A Stochastic Frontier Approach to Running Performance

Draft version 4: November 2004

Elmer Sterken

Department of Economics
University of Groningen
PO Box 800
9700 Av Groningen
The Netherlands
e-mail: e.sterken@eco.rug.nl

Abstract

Long-distance running performance depends on age of the runner and race distance. In this paper we apply a stochastic frontier approach to estimate optimal running performance across distances from 5000 meter up to and including the marathon. We present age correction factors for 5000, 10000, 15000 meters, half marathon and marathon distances for women and men. Two conclusions emerge: (1) official age grading tables are too optimistic for older runners, and (2) running performance-age curves differ across distances.

Keywords

Running performance, stochastic frontier estimation, age correction factors.

1 Introduction

One of the interesting fields in the analysis of sports is the impact of age on performance. In this paper we model long-distance average running speed *V* (*V=DIST/T*, where *DIST* is the race distance in meters and *T* is the time to run distance *DIST* in seconds) as a function of age (*AGE*, measured in years) per distance (*DIST*). We do so for observed *record* times of road race events in the U.S., so we model 'optimal' average running speed (since we assume constant speed across during a record race we label our dependent variable by 'average' speed). In aerobic events (such as the marathon) endurance is the most prominent physical characteristic of success. The ability of the body to transport oxygen to the muscles (mostly described by the maximum oxygen intake or VO_{2max}) plays a key role. As the maximum oxygen intake increases when young but decreases at higher ages (see hereafter) the running performance curves are likely to be inverse-U-shaped.

The topic of the relation between running performance and age is not new and attracted some interest in the recent past (see e.g. Fair, 1994, and Sterken, 2003, for reviews). Fair is interested in determining the age where serious decline of the male body starts and uses a special kinked frontier estimation method to estimate the athletic-performance-curve for men over 35 in track and field events. Sterken uses a pooled (across distances and sexes) least-squares model to estimate lifetime curves for both men and women for road race record running speeds. However, road race data are known to be less accurate in measuring distance and could so be subjected to measurement error. Least squares estimation treats outstanding performances as measurement error, which is in conflict

with estimating the maximum performance curves. In this paper we use an alternative econometric method to treat measurement error in road-racing data. We apply a stochastic frontier approach, which estimates an envelope subject to stochastic errors. So outstanding performances are likely to shift the frontier instead of being canceled out as measurement error. The stochastic frontier method is taken from the economic analysis of production functions. Production functions explain output as a function of inputs, like labour and capital. Knowing the frontier one can measure technical inefficiencies: the distances of individual cases to the frontier. In this paper we treat average maximum running speed per age and distance as 'output' and age and distance as 'input' factors. So the main innovation of this paper is the technique used to estimate the human running frontier. Knowing the frontier we can produce age-dependent correction factors that can be used in road-racing events.

The *World Masters Athletes* (WMA), formerly the *World Association of Veteran Athletes* (WAVA), publishes age-group dependent correction factors. One of the serious flaws of the WMA data is the rather arbitrary subdivision of ages into age classes (of five years). One of our goals is to present more precise correction factors per age. The WMA-tables probably also overestimate the capabilities of very old runners (see Fair, 1994). Using our statistical model we simulate “normal” values for average optimal running speed depending on age for various distances (varying from 5000 to 42195 meters).

In other studies like Fair (1994), Grubb (1998), and Sterken (2003) theoretical discussions of the relation between age and average optimal running speed are presented.

We repeat that for long-distance running the ability to transfer oxygen to the muscles (the VO_{2max} or maximum oxygen intake) is the most important variable driving the results of the impact of aging on average running speed. Actual performance seems to co-vary with meta-estimates of the lifetime development of the VO_{2max}. After a certain age the maximum oxygen intake decreases. Although the estimates vary, a yearly decline of 0.5 to 0.9 per cent per year of VO_{2max} seems to be a meta-outcome (see e.g. Rogers *et al.*, 1990). For younger ages the lower figure applies, while the decline is larger for older people. In this paper we will not focus on these theoretical notions again, but concentrate fully on the methodological issues. We use long-distance running data to analyze the impact of aging on endurance, because these data are age-specific and publicly available. We include data on U.S. record times on five distances between 5000 and 42195 meters (the marathon) for ages between 6 and 85 (a more elaborated description of the data is presented in Section 2). In Section 3 we briefly discuss the estimation method used. For more insights concerning the stochastic production frontier estimation we refer to Coelli (1997). We present the model that predicts the optimal running speed based on age and distance in Section 4 and compare the results with previous outcomes. A summary and conclusions are given in Section 5.

