LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

VOLKSWIRTSCHAFTLICHE FAKULTÄT

John Komlos:

How to (and How Not to) Analyze Deficient Height Samples

Munich Discussion Paper No. 2003-12

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Online at http://epub.ub.uni-muenchen.de/56/

How to (and How Not to) Analyze Deficient Height Samples

John Komlos

Department of Economics

University of Munich

Ludwigstraße 33/IV

D-80539 Munich, Germany

Telephone: +49-89-2180-5824

+ 49-89-8983-9700

Voice Mail: +49-89-2180-3169

Fax: +49-89-33-92-33

email: jk@econhist.de

How to (and How Not to) Analyze Deficient Heights Samples

The systematic study of the physical characteristics of human beings reach back well into the eighteenth century (Tanner, 1981). By the 1830s Adolphe Quetelet and Louis R. Villermé recognized that biological outcomes, such as physical stature were influenced both by the natural, as well as the socio-economic environment (Villermé, 1829; Quetelet, 1831). However, until the late 1960s, when French historians of the Annales tradition began to explore the socio-economic correlates of human height, the topic was of interest primarily to scholars in sister disciplines such as anthropology, medicine, or military history (Le Roy Ladurie, Bernageau, and Pasquet, 1969). The expansion of anthropometric history was brought about beginning in the mid-1970s by American cliometricians (quantitative economic historians), in search of a new measure of well-being in historical populations during the course of the last two centuries (Harris, 1994; Komlos, 1995; Steckel, 1995; Komlos and Cuff, 1998).

Anthropometric historians are generally interested in the changes in mean height of a population over time, insofar as that provides an indication of how well the biological organism was able to thrive in its socio-economic and epidemiological environment. From trends one can infer changes in income, in the variability of income, in the distribution of income, or how well self-sufficient peasants were able to care for their children. They provide an overall indication of access to nutrients, influenced by the relative price of nutrients. In addition, measures of social, gender, and spatial differences in nutritional status can be obtained. This is a considerable advantage insofar as height data are plentiful for segments of the population at the regional level for a time when conventional indicators of economic welfare are sparse.

New windows were opened and new vistas discovered. We now know that during the pre- and early industrial period the biological standard of living depended on such socio-economic factors as the level, variability, and distribution of income, as well as on the relative price of

1

nutrients, particularly of protein. Urbanization and the degree of commercialization of the economy also had an impact on the human growth process. We also learned that prior to the onset of modern economic growth, people who were self-sufficient in food production, were living on productive land and in regions with low population density (i.e. relatively removed from urban markets and their disease pools) tended to be relatively tall even if they were poor in conventional terms (Komlos, 1998). Propinquity to nutrients invariably conferred considerable nutritional advantages in the early-industrial period vis-à-vis urban populations, and those engaged in industrial activity prior to the emergence of refrigerated trucks and railroad cars (Cuff, 1998; Craig and Weiss, 1998; Haines, 1998). Physical stature declined at the onset of modern economic growth in the then developed world, and did not begin to improve substantially until well into the second half of the 19th century. In marked contrast, the secular trend in stature in OECD countries in the 20th century was less subject to cyclical downturns, except during wars. The reason is that markets in food products became better integrated, so that local shortages in Europe and North America were quickly alleviated. Government expenditures on public health and medical care increased sufficiently to make a major impact on the biological well being of the population. Children's work declined or was entirely eliminated, freeing up calories for the growth process. Welfare programs increased, so that the effect of short term fluctuations in income had a negligible effect on children's heights (Komlos, 1996).

Yet, the aim of this paper is not to provide an overview of the accomplishments of anthropometric history. Rather, it is to offer some basic guidelines on how to analyze deficient height samples. With two to three decades of experience behind us, it should be useful to share our understanding of current practice. To be sure, in many cases one can apply standard procedures to estimate trends and cross-sectional patterns using ordinary least squares (OLS) linear regressions analysis. Such institutions as prisons, passport agencies, and armies with universal conscription did

not impose a height requirement (HR) on entrants, and, therefore, the samples drawn from their records are representative, in the main, of the underlying population from which the institution drew its members. In such cases one can obtain a random sample of heights representative of the universe of observations upon which the archival information is based.[1] Such samples are in principal unbiased, and standard procedures of analysis apply, even if the universe might be limited by gender or confined to certain social classes.[2]

However, the analysis of height samples is frequently markedly different from the study of other data insofar as height is the outcome of a biological process, and therefore conforms to certain biological laws - not discussed in this paper except to mention that the adult height of a population is (approximately) normally distributed, because this fact will be used in the subsequent analysis (Tanner, 1978; Bogin, 1999).

<u>Deficient Samples</u>

Height data often stem from institutions that imposed a HR as a precondition of acceptance so that some members of the population had a lower probability of gaining entrance. Samples drawn from such records are perforce deficient i.e., incomplete, inasmuch as a substantial portion of the underlying population's height distribution is unavailable for analysis. Obviously, samples drawn from such records are not representative of the universe of observations: people below or above a certain threshold are either missing completely, or are underrepresented. There might be a minimum HR on the left side of the distribution (HR_{min}, at τ_m,), or on the right side, due to a maximum height requirement (HR_{max}, at τ_x) (or both). HR_{min} are most common, but HR_{max} were also imposed, if being much taller than average was a disadvantage, as aboard ships, because of the high center of gravity. Estimating procedures devised to correct for a bias caused by the HRs are based on the biological fact that the height distribution of a (homogeneous) population is (approximately) normal $N : (\mu, \sigma)$.

A complicating factor is that the HRs varied over time, and were enforced with varying stringency: in effect, the records went through a (time-variant) filter: people shorter than τ_m, or taller than τ_x, had a lower probability of entering the sample. Such distributions have a "shortfall" beyond the HRs, and are thus inherently biased (Wachter, 1981). Shortfall implies that the sample is not random. The amount of shortfall is the (unknown) percent of missing observations in the sample beyond the HRs, i.e., the percent of people who were not accepted into the institution because of the HRs. Without proper statistical adjustments, the analysis of such deficient samples are usually unreliable, misleading, or even meaningless[3] (Komlos, 1993; Heintel and Baten, 1998). (See Appendix A for definitions and abbreviations.) In the presence of shortfall the sample is normally distributed only within the range (τ_m, τ_x).[4] The height distribution (histogram) of 11,000 adult French soldiers recruited prior to 1740 is depicted in Figure 1. An erosion of the distribution below 62 French inches (F.i.) is quite apparent, implying the existence of shortfall. Note, however, that the truncation is not perfect, indicating that the HR_{min} was not applied consistently.[5] This is the typical pattern in armies that did not have universal conscription. An example of the effect of HR_{max} is shown in a sample of poor boys admitted into the Marine Society of London, a charitable institution, between 1792 and 1798 (N=559) (Figure 2). The institution prepared boys for life at sea, did not have an official HR_{max}, but obviously an informal one.[6]

