required and where they must be placed to produce the sum-of-squares minimizing lower convex envelope. The formulation employed is truly flexible in that *any* piecewise quadratic frontier could have resulted depending upon the data. Such flexibility does require all 64 parameters; it is cheating to only count the number of segments and change points estimated after the fact. Thus, one is back to a ratio of 3.1 observations per parameter for the QP, which of course suggests the potential for overfitting.

All told, the NP does what is is asked to do, namely, produce the least squares lower convex envelope for the data presented subject to the constraints stated earlier. Biologically, it is difficult to argue why the specific number of resulting breakpoints occurs, in addition to the specific placement of these breakpoints. Ease of model interpretation plus biological plausibility argues in favor of the LQ model, in spite of the slightly better fit offered by NP.

The extreme value estimation also reveals information about the the comparative fit of the NPext versus LQext estimates. The nonparametric log likelihood equals 314.5, only marginally larger than the LQ log likelihood of 313.6. To gain a sense for how close these numbers are, in standard maximum likelihood problems, twice the difference in log likelihoods has a chi-square distribution with degrees of freedom equal to the difference in the number of parameters estimated. Comparing LQ and the NP models, twice the log likelihood difference equals 1.8, while the difference in the number of parameters estimated equals 55. To reject the LQ model in favor of the nonparametric alternative at the usual 5 percent level of significance would require twice the difference in log likelihoods to equal or exceed 73.3, a number far larger than the observed value of 1.8. While strictly speaking the standard maximum likelihood results do not apply to our problem owing to the constraints imposed, it is clear that the improvement in fit from the NP model over the LQ model is insufficient to warrant rejecting LQ. That both models produce similar values of $\overline{\Delta}_{\log}$ is further testimony to the satisfactory fit of LQ.

8 Conclusion

The support for the LQ model in this paper suggests that there is linear decline in percentage terms up to the late 70's and then quadratic decline after that. The estimates at the very old ages must be interpreted with caution because of the small number of observations and the fact that some of the observations may be soft. It is interesting that even though the extreme value theory requires a strong assumption about the distribution of the gaps, the estimated decline rates using this theory are quite similar to those estimated by minimizing the sum of squared deviations. If the extreme value results are to be trusted, they suggest that on average the true biological frontier is about 8 percent below the currently estimated frontier.

Regarding the WMA age factors that are currently used in some races, they are fairly close to the LQ age factors except at the old ages, where they are noticeably lower. They also do not meet the second derivative restriction, which should be corrected.

As noted in the Introduction, the LQ results are encouraging to humankind in that there is only linear percentge decline up to the late 70's and that after age 90 the age factors are only a little over two.

The age factors for LQmin are presented on the website: *https://fairmodel.econ.yale.edu/aging/upd2017.htm*. This site can be used for age grading running times. An individual male runner can also use the site to compute how fast he should be slowing down as he ages, assuming that he slows down at the same percentage rate as the age factors imply. Women runners can also use the site under the further assumption that the age factors also pertain to them. Until more data are available for old women runners, using the male age factors is probably the best that one can do.

References

[1] David, H.A., 1970, *Order Statistics*, New York: John Wiley & Sons, Inc.

[2] Einmahl, J.H.J. and J.R. Magnus, 2008, “Records in Athletics Through Extreme-Value Theory,”, *Journal of the American Statistical Association*, 103, 1382–1391.

[3] Fair, R.C., 1994, “How Fast Do Old Men Slow Down?” *The Review of Economics and Statistics*, 76, 103–118.

[4] Fair, Ray C., 2007, “Estimated Age Effects in Athletic Events and Chess,” *Experimental Aging Research*, 33, 37–57.

[5] Gnedenko, B., 1943, “Sur la distribution limite du terme maximum d’une série aséatoire,”, *Annals of Mathematics*, 44, 423–453.

[6] Hill, A.V., 1925, “The Physiological Basis of Athletic Records,” *Lancet*, 209, 481–486.