2 Running data

We start our analysis by describing the data. The key variable is the average maximum running speed during road race events of distance *i* at the age of the runner *t*. To that purpose we observe actual record times per age and distance. We use data from the U.S.-based Long Distance Running Association (2001). This association keeps records of road racing events on various distances in the U.S. (in this paper data are compiled up to September 29, 2002). For each age the fastest time ever ran by an American citizen is recorded with the name, sex, and date of birth of the runner, as well as the name and date of the race. Moreover, the set contains information on special conditions of the event (downhill track, wind, etc.) and presents corrected data if necessary. The following distances are recorded: 5000, 8045 (5 miles), 10000, 12000, 15000, 16090, 20000, 21097,5 (half marathon), 25000, 30000, 42195 (marathon) and even ultra-long events, such as 24-hour races. The recorded ages vary from 3 to 95 years. It might seem surprising that children of 5 year old run marathons, but the data show that it happens (despite age restrictions at major marathon events). At the extreme ends though the number of runners in those age classes is much lower than in the range 20 to 50 years. Although we cannot observe the underlying distribution of observations per age class (we only observe the best time), it is valid to assume that the density functions are less accurate for very young and old ages. Since this is a potential source of noise, one should take this into account in the remainder of the paper. In order to get rid of the most extreme observations we use the range 6 to 85 years in our sample. We start at the age of 6, because we want to model the fast rate of progress in running speed of children. We

stop at the age of 85: thereafter especially the observations for men decrease seriously. We use data on five popular distances: 5000, 10000, 15000, 21097.5 and 42195 meters. These distances seem to be the most competitive ones, are the least subjected to measurement error (see also Sterken, 2003), and appeal mostly to the idea that running performances have an impact on the shape of the human running frontier. The record times are transformed into average running speed in meters per second and it is implicitly assumed that the speed during the record event is constant. Finally we take natural logs of the average running speed. Taking the logs of the speed improves the fit of the model and is consistent with the assumptions on the estimation of the (exponential) stochastic production frontier model.

Insert Figure 1 about here

Figure 1 gives an overview of the nature of the data. We plot the natural log of the average running speed (vertical axis in logs of meters per second) of men of various ages (horizontal axis) for the marathon distance (one of the five different events included). So for each age we plot the highest running speed for actually observed data and estimated frontiers. This figure shows three lines: V for the raw data and VOLS for a least-squares approximation and FRONTIER for the newly suggested stochastic frontier approach. Figure 1 reveals a pattern that can be observed in the plots of age-dependent velocities of other distances for both women and men. At younger ages one can observe a steep slope and speed increases substantially in age. Around the ages 26-30 the highest running

speeds are observed. After the age of 30 a decrease in running speed occurs and men over 70 are confronted with a more serious loss of performance.

There are various reasons to believe that measurement error is present in the data presented in Figure 1. Individual strong performance can distort the general pattern. Moreover, for very young and very old ages there are relatively few observations per age class, which increases the (unobserved) variance of the distribution of record times per age. Since we are interested in the general development of human endurance we want to analyze series without measurement error. We can solve the measurement problem in two ways. First, as is done in Figure 1, we can smooth the individual series and compute “normal” values using least-squares estimation (this is the approach followed in Sterken, 2003). Sterken estimates pooled male and female average running speed using 12 parameters to model speed in 22 distances (11 for male and female runners). Using relatively few parameters makes the model parsimonious, but throws away information especially available in the tails of the series. Moreover, averaging out measurement error does not appeal to approaching 'ultimate' performance. Therefore we use the stochastic frontier approach in this paper. As one can see in Figure 1 this leads to an envelope that smoothly touches on the top of the functional form. The stochastic approach allows still some major peaks to be considered to be stochastic outliers. Especially at the younger and higher ages measurement error might still occur. This approach seems to be preferred if one wants to calculate 'ultimate' relative age performances.