In order to discover such deficiencies in the sample, the very first step in any analysis of height should be a visual inspection of the sample histograms. This is mandatory before proceeding further, even if one has reason to think that HRs did not exist, because informal ones were often enforced, and by examining such evidence one gains a sense of the degree of shortfall, of rounding, and other possible sample deficiencies. Note that heights were not always recorded with great precision, particularly in military samples. Observations were often rounded to a nearest unit (e.g., inch or cm). Heaping on whole, or on even numbers, or numbers divided by 5 is often evident.[7]

However, symmetric rounding (heaping) does not introduce appreciable systematic error into the analysis.[8] The distribution within the range (τ_m, τ_x) should be approximately normally distributed, but heaping on favorite numbers might well produce samples which deviate from this general rule.[9]

<u>Height Distributions</u>

The primary aim of the examination of the histograms is to identify the actual HRs for the period under consideration, and if they changed over time, because one needs the de facto, and not the prescribed HRs for further analysis.[10] In creating the histograms for initial inspection, the height measurements must be left in the original units, because conversion into another metric system can introduce distortions.[11] Samples from different military units, i,e. cavalry, artillery, navy, infantry, and ancillary troops should not be conflated, because different units generally had different HRs, and the combined sample would become a mixture of truncated normal distributions for which estimating procedures are more complicated.[12] In historical populations those who were older than 23-years old have reached their final height and can therefore be considered adults. The height of younger soldiers should be analyzed separately, because the time during which their growth could be influenced by environmental circumstances are not perfectly coincident with those of adults.[13] Insofar as height begins to diminish after age 50 – those above this age should not be considered in the analysis.

In samples covering over an extended time period, several histograms should be produced for various time periods. Histograms by decades are desirable if there are sufficient number of observations, but longer intervals are appropriate otherwise. It is usually necessary to separate peace-time from war-time recruiting, because HRs were frequently overlooked during times of need.[14] The dates at which legal changes were made in the conscription laws are also obvious choices for periodization.[15] For this part of the analysis the relevant date is that of recruitment, and not that of the time of birth, because HRs changed for recruitment years and not for birth years.

There should be about 500 records for each histogram in order to minimize variations due to small sample size.[16]

Only after the HRs are determined, can the actual process of analysis commence. The aim is to estimate: the height of the population from which the soldiers were drawn, μ, and the effect of covariates (independent variables) on height including time, i.e., the trend.[17] Usual explanatory variables available for military samples include age, date of birth, region of birth and of residence, urban/rural provenance, date of enlistment, and occupation. Physical stature is affected by many socio-economic variables. These include, but are not limited to the epidemiological environment – hence the inclusion of mortality rates in the region of birth or of residence is a legitimate explanatory variable. Social stratification is also a crucial determinant of height insofar as income determines the budget constraint. Education (literacy) has an effect, because better educated parents have superior consumption skills, are better informed about long-range health effects of consumption patterns, and, thus, are usually able to take better care of their off-springs. Education also correlates positively with income. Height is a function of income inasmuch as the consumption of nutrients, particularly of proteins, vitamins, and minerals, and the regularity with which those nutrients are consumed, all influence height at a particular age until adulthood. Urban/rural differences (population density) are also useful predictors of height insofar as the disease environment and medical services vary spatially. Population density also determines the speed and regularity with which disease vectors are transmitted throughout the population. From the region of birth many ecological variables might be inferred, depending on collateral information, such as the distribution of income. Occasionally some data might be available on the parents as well.[18] All of these variables can be useful, though usually a small percentage of the variation in height can be explained at the individual level, because most of the individual variation in height is genetic.[19]

Estimation Procedures on Deficient Samples

Four estimating procedures, available for analyzing normal distributions with shortfall, are discussed below: the Quantile Bend Esimator (QBE), the Komlos-Kim method (K&K), Truncated Ordinary Least Squares (TOLS), and Truncated Regression (TR) using maximum likelihood procedure (See Appendix B).[20]

The QBE was the first method proposed to estimate m, the true mean height of the underlying normally distributed population from a sample with shortfall: theoretically $E(\hat{m}_{QBE})$ $= m$[21] (Wachter and Trussell, 1981, 1982). The algorithm estimates the amount of shortfall by filling in the missing observations until the sample distribution becomes normal (Gaussian).[22] In contrast, the first step for the other three methods is to discard all observations outside of the range (τ_m, τ_x) in order to equalize the bias across the whole sample. The K&K method is the simplest: calculate the mean of the part of the sample remaining after truncation, i.e., after discarding the portion outside of the range (τ_m, τ_x).[23] While $\hat{m}_{K\&K} \neq \mu$, the direction of the trend of the truncated mean using the K&K method is also the direction of the trend in μ, the mean height of the population from which the sample originates.[24]

The usual approach using ordinary least squares regression (OLS) is unsuitable (biased) in case of samples with shortfall, and should not be used on such data in any circumstances whatsoever. Two regression methods are appropriate to analyze the correlates of height in case of truncated normal distributions: TOLS and TR.[25] The easier approach is the TOLS, an extension of the K&K method to estimating the effect of covariates of height. TOLS can be performed with most statistical computer programs, such as SPSS. This is not the case with TR, which is confined to a more limited set of statistical packages such as STATA (2001, p. 209) and EViews (2000, p. 438). TOLS is simple OLS regression analysis after the data outside of the range (τ_m, τ_x)

have been eliminated. Clearly, the coefficients obtained with TOLS are biased, but their relative sizes, as well as their signs are correct.[26] Note, however, that accurate statistical inference (confidence intervals, hypothesis testing) is not possible. For this purpose the fourth method, the TR, using maximum likelihood estimation, is required.[27]

The QBE is the only one of the four methods that does not need information on the HRs, and as K&K, it cannot estimate the impact of covariates on height; moreover, it cannot be used on doubly truncated samples, whereas the other three methods can. However, it was clear from the very beginning that $\hat{\mathbf{m}}_{QBE}$ was inaccurate (Komlos 1985, 1989).[28] Simulations have confirmed that $\hat{\mathbf{m}}_{QBE}$ is an inefficient estimator of μ, and might, in fact, be biased. Moreover, it is unambiguously it is inferior to all three other procedures (Table 1). This conclusion is derived from an experiment which used twelve specifications with various sample sizes and varying amounts of shortfall (Heintel, 1996a). The 12,000 simulations indicate that the average bias $\left| \hat{\mathbf{m}}_{QBE} - \mathbf{m} \right|$ of the QBE estimates was twice as large, and its mean square error (σ^2 + bias^2) was three times as large, as that of the TR estimator[29] (Table 1). Another set of experiments using discreet samples, yielded similar results, and the bias of the QBE estimates was even larger (Table 1). In sum, $\hat{\mathbf{m}}_{QBE}$ is inefficient and in practice the estimates vary implausibly from one decade to another.