3 The stochastic frontier approach

The stochastic frontier production function is independently proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). The original specification involved a production function specified for cross-sectional data, which had an error term with two components, one to account for random effects and another to account for technical inefficiency. This model can be applied to our data and be expressed in the following form:

$$y_i = x_i\beta + (V_i - U_i) \qquad i=6,...,85, \qquad (1)$$

where y_i is the log of the average running speed of the best performing runner of age i, x_i is a $k\times1$ vector of (transformations of the) input quantities of the i-th runner (e.g. age or higher-order age terms), and β is a vector of unknown parameters. V_i is a random variable, which is assumed to be i.i.d. $N(0,\sigma_V^2)$, and independent of U_i, which is a non-negative random variables and assumed to account for inefficiency in running speed (often assumed to be i.i.d. $|N(0,\sigma_U^2)|$). Note that the assumption of homoskedasticity of the distribution of error terms across age classes might not be valid. Given the differences in the theoretical number of observations per age class, it would be more valid to assume heteroskedasticy. The general specification can be changed and extended in a number of ways. These extensions include the specification of more general distributional assumptions for the U_i, such as the exponential, the truncated normal or two-parameter gamma distributions; the consideration of panel data (see hereafter) and time-varying

technical efficiencies; the extension of the methodology to 'cost' functions and also to the estimation of systems of equations. We decided to use the exponential distribution for U_i to estimate models per distance for men and women, because this allows a direct comparison with the results obtained with the panel model described below. We apply model (1) in estimating our model per distance, which gives us full flexibility in determining parameters per distance and sex.

Apart from estimating a model per distance we can combine various distances in one model (and so reduce the number of parameters to be estimated) by employing a panel data approach. Especially if one believes that there is heterogeneity in the intercept of various cross-section units, but homogeneity in the dependence on the determining variables, e.g. age in our case. So if the shape of Figure 1 holds for all distances, one could keep the functional form constant, while allowing for different levels. The main advantage of a panel data model is the drop in the number of parameters to be estimated. The panel model reads:

$$y_{it} = x_{it}\beta + (V_{it} - U_{it}) \qquad i = 6,\ldots,85,\ t=1,\ldots,5, \qquad (2)$$

y_{it} is the logarithm of the average running speed (in meters/second) of a runner of age i on distance t. x_{it} is a matrix of higher-order terms of distance (in kilometers) and age (in years). The V_{it} are random variables, which are assumed to be i.i.d. $N(0,\sigma_V^2)$, and independent of the $U_{it} = U_i \exp(-\eta(t-5))$, where the U_i are non-negative random variables which are assumed to account for technical inefficiency in production and are assumed to

be i.i.d. as truncations at zero of the $N(\mu,\sigma_U^2)$ distribution. η is a parameter to be estimated. There are obviously a large number of model choices that could be debated. For example, does one assume a half-normal distribution for the inefficiency effects or the more general truncated normal distribution? Should one assume time-invariant or time-varying efficiencies? Are inefficiencies identically distributed across ages? As discussed above, in our specific model it is likely to assume that the distribution of errors for extreme ages differs from the distribution of ages between e.g. 20 and 40 years. In the literature it is recommended that a number of alternative models should be estimated and that a preferred model is to be selected using various fit criteria (e.g. likelihood ratio tests). We experimented with various model specifications and report the best fit (using the exponential distribution for U_i) here. Note that we have no degrees of freedom to estimate age-dependent variances of the inefficiencies, so we are not able to address the heteroskedasticity issue. We employ the computer program FRONTIER Version 4.1 (see Coelli, 1997), which we use to obtain maximum likelihood estimates of the stochastic frontier. The program also provides indications of the technical efficiency of each of the data points.