For all of the other three methods observations outside of the range (τ_m, τ_x) are discarded. As a consequence, the height distributions have to be examined carefully and if (τ_m, τ_x) changed over time, then one has to consider which of the (τ_m, τ_x) to use. For the TR procedure this is not a problem, insofar as it can be used with several sets of (τ_m, τ_x) simultaneously; for the K&K and the TOLS methods, however, the largest τ_m ($\mathbf{t}_m^{max}$), and the smallest τ_x ($\mathbf{t}_x^{min}$) are the relevant ones to use for truncation, if (τ_m, τ_x) changed over time. These then become the effective (τ_m, τ_x) for further analysis. In subsequent notation we drop the superscript notation on the τ's and assume that

8

they are the appropriate ones for the particular analysis. For instance, in the case of the French army, τ_m was lowered from 62 to 60 French inches between 1740 and 1762.[30] Hence, with K&K and TOLS, one would choose the larger of the two, i.e. 62 French inches, as the binding τ_m for the whole period 1716-1789.[31] Within the range (τ_m, τ_x) the distribution of heights has been conditioned to "have identical biases", and can be used for further analysis. In other words, we make certain that the biases are the same throughout the period under consideration. The observations outside of the range (τ_m, τ_x) are then not used for further analysis.[32]

Ascertaining accurate trends of physical stature is an important issue in anthropometric history, and, consequently we proceed to estimate the probability of correctly estimating whether the difference between the estimated means of two samples has (at least) the correct sign. Consider two samples drawn from two populations $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$. Discard a percentage a_1 and a_2 of the sample to the left of τ_{m1}, τ_{m2}. Then calculate the probability that using the three methods, TR, QBE, and K&K we estimate the direction of the trend correctly. That is to say, in how many cases out of 1,000 is it true that if $\mu_1 < \mu_2$ then $\hat{m}_{1,K\&K} < \hat{m}_{2,K\&K}$ or $\hat{m}_{1,QBE} < \hat{m}_{2,QBE}$ or $\hat{m}_{1,TR} < \hat{m}_{2,TR}$. The results of 20,000 simulations indicate that K&K method is the most robust, yielding a probability of 0.88, the TR is second with 0.72, followed by the QBE with 0.66 (Table 2) (Heintel 1996b).[33] In other words, the QBE procedure is not able to estimate even the direction of the trend correctly about one-third of the time.[34] Heintel (1996, p. 21) also shows that the relative accuracy of the K&K procedure increases with larger changes in mean heights.[35] Hence, by all measures, the QBE is unacceptably unreliable, and should not be used as an estimator of either the true population means, or of trends, insofar as more accurate procedures are available.

In contrast, the K&K method is accurate and can be calculated easily with all statistical programs.[36] While it is suitable for estimating only the direction of the trend, and not the mean height of the population from which the sample was drawn, it can be used as a preliminary exercise

in order to obtain an accurate impression of trends over time. Insofar as it cannot be used to estimate correlates of height, the sample first has to be subdivided in order to ascertain the trends in height for particular subgroups, by geographic provenance, for example.[37] For the K&K method to be more accurate than the TR, one needs one of three conditions to hold: a) σ remain constant; or b) if σ is not constant, it changes in the same direction as the mean height, μ;[38] or c) if σ and μ move in opposite direction, σ should not change by more than about six percent (about 4 mm) (Heintel 1996b). If the true heights increase and at the same time σ declines by about 4 mm the TR becomes a superior method. These are weak restrictions, however, insofar as σ tends to remain relatively stable over time, even over centuries, and even as mean heights change considerably. In sum, the K&K method is a good first step in the analysis. It is not foolproof, though, and it is important to supplement it with regression methods (such as TOLS or TR) that are able to explore the effect of a vector of covariates (i.e., age, time, birth place, socioeconomic environment) on heights, provided such data are available.

The advantage of TOLS is its ease of application – almost as convenient as the K&K method, but in contrast to K&K, it also estimates the relative effect of the covariates on heights. The estimates $\hat{b}_{TOLS}$ are first converted into (truncated) height estimates. For example, in order to estimate the height of French soldiers who have not yet reached adult height, dummy variables were included for ages younger than 22. The constant (171.62 cm) then pertains to 22-year-old soldiers: $\hat{m}_{TOLS(age\,=22)}$=171.62 cm, and insofar as the estimated coefficient of 20-year-old soldiers is –0.44 cm, this implies that 20-year-old soldiers above the MHR were shorter than 22-year-olds by this amount. In turn, this coefficient estimate can then be converted into estimated height of 20-year old soldiers: $\hat{m}_{TOLS(age=20)}$=171.62–0.44= 171.18 cm, but $E(\hat{m}_{TOLS})$ ^{1} m. However, these heights can be converted into estimates of the mean height of the sub-populations from which the soldiers

were selected, even if the conversion is relatively complicated, and as a consequence, if available, the TR method is more effective in providing true population estimates.[39]

Though TR is often a bit more difficult to implement in practice, insofar as the procedure is available on fewer statistical computer programs, it has the considerable advantage of immediately providing consistent and unbiased estimates of the coefficients of the independent variables, as well as their standard errors, thereby allowing for further statistical inference, such as the calculation of the t-values of the estimates.[40] Just as importantly, with TR one can use several truncation points for different sub-periods, in contrast to TOLS. One only has to discard the part of the sample outside of the range (τ_m, τ_x) pertaining to the various subsamples, and specify these upper and lower limits within the command and let the computer calculate the values.[41]

Conclusion

The analysis of height data can be challenging if a part of the underlying height distribution is missing from the sample, as is the case for records stemming from institutions which imposed a minimum or maximum height requirement on entrants, as many military establishments, in fact, did.[42] We have reviewed several methods that can be used on such deficient height distributions and shared our experience working with them over the course of the last two decades. We suggest that one begin the analysis with a visual inspection of the sample height distributions, and a determination of the actual height requirements in conjunction with prior information on the legal height requirements. We have discussed four methods (QBE, K&K, TOLS, TR) for estimating means, trends and covariates of heights from a deficient sample with shortfall. Only to the extent that the conclusions obtained with the various methods support each other, should the results be accepted as valid. The mean height of a population generally changes very slowly over time. During the course of the last century and a half mean height of European populations increased at a rate of

11

between 1-1.5 cm per decade. Hence, if the estimates diverge from this order of magnitude substantially, the presumption is that the estimates might well be incorrect, and in need of further consideration.