In Section 4 we use both approaches. After a simple least-squares polynomial estimation to get a first impression of the shape of the curves we estimate a panel model (for men and women separately) to get a first idea of the shape of the contours. Next, we estimate models per distance and allow for heterogeneity in the functional form. Again, although we can accept the pooled model on statistical grounds, we still have an interest in improving the fit of the curves per distance. Knowing the maximum likelihood parameter

estimates, we generate the running frontier and use the frontier to compute age correction factors. We use the highest speed on the 5000 meters as the normalization value. This implies that all results are expressed in terms of the relative position to the fastest performance on the 5000 meters.

4 Results

The dependent variable in our models is the natural logarithm of average running speed in meters/second on the five distances (from 5000 to 42195 meters). We link average running speed with age and eventually with distance. In order to get a rough idea of the shape of the contour we first use simple least squares estimation of average running speed as a function of a higher-order polynomial of age for various distances. Experimenting with the maximum order of the polynomial illustrates that the shape of the frontier can be approximated by a fourth-order polynomial in age (see Figure 1 for an example for the male's marathon). Given the fact that the shape of the relationship between average speed and age is rather similar for various distances we can first combine the distances into one model too get a rough idea of the shape of the contour of the stochastic frontier. So, in the following we present the results of two model classes. First, we estimate the pooled model (across the five distances) for men and women separately. Because we use fully deterministic explanatory variables with limited variation, estimating a stochastic frontier model will be mainly concerned with estimating the intercept. Therefore we focus the discussion on the relative change in the intercept if we compare the outcomes of the stochastic frontier estimation with the previously obtained least squares results. So we estimate a fourth-order polynomial model in age and included moreover a second-order polynomial in distance and age and distance cross-terms. The latter allows for a differential impact of age on various distances. We estimate this model for both men and women separately. It should be noted that the higher-order terms show multicollinearity with the lower-order terms, which renders interpretation of the signs and magnitudes of

the parameters less useful. Our interest though is in the total fit of the model. In Table 1 we present both the Ordinary Least Squares results and the Maximum-Likelihood Stochastic Frontier adjustment for both men and women. The parameter estimates (except for the intercept) are equal for the OLS and Stochastic Frontier routines. Table 1 shows that the fit of the model is comparable across men and women. We exclude the insignificant cross-term age times the cubic distance in the model for women. For men the intercept increases using the stochastic frontier routine by 1.014% to 1.036 and for women by 0.598% to 1.247. This implies that averaging out the disturbances by OLS leads to an underestimation of the maximum running speed by 0.5 to 1% of the speed. The table moreover shows a slightly higher technical efficiency for women. Technical efficiency is an average (over various ages, 6 to 85) of individual relative performance to the frontier. Interpreting average technical inefficiency can be troubled by the assumption of homoskedasticity of the residuals across ages. As argued before, it is likely that the estimated variances of the error terms of the stochastic frontier models will vary over ages. Since we are unable to identify age-specific variances, the estimates of technical inefficiency should be considered with care. Inspection of the estimated inefficiencies per age reveals that these are rather constant across ages (for all distances for both men and women). With respect to the shape of the functional form the invariance of the parameter estimates to the estimation method implies that age correction factors are not influenced by pooled stochastic frontier estimation (as compared to OLS estimation), because age correction factors are relative figures.

Insert Table 1 about here

Inspection of the residuals though reveals systematic underestimation of the half-marathon and marathon running speed for ages 45 to 75. Therefore we proceed in estimating models per distance. Table 2 gives the fourth-order polynomial estimation results per distance. All parameters are significant at the 1% confidence level. As one can observe the parameters for the male's marathon differ to some extent from the other estimates. Table 2 moreover shows considerable differences between the intercept adjustments across distances and sex. Especially the men's half marathon and marathon get rather substantial intercept adjustment. Moreover, it seems that especially the half marathon suffers from relatively low mean technical efficiency. This justifies an approach per distance. Therefore we conclude that Table 2 provides the relevant information to construct age correction tables.