On the basis of simulation experiments we conclude that the QBE is inefficient, and should no longer be used to estimate heights. The K&K and the TOLS methods have advantages and disadvantages, and hence, should not be used by themselves, but in conjunction with at least one other method in order to confirm the obtained results. The K&K method is a quick method to obtain an overall impression of trends. Its major shortcoming is its inability to estimate the covariates of height. In contrast, the TOLS is at least able to estimate the relative size of the covariates, while the TR method is able to estimate both the covariates and their standard error. As a consequence, the TR method should be preferred over the other methods.

Appendix A. Abbreviations – Definitions

amount of shortfall - the percent of missing observations from the sample beyond the HRs.

E – expectation operator.

F.i. – French inches

HR - height requirement(s) (minimum and/or maximum) in order to be accepted into an institution.

m – The true mean height of the population from which the sample is drawn.

$\hat{m}_{K\&K}$ – mean height obtained from the sample using the K&K method: $E(\hat{m}_{K\&K}) \neq m$.

$\hat{m}_{TR}$ – Estimator of m using the maximum likelihood method such that $E(\hat{m}_{TR}) = m$.

$\hat{m}_{QBE}$ – Estimator of m using the QBE method, but $\hat{m}_{QBE}$ is not an efficient estimator of m.

$\hat{m}_{TOLS}$ – Estimator of m using TOLS method such that $E(\hat{m}_{TOLS}) \neq m$.

HR_{min} - minimum height requirement.

HR_{max} - maximum height requirement.

$N(\mu,\sigma)$ – Normal distribution with mean μ and standard deviation σ.

OLS - ordinary least squares regression analysis.

QBE – Quantile Bend Estimator – a method of estimating μ and σ by estimating the amount of shortfall. Has been found to be inefficient.

RSML – reduced sample maximum likelihood – see TR.

σ – the true standard deviation of the population from which the sample is drawn.

shortfall – missing observations from the sample on account of the HR.

τ_m - the value of the minimum truncation point.

τ_x - the value of the maximum truncation point.

TOLS – truncated ordinary least squares; linear regression analysis after the values of the sample outside of the range (τ_m, τ_x) have been eliminated.

TR – truncated regression: uses maximum likelihood procedure to estimate μ, its standard error, as well as σ; previously referred to as RSML, reduced sample maximum likelihood or as TML, truncated maximum likelihood.

truncated sample – the sample obtained after the observations outside of the range (τ_m, τ_x) are eliminated.

truncation point – the value at which shortfall begins.

Appendix B. The Truncated Maximum Likelihood Estimate

Let x_i, $i = 1,..., n$, the n observations in a (sub)-sample fully truncated below τ_m. The density of this sample is the product of the individual densities $f(x_i)$, which are densities of a normal distribution with mean μ and a standard deviation σ, normalized by the factor $1-\Phi[\frac{\tau_m - \mu}{\sigma}]$, so that the density integrates to one over the range above the truncation point. Φ is the standardized normal distribution function. The estimates can be found numerically by maximizing the sample density subject to the unknown parameters μ and σ (maximum likelihood estimation). With rounding, the likelihood function to be maximized is given by:

$$L = \prod_{i=1}^{n} \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{[-\frac{1}{2\sigma^2}(x_i - \mu)^2]}}{1 - \Phi[\frac{\tau_m - z - \mu}{\sigma}]}, \text{ where } z \text{ the linear measure (inch, cm) of the rounding interval}$$

below the truncation point (Mokyr and O´Gráda. 1996). That is, in case of the 18th century French military τ_m was 62 French inches, and we assumed that $z = 0.25$, i.e., that men whose height equaled 61.75 French inches were allowed to pass muster, by having their height rounded up to 62 French inches. This is a crucial assumption, insofar as it increases the height estimates by about 1 cm. Even if the assumption is arbitrary – it is plausible, and should be implemented as the best guess estimate.

Appendix C. Density Estimation (Smoothing)

To estimate the continuous density of a sample $X = (x_i, i = 1,..., n)$, perform the following steps Scott (1992, ch. 6): Calculate the optimal bandwidth $h = 2.78 \, \sigma_e \, n^{-1/5}$, where σ_e is the estimated standard deviation.[43] Define a stepwidth s (for heights in cm $s = 0.1$, for heights in inches $s = 0.05$ is a reasonable choice). With this stepwidth generate points P_j, $j = 1,...,$ the smallest element being $\min(X) - h$, the largest about $\max(X) + h$. Calculate the estimated density

on every point P_j by weighting the observations X on each of these points: $f_e(P_j)$

$$= \frac{1}{nh} \sum_{i=1}^{n} K[\frac{P_j - x_i}{h}]. \text{ The function } K(u) \text{ is } \frac{15}{16}(1 - u^2)^2 \text{ if } |u| \le 1. \text{ Otherwise } K(u) \text{ is zero.}$$
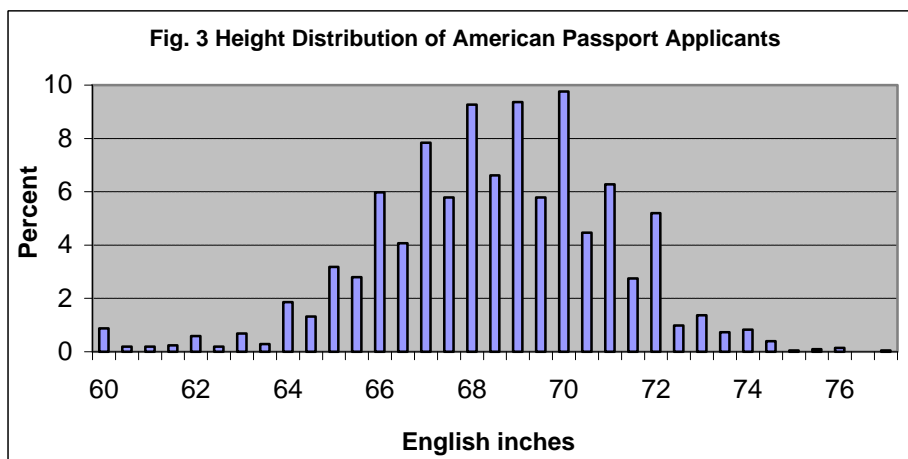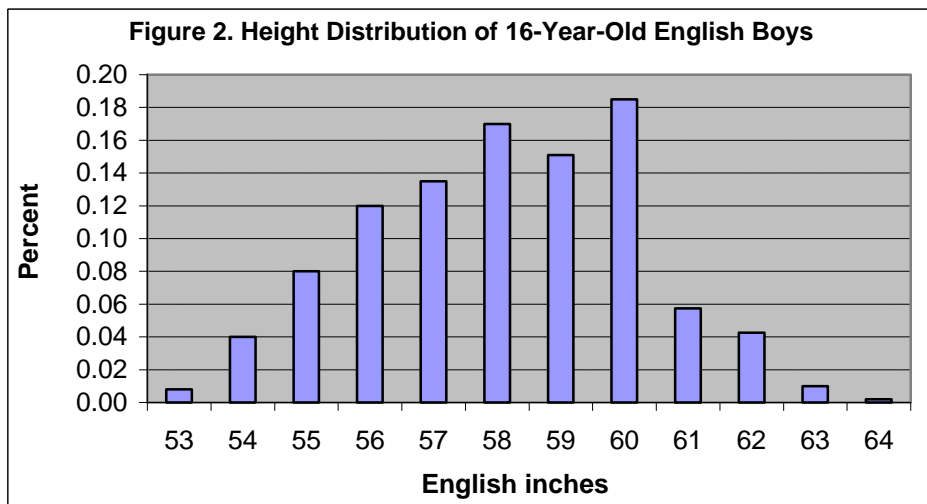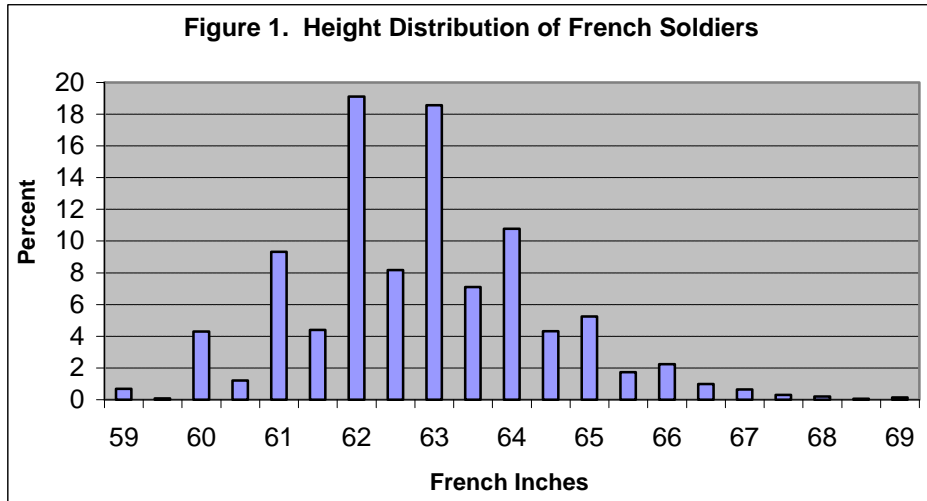
15

Figure 1. Height Distribution of French Soldiers

Figure 2. Height Distribution of 16-Year-Old English Boys

Fig. 3 Height Distribution of American Passport Applicants

Table 1. Comparison of the accuracy (cm) of the QBE and TR Procedures

Continuous Distributions[a]

N	Mean Bias		Mean σ of estimates		Mean Square Error		
	QBE	TR	QBE	TR	QBE	TR	Ratio
12,000	-0.40	+0.21	1.70	0.91	3.05	0.87	3.52

Discrete Distributions[b]

N	Mean Bias[c]		Mean σ of estimates		Mean Square Error		
	QBE	TR	QBE	TR	QBE	TR	Ratio
16,000	0.62	0.19	1.72	1.00	3.36	1.03	3.26

Note: N = number of samples for which means were estimated with both methods; σ = standard deviation of the estimates. Each simulation was done with a sample size of between 250 and 500 observations,
[a] The amount of shortfall was between 10 and 40 percent. True mean of the population from which the samples were drawn was 165 cm with a σ of 6.5 cm.
[b] The amount of shortfall was between 5 and 45 percent. The case in which shortfall was = 0 was not considered in the above averages. The true mean of the population from which the samples were drawn was 65 Austrian inches (A.i.) with a σ of 2.6 A.i. (1 A.i.= 2.633 cm). Once the sample was obtained, the observations were rounded to the nearest A.i..
[c] Mean of absolute values.
Source: Heintel, 1996a, 1998.

Table 2. The Accuracy of Trend Estimates. The Probability of Obtaining Correct Direction of Change with Three Methods

	Probability				K&K
N	QBE	TR	K&K	n	Advantage
10,000	0.64	0.69	0.84	250	0.15 - 0.20
10,000	0.67	0.74	0.91	500	0.17 - 0.24
Average	0.66	0.72	0.88		0.16 - 0.22

Note: N = number of pairs of samples for which direction of change was estimated with all three methods; N= sample size of each simulation. Simulations were made with 10 different specifications, with 1,000 simulations each. The true mean of the population from which samples were drawn increased in 9 specifications from 165 cm to 166 cm, and from 165 to 165.5 in one specification. The σ was 6.5 cm throughout. Shortfall varied from 10 to 50 percent, and the truncation points from 160 to 165 cm.
Source: Heintel, 1996b.

References

Baten, J. and M. Heintel. 1995. "Zum Problem der Verteilung mit Shortfall bei der Nutzung des
Indikators 'Durchschnittliche Körpergröße'," Historical Social Research 20 4: 135-154.

Bogin, Barry. 1999. Patterns of Human Growth. Cambridge: Cambridge University Press, 2nd
edition.

Cheung C. and A. Goldberger. 1984. „Proportional Projections in limited Dependent Variable
Models," Econometrica 52: 531-34.

Chay, Kenneth Y. and James L. Powell, (2001) "Semiparamertic Censored Regression Models,"
Journal of Economic Perspectives, 15, 4: 29-42.

Craig, Lee, and Weiss, Thomas. "Nutritional Status and Agricultural Surpluses in the Antebellum
United States," in John Komlos and Jörg Baten, eds*., The Biological Standard of Living in
Comparative Perspective*, Stuttgart: Franz Steiner, 1998, pp. 190-207.

Cuff, Timothy. „Variation and Trends in the Stature of Pennsylvanians, 1820-1860," in John
Komlos and Jörg Baten, eds*., The Biological Standard of Living in Comparative Perspective*,
Stuttgart: Franz Steiner, 1998, pp. 208-235.

DiNardo, Hohn and Justin L. Tobias, (2001) "Nonparametric Density and Regrssion Estimation,"
Journal of Economic Perspectives, 15, 4: 11-28.

EViews 4 User's Guide Quantitative Irvine CA: Micro Software, 2000.

Floud, R. and K. Wachter. 1982. Poverty and Physical Stature: Evidence on the Standard of Living
of London Boys 1770-1870. Social Science History 6: 422-52.

Floud, R., K. Wachter and A. Gregory. 1990. Height, Health and History. Cambridge: Cambridge
University Press.