Insert Table 2 about here

It seems that pooled models (models that combine various distances) under-parameterize the functional forms per distance. Moreover, some distances seem to have higher frontier-intercepts than others. We will illustrate this notion by transforming the rather abstract Table 2 into a table with age correction factors and compare the results found with the ones presented in the WMA-tables, Fair (1994) and Sterken (2003). Transforming the data into age correction factors moreover has the advantage of reducing the influence of specific cases (like our US road-racing case) in generalizing to global facts.

Table 3 presents age correction factors for men and women for the 5000 meters and marathon respectively (correction factors for the other distances can be found on www.eco.rug.nl/medewerk/sterken). Table 3 shows that male runners should be able to run the marathon at 91.3% of their 5000 meters average speed. The world's fastest road race time on 5000 meters is 13:00. This would imply a fastest marathon time of 2:00:07. This is still rather far from the current world best marathon time 2:04:55. Our estimate of 91.3% also exceeds the 87.7% estimate in Sterken (2003). This fact is due to the approach per distance as explained above. With respect to ageing we observe that a 70-year old male runner should be able to reach 72.8% of his maximum performance on the 5000 meters. This is in line with Fair (1994) and Sterken (2003) and so below the correction factor used by the WMA (73.7%). For older ages our results show that the WMA-times are overoptimistic. For the marathon we find that a 70-year old runner should be able to reach 70.5% of his maximum marathon performance. This is again below the WMA-value (75.4%). In general, our current results are more optimistic for older men on longer distances than Sterken (2003), but still point at lower levels of result as compared to the WMA-tables.

Insert Table 3 about here

Table 3 shows similar results for women. A female runner is able to perform at about 88.3% on the marathon, as compared to the maximum 5000 meters average running speed. This result is in line with Sterken (2003). According to the WMA a 70-year old women should be able to run at 69.7% of her maximum speed on the 5000 meters. We

find 68.2% for his figure, while for the marathon we get 67.7% instead of the 71.8% by the WMA. Again, our new approach favors older women a bit more than in Sterken (2003), but still cannot meet the quite optimistic WMA-factors.

5 Summary and conclusions

In this paper we model average maximum running speed on distances from 5000 meters to the marathon for men and women. We use a new approach to model running speed: stochastic frontier estimation. This method appeals more to the idea of ultimate record running performance by estimating an envelope to raw data than simple OLS-estimation. We show two sets of results: a pooled model (combining the data various distances) and models for each separate distance. There are clues to underparameterization of the pooled model, so we prefer to use the models per distance. It seems that there is relatively large inefficiency in the half marathon data (and male marathon data) in the US case. We present estimates of technical efficiency. We also note that, due to the assumption of homoskedastic error variances, we should be careful in interpreting average technical inefficiency. The nature of our data, fewer (unobserved) data for extreme ages, likely supports the assumption of heteroskedasticity (which we are not able to tackle). We use the models per distance to compute age correction factors. These factors give insight into the impact of aging on performance and can be used in road-race events. The results show that official correction tables tend to be overoptimistic, especially for older male and female runners.

References

Aigner, D.J., Lovell, C.A.K. and Schmidt, P., 1977, Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, **6**, 21-37.

Coelli, T.J., 1997, A guide to FRONTIER version 4.1: A computer program for stochastic frontier production and cost function estimation, Center for Efficiency and Productivity Analysis, University of New England, Armidale, Australia.

Fair, R.C., 1994, How fast do old men slow down, *Review of Economics and Statistics*, **76**, 103-118.

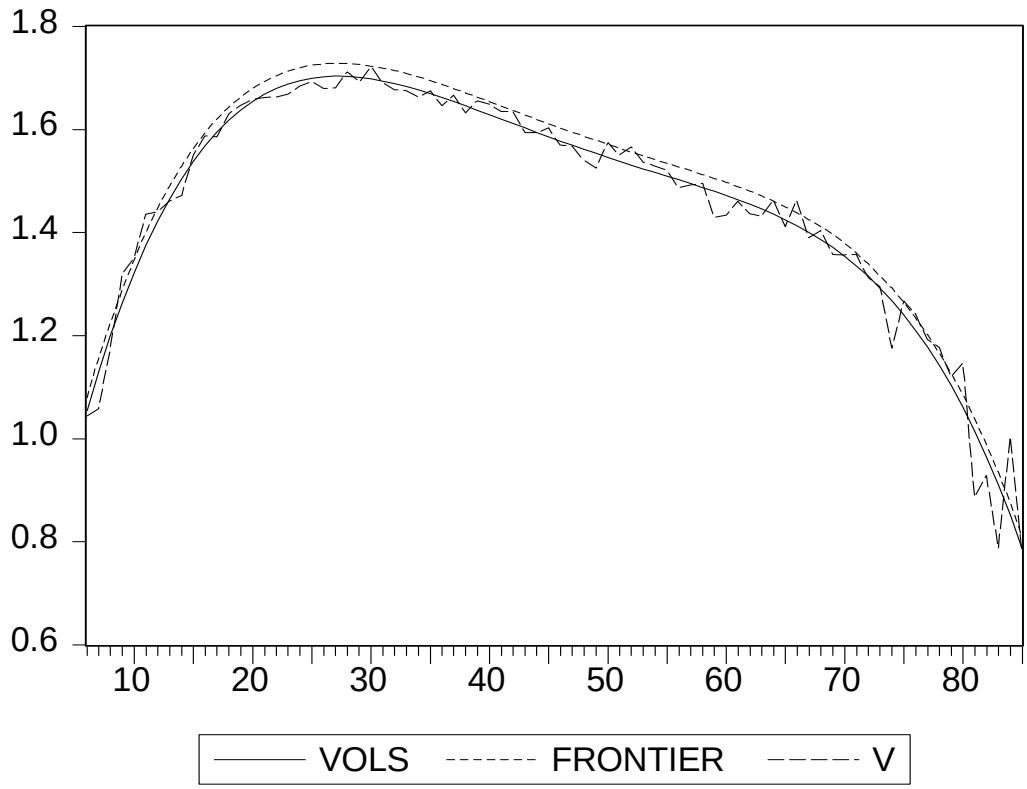
Grubb, H.J., 1998, Models for comparing athletic performances, *The Statistician*, **47**, 509-521.

Meeusen, W. and van den Broeck, J., 1977, Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review*, **18**, 435-444.

Rogers, M.A., Hagberg, J.M., Martin III, W.H., Ehsani, A.A., and Holloszy, J.O., 1990, Decline in VO_{2max} with aging in masters athletes and sedentary men, *Journal of Applied Physiology*, **68**, 2195-2199.

Sterken, E. 2003, From the Cradle to the Grave: How Fast Can We Run?, *Journal of Sports Sciences*, **21**, 479-491.

Figure 1 - Average running speed on the marathon per age (male observations)



V = natural logarithm of average running speed for various ages from 6 up to and including 85 (denoted on the horizontal axis);
VOLS = fitted running speed from OLS estimation (see Table 2 for full results);
FRONTIER = estimated stochastic frontier of running speed;

Source of the data: www.usaldr.org

Table 1 - Average running speed as a function of age and distance: OLS-results

	Men		Women	
	Parameter	**t-value**	**Parameter**	**t-value**
AGE	0.086	32.268	0.046	15.260
$AGE^2/100$	-0.299	-27.842	-0.134	-11.026
$AGE^3/10000$	0.411	23.860	0.146	7.524
$AGE^4/1000000$	-0.207	-21.956	-0.064	-6.037
DIST	-0.015	-9.785	-0.013	-13.787
$DIST^2/100$	0.014	4.567	0.013	8.252
AGE*DIST/100	0.036	8.974	0.018	5.835
AGE^2*DIST/1000000	-3.131	-10.530	-1.874	-5.584
AGE*$DIST^2$/1000000	-1.657	-2.847		
C	1.026	40.626	1.239	47.897
Stochastic frontier C	1.036		1.247	
Relative increase	1.014		0.598	
Mean efficiency (%)	98.977		99.264	
R-squared	0.965		0.959	
SSR	0.524		0.669	

Notes to the table:
- Dependent variable: natural logarithm of running speed (meters.second^{-1})
- Number of observations: 400. The data include five distances (5000, 10000, 15000, 21097.5 and 42195 meters) for the ages 6 up to and including 85.