Freiman, J.A. Chalmers, TC. Smith, H. and Kuebler, RR. 1978, "The importance of beta, the
type II error and sample size in the design and interpretation of the randomized control

trial," <u>New England Journal of Medicine</u>, 299 690-4; cited in C.G. Nicholas Mascie-
Taylor, "Statistical Issues in Anthropometry," in S.J. Ulijaszek and C.G.N. Mascie-Taylor
eds., <u>Anthropometry: the Individual and the Population</u>. Cambridge University Press, 1994,
56-77.

Greene, W. H. 1993. <u>Econometric Analysis</u>. New York: MacMillan.

Haines, Michael R. „Health, Height, Nutrition, and Mortality: Evidence on the ‚Antebellum Puzzle'
from Union Army Recruits for New York State and the United States," in John Komlos and
Jörg Baten, eds*., The Biological Standard of Living in Comparative Perspective*, Stuttgart:
Franz Steiner, 1998, pp. 155-180.

Harris, B. 1994. Health, Height, and History: An Overview of Recent Developments in
Anthropometric History. <u>Social History of Medicine</u> 7: 297-320.

Heintel, Markus. 1996a. "Historical Height Samples with Shortfall: A Computational Approach,"
<u>History and Computing</u> 8: 24-37.

Heintel, Markus. 1996b. "Estimating Means and Trends of Normal Samples with Shortfall,"
Working Paper no. 96-20, Department of Economics, University of Munich.

Heintel, Markus. 1997. Über einige statistische Probleme und Möglichkeiten der
anthropometrishen Geschichte," in Jörg Baten and Markus Denzel (eds). <u>Wirtschaftsstruktur
und Ernährungslage, 1770-1870</u> (St. Katharinen: Scripta Mercaturae Verlag, 1997), pp. 108-
130.

Heintel, Markus and Jörg Baten. 1998. Smallpox and Nutritional Status in England, 1770-1873: On
the Difficulties of Estimating Historical Heights. <u>Economic History Review</u> 51: 360-371.

Heintel, Markus, Sandberg, L. and R. Steckel. (1998). "Swedish Historical Heights Revisited. New
Estimation Techniques and Results," in in John Komlos and Jörg Baten, eds*., The Biological
Standard of Living in Comparative Perspective*, Stuttgart: Franz Steiner, 1998, pp. 449-458.

Komlos, John, 1985. "Stature and Nutrition in the Habsburg Monarchy: The Standard of Living and Economic Development in the Eighteenth Century" <u>American Historical Review</u> 90: 1149-61.

Komlos, John, 1989. <u>Nutrition and Economic Development in the Eighteenth Century Habsburg Monarchy</u> (Princeton: Princeton University Press).

Komlos, John, 1993. "The Secular Trend in the Biological Standard of Living in the United Kingdom," <u>Economic History Review</u>, 46: 115-44.

Komlos, John, 1995. <u>The Biological Standard of Living in Europe and America 1700-1900</u>. Aldershot: Variorum Press.

Komlos, John, 1996. "Anomalies in Economic History: Toward a Resolution of the 'Antebellum Puzzle'," <u>Journal of Economic History</u> 56: 202-214.

Komlos, John, 1998. "Shrinking in a Growing Economy. The Mystery of Physical Stature during the Industrial Revolution," <u>Journal of Economic History</u> 58 (1998) 3: 779-802.

Komlos, John, 1999. "On the Nature of the Malthusian Threat in the Eighteenth Century," <u>Economic History Review</u> 52, no. 2 (Nov. 1999): 730-48.

Komlos, John and Joo H. Kim. 1990. Estimating Trends in Historical Heights. <u>Historical Methods</u> 23: 116-120.

Komlos, John and Jörg Baten 1998. <u>The Biological Standard of Living in Comparative Perspectives: Proceedings of a Conference Held in Munich January 18-23, 1997</u>. Stuttgart: Franz Steiner Verlag.

Komlos, John and T. Cuff. 1998. <u>Classics of Anthropometric History</u>. St. Katharinen: Scripta Mercaturae.

Le Roy Ladurie, E., Bernageau, N. and Y. Pasquet. 1969. "Le Conscrit et l'ordinateur: Perspectives de recherches sur les archives militaires du XIXe siècle francais," <u>Studi Storici</u> 10: 260-308.

Mokyr, J. and C. Ó Gráda. 1994. "The Heights of the British and the Irish c. 1800-1815. Evidence from Recruits to the East India Company's Army." In <u>Stature, Living Standards, and Economic Development: Essays in Anthropometric History</u>, edited by J. Komlos, Chicago: The University of Chicago Press, pp. 117-28.

Mokyr, J. and C. Ó Gráda. 1996. Height and Health in the United Kingdom 1815-1860: Evidence from the East India Company Army. <u>Explorations in Economic History</u> 33: 141-68.

Preece, M. A. and M. J. Baines. 1978. A New Family of Mathematical Models describing the Human Growth Curve. <u>Annals of Human Biology</u> 5: 1-24.

Quetelet, A. 1831. Recherches sur la loi de croissance de l'homme. <u>Annales d'Hygiène Publique et de Médicine Légale</u> 6: 89-113.

Salvatore, Ricardo and Jörg Baten, "A Most Difficult Case of Estimation: Argentinian Heights, 1770-1840," in John Komlos and Jörg Baten, eds*., The Biological Standard of Living in Comparative Perspective*, Stuttgart: Franz Steiner, 1998, pp. 90-96.

Sandberg, L. and R. Steckel. 1987. "Heights and Economic History: The Swedish Case," <u>Annals of Human Biology</u> 14: 101-10.

Scott, D. W. 1992. <u>Multivariate Density Estimation</u>. New York: Wiley.

<u>Stata Reference Manual. Release 7</u>, College Station, TX: Stata Press, 2001, vol. 4.

Steckel, R. 1979. "Slave Height Profiles from Coastwise Manifests," <u>Explorations in Economic History</u>, 16: 363-380.

Steckel, R. 1995. "Stature and the Standard of Living," <u>Journal of Economic Literature</u> 33: 1903-40.

Tanner, James M. 1978. Foetus into Man: Physical Growth from Conception to Maturity. Cambridge: Cambridge University Press.

Tanner, James M. A History of the Study of Human Growth. Cambridge: Cambridge University Press, 1981.

Trussell, J and Kenneth Wachter, "Estimating the Covariates of Historical Heights," NBER Working Paper No. 1455, 1984.

Twarog, S. 1997. Heights and Living Standards in Germany, 1850-1939: The Case of Württemberg. In Health and Welfare during Industrialization, edited by R Steckel and R. Floud. Chicago: The University of Chicago Press.

Villermé, L. R. 1929. Mémoire sur la taille de l'homme en France. Annales d'Hygiène Publique et de Médicine Légale 1: 551-59.

Wachter, K. 1981. Graphical Estimation of Military Heights. Historical Methods 14: 31-42.

Wachter, K. and J. Trussell. 1982. Estimating Historical Heights, Journal of the American Statistical Association 77: 279-293.

Endnotes

[1] Because of regulations and financial constraints, archival sampling can be challenging, and it is often impractical to obtain a truly random sample.

[2] Obviously prison samples, for example, pertain essentially to the lower classes of the society.

[3] For example, the French military had a HR_{min} of 62 French inches (F.i.) (or 167.81 cm), implying that men shorter than this height had a lower probability of being accepted into the infantry and therefore into the sample. The mean height of the soldiers was about 171.5 cm, but this obviously does not pertain to the height of the population of able-bodied men living in France at the time: this mean is upwardly biased. However, one can use this information to estimate the mean height of the population of all men by assuming that 18th century heights were also normally distributed with a standard deviation equal to that of modern adult populations (about 6.86 cm). On the basis of this assumption one can calculate that the mean height of the male population must have been about 161.7 cm. That, in turn implies, that 82.5 percent of the population must have been shorter than the HR_{min} of 167.8 cm. However, that is far from being the case in the sample: only 20.5 percent of the sample is to the left of 167.8 cm, and the difference between the two values (62 percent) is the amount of shortfall, i.e., 62 percent of observations below 167.8 cm are missing from the data set.

[4] Deviations from the normal distribution might be due to insufficient sample size. Thus, in case of archival sampling, one should examine the histograms during the sampling process, in case the sample needs to be enlarged.

[5] The value at which shortfall begins is referred to as the truncation point, even if the HR is never perfectly enforced.

[6] Heintel and Baten (1998) describe some statistical difficulties with these data. Boys of this age who were extraordinarily tall might have been deemed unsuitable for life at sea or might have had the opportunity to go directly into the labor force, and might not have required the charitable support of the Marine Society.

[7] There might be slight deviations from normality due to rounding. Americans who traveled abroad applied for passports as proof of citizenship. There were obviously no HR for traveling abroad, so one would not expect such a height distribution to be truncated. The distribution of height of a sample of 2040 male adult (older than 22 years) applicants, extracted from archival

records in Washington D.C. confirms that the heights were not truncated (Figure 3). Thus, the sample can be analyzed as usual. Nonetheless, a frequent pattern is evident, i.e., that there are departures from a perfectly normal distribution due to rounding on 70 inches. Heaping on favorite numbers is observed in almost all height samples, even in modern ones. Thus, slight departures from normality, even in the absence of height requirements is possible, usually due to rounding on favorite numbers. However, such biases tend have only a marginal affect on results, and can be ignored generally (Komlos 1999).

[8] It might be the case, say, that 67 inches has fewer observations than warranted, but the rounding to 68 and 66 inches would tend to cancel each other. However, there might be some situations when rounding does have a significant impact on the analysis. For instance, in an Argentinean data set analyzed by Baten and Salvatore (1998) a large number of the height of recruits were recorded as being exactly equal to the HR_{min}. Obviously, the height of recruits was rounded up to the HR_{min} in order to enable them to enter the military. In order to obtain accurate estimates, the sample had to be truncated one unit above the HR_{min}. This example illustrates again the utmost importance of a visual inspection of the distributions, in order to ascertain data anomalies.

[9] Heaping might also affect information on age. Ex-slaves of the union army, for instance, did not know their age exactly. Hence, one can observe heaping on certain ages, and the age-by-height profile is not accurate.

[10] A visual inspection of the distributions suffices in most cases. However, Heintel (1996a, 1998) developed procedures to estimate truncation points, based on the fact that the increase in sample density is steepest at τ, because this point is where the distribution changes discontinuously from a complete normal distribution to one with a shortfall (see Appendix C). However, this procedure does not produce superior estimates to visual inspection.

[11] Youth and adult height distributions should also be inspected separately. This is the case, because the HR_{min} were at times not as stringently applied to youth, inasmuch as they were expected to grow subsequently.

[12] Floud, Wachter and Gregory (1990) merged data from the British Army and from the Royal Marines even though they had different HRs, which also were enforced with different degrees

of stringency. Additionally, the Royal Marines had a maximum height requirement. Therefore, their results fluctuated unreasonably (Komlos, 1993).

[13] An adult born in 1740, for instance, was influenced by the nutritional circumstances between 1740-63, while an 18-year-old born in the same year was influenced by those of 1740-1758. Hence, they do not overlap for the years 1759-63.

[14] In fact, in the French example, cited above, the HR_{min} was lowered to 60 French inches after 1740.

[15] For 18th-century samples, one might also consider examining the histograms for different regiments in order to see if there were some anomalies in the recruitment procedures used in the field, because there was still considerable local autonomy in recruiting.

[16] The sample size (n) needed, for a given degree of accuracy, can be calculated from the following equation: $n = \dfrac{2s^2(Z_{1-a} + Z_{1-b})^2}{x^2}$, where σ is the standard deviation, α the is desired level of significance, β is the power of the test (defined as the probability of correctly rejecting the null hypothesis, if it is false), and x is the difference between two means (Freiman et al, 1978, 58). Hence, the sample size needed to detect a 1 cm difference at a 0.5 significance level, for $\beta = 0.7$, would be: $n = \dfrac{2(6.8)^2(1.96 + 0.5244)^2}{1^2} = 571$ observations in each cell. Admittedly, this sample size is often impractical in real world situations given financial and archival constraints, but at least it gives one a sense of the ideal sample size to strive for.

[17] The appropriate time interval (annual, quinquennial, decadal, quarter century, etc.) is determined by the number of observations available for analysis. The number of observations per interval should be several hundred for robust estimates.

[18] Admittedly, the occupation of an adult soldier would not have had an effect on his height during his growing years. Nonetheless, it is legitimate to include this variable insofar as social mobility was not very pronounced in the 18th and early 19th centuries, and therefore, the soldier's occupation can serve as a proxy of his family's social status during childhood and adolescence.

[19] One has to make sure that the explanatory variables entered into the analysis are legitimate in the sense that they are variables that could have affected the height of the soldiers prior to

adulthood. The "year of recruitment" during the American Civil War has been used as an explanatory variable, although it is not pertinent as far as the determination of the soldiers' heights were concerned. The height of the soldiers was not determined by the year of recruitment, in the same sense as the income of the soldier's parents was a determining factor in the height of the soldiers. In contrast, different unobserved recruitment procedures meant that a particular birth cohort was sorted into different recruitment years by height. In other words, heights could have determined enlistment year, and not the other way around. Hence, recruitment year is not a useful right hand side variable.