AGE = age (years);
DIST = distance (1000 meters);
C = intercept;
Stochastic frontier C = intercept estimated by the stochastic frontier routine;
Relative increase = 100*[(Stochastic frontier C/C)-1];
Mean efficiency = average (over ages) technical efficiency (percentages);
R-squared = adjusted R-squared;
SSR = sum of squared residuals.

Table 2 - Average running speed as a function of age

	Men				
	5k	**10k**	**15k**	**21.0975k**	**42.195k**
AGE	0.081	0.080	0.081	0.091	0.125
$AGE^2/100$	-0.275	-0.267	-0.258	-0.293	-0.432
$AGE^3/10000$	0.377	0.358	0.329	0.380	0.612
$AGE^4/1000000$	-0.193	-0.179	-0.157	-0.186	-0.319
C	0.993	0.964	0.893	0.757	0.449
Stochastic C	0.998	0.982	0.923	0.802	0.473
Relative increase	0.452	1.922	3.315	5.902	5.551
Mean efficiency	0.995	0.982	0.971	0.958	0.976
	Women				
	5k	**10k**	**15k**	**21.0975k**	**42.195k**
AGE	0.032	0.036	0.042	0.069	0.066
$AGE^2/100$	-0.070	-0.098	-0.104	-0.219	-0.194
$AGE^3/10000$	0.046	0.098	0.079	0.280	0.229
$AGE^4/1000000$	-0.012	-0.042	-0.021	-0.139	-0.107
C	1.263	1.230	1.104	0.858	0.812
Stochastic C	1.268	1.234	1.138	0.890	0.817
Relative increase	0.383	0.348	3.076	3.683	0.623
Mean efficiency	0.995	0.996	0.967	0.969	0.995

Notes to the table:
- Dependent variable: natural logarithm of running speed (meters.second^{-1}).
- Number of observations per distance: 80, for ages from 6 up to and including 85.

AGE = age (years);
C = OLS-estimate of the intercept;
Stochastic C = Stochastic Frontier estimate of the intercept;
Relative increase = 100*[(Stochastic C/C) -1];
Mean efficiency = mean efficiency estimate of the individual running performance to the estimated stochastic frontier.

Table 3 - Age correction factors

Age	Male - 5k	Male - Marathon	Female - 5k	Female - Marathon
10	0.778		0.837	
12	0.831	0.690	0.867	0.717
14	0.877	0.748	0.894	0.756
16	0.915	0.798	0.918	0.789
18	0.946	0.839	0.940	0.817
20	0.969	0.870	0.958	0.840
22	0.985	0.892	0.972	0.858
24	0.995	0.906	0.984	0.872
26	1.000	0.912	0.992	0.880
28	0.999	0.913	0.998	0.885
30	0.995	0.908	1.000	0.886
32	0.988	0.900	0.999	0.884
34	0.978	0.889	0.996	0.879
36	0.967	0.876	0.990	0.871
38	0.954	0.862	0.982	0.862
40	0.940	0.848	0.972	0.850
42	0.927	0.833	0.960	0.838
44	0.913	0.819	0.946	0.824
46	0.899	0.805	0.930	0.809
48	0.886	0.792	0.913	0.794
50	0.872	0.780	0.895	0.778
52	0.859	0.769	0.876	0.762
54	0.847	0.758	0.856	0.746
56	0.834	0.747	0.835	0.729
58	0.822	0.736	0.814	0.712
60	0.808	0.725	0.792	0.694
62	0.795	0.712	0.770	0.676
64	0.780	0.699	0.748	0.658
66	0.764	0.683	0.726	0.640
68	0.747	0.665	0.704	0.620
70	0.728	0.644	0.682	0.600
72	0.706	0.619	0.660	0.580
74	0.681	0.591	0.638	0.558
76	0.654	0.558	0.617	0.536
78	0.623	0.521	0.596	0.513
80	0.590	0.480	0.575	0.488
82	0.553	0.436	0.555	0.463
84	0.513	0.389	0.535	0.437

See for the other ages and distances (10k, 15k and the half marathon): www.eco.rug.nl/medewerk/sterken