[20] The discrete histograms can be turned into a continuous smooth distribution using a kernel density estimator (which incorporates a smoothing function), as height is a continuous variable (see Appendix C). The continuous distribution so obtained can also be used to estimate τ, as well as the mode, which in turn is an estimator of $\hat{m}$, as the mode and the mean are identical in a normal distribution. Successive such estimates provide trend estimates of the mean height. The accuracy of the kernel density estimator of $\hat{m}$ has not been explored. The method also does not allow for estimation of the effect of covariates on heights (Heintel 1996a, 1998; DiNardo and Tobias 2001).

[21] It also provides estimates of the variance of the height of the population. For the QBE procedure the observations below the HR_{min} do not have to be discarded.

[22] The algorithm estimates the amount of shortfall by minimizing the bending in a quantile-quantile plot. The method completes the sample in such a way that the part of the plot above the truncation point forms a straight line. This line is estimated by means robust regression. If a is the amount of shortfall, $\mathbf{F}$ is the standard normal distribution function and F(y) is the empirical distribution function of the sample at point y, then the aim is to estimate a such that the line: $\mathbf{x} = \mathbf{F}^{-1}[(1 - \mathbf{a})(1 - \mathbf{F}(\mathbf{y}))]$ is as straight as possible. Yet, since we usually have only a handful of points, the estimation of the straight line has a relatively large variance.

[23] If one would like to estimate urban heights distinctly from rural heights, one has to first divide the sample into urban and rural provenance before proceeding with the K&K analysis. One cannot do it in one step as with TR or with TOLS by including a dummy variable for urban heights.

[24] This is the case, because μ is a monothetic function of $\hat{m}_{K\&K}$: if $\hat{m}_{1,K\&K} < \hat{m}_{2,K\&K}$ then it follows that $\mu_1 < \mu_2$.

[25] The truncated regression is essentially the model presented in Appendix B replacing the overall mean μ by the individual mean in a particular category.

[26] By eliminating the sample outside of the range $(\tau_m \ \tau_x)$ the artificial impact of potentially different shortfall patterns in the sub-samples (caused by different truncation points and/or different amounts of shortfall) is avoided by "equalizing" the bias over the complete sample (see Heintel and Baten, 1998, footnote 17, for an interesting study of artificial correlations if one fails to equalize the bias). In fact, the theoretical bias β_i of the respective regression coefficient vector β_i, $i = 1,...$, has the property $\beta_i = \beta_i(\kappa - 1)$ for some factor τ with $0 < \kappa < 1$ (Cheung and Goldberger, 1984). This means that the sign, and the relative ordering of the β_i's are not affected by the bias. As a consequence, the coefficient of the time dummy variables indicates the true direction of height trend over time. Furthermore, as κ is the same for all regression coefficients, one can infer the relative magnitude of the covariates, because the values of the coefficients reflect their importance relative to one another.

[27] The coefficient vector is found numerically by maximizing the likelihood function subject to the unknown parameters (Chay and Powell, 2001). Another considerable advantage of the TR method is that one does not have to assume that the standard deviations remain constant as one does if one converts $\hat{m}_{TOLS}$ into m.

[28] See, for example, the unstable QBE estimates in Floud and Wachter (1982), Sandberg and Steckel (1987), Twarog (1997). See Heintel, Sandberg & Steckel (1998) for empirical investigations confirming the unreliability of the QBE.

[29] Inspite of the fact that theoretically one expects the bias to be zero.

[30] This distribution is not reported here. John Komlos, in collaboration with Michel Hau and Nicolas Bourguinat, "The Anthropometric History of Early-Modern France, 1666-1766," unpublished manuscript, University of Munich.

[31] The collection of height data began in 1716, and the recruiting system remained in effect until the revolution. Only between 1740-1762 was the τ_m lowered to 60 French inches.

[32] However, the data within a particular HR regime can be used for further analysis. For the French example, the soldiers recruited during the war years 1740-1762 can and should be also

analyzed separately, i.e., without merging with those recruited earlier or later. These results could corroborate the findings obtained with the combined data set. In effect, one create two samples for analysis. Sample (A) encompasses the whole period 1716-1786, with an effective τ_m of 62 French inches (F.i.), and sample (B) for the period 1740-62 with τ_m of 60 F.i.. The principle is that soldiers recruited with different HRs should not be conflated, unless the HRs has been equalized between them. They can be analyzed together using TOLS only if the largest of the two τ_m is applied to both of them – in this case 62 French inches. With TR, however, one can use both 60 and 62 F.i. as the appropriate τ_m by specifying them in the program as the lower limit for the appropriate time interval.

[33] The known HR was used in the estimations. 60,000 simulations were run with 10 different specifications: n= 250, 500; σ = 6.5 cm and unchanged, ΔH = 1 cm in 9 specifications and 0.5 cm in one specification.

[34] None of the methods shows substantial improvements in accuracy with increased sample size (Table 2).

[35] If μ_1=165 cm and μ_2=165.25 cm (with σ remaining unchanged at 6.5 cm), the probability that the K&K method estimates the direction of the trend correctly is 10 percent higher than that of the QBE. However, if μ_2=166 cm, K&K advantage rises to 22 percent. In other words, in that case K&K has 22% higher probability of estimating the trend correctly than QBE (Table 2). In contrast, the relative accuracy of the TR v.s. the QBE does not change substantially in this range of height changes.

[36] While $\hat{m}_{K\&K}$ can be calculated with all statistical packages, $\hat{m}_{TR}$ cannot be calculated with SPSS.

[37] This is also the case with the QBE, but not with TR or TOLS.

[38] That is to say, if height decreases, then σ can decrease without affecting the advantage of the K&K method relative to the TR method.

[39] Assume that σ = 6.86 cm, the same as for modern populations, and that the effective τ_m was not actually equal to 62 F.i. (167.81 cm), as indicated by the histograms, but slightly lower, at 61.75 F.i., (167.14 cm), because it is plausible to suppose that those recruits who were slightly shorter than the τ_m were probably allowed to slip through by having their height measurement

rounded up to the nearest whole French inch. (This assumption increases the estimated

population means by about 1 cm.) Next take a normal distribution with $\mu = 170$ cm and σ

$=6.86$ cm and discard all observations below 167.14 cm. Then calculate the mean of the

truncated distribution, to obtain 173.8 cm. Thus, we can reverse the calculation and assert that

if $\hat{m}_{TOLS}=173.8$ cm then $m=170$ cm. In this manner we obtain a conversion schedule for

$\hat{m}_{TOLS(converted)}$. This estimator is unbiased, i.e., it has the property that $E(\hat{m}_{TOLS(converted)})=m$. In

this manner we obtained the following schedule (millimeters):

$\hat{m}_{TOLS}$	$\hat{m}_{TOLS(converted)}$
1738	1700
1733	1690
1729	1680
1726	1670
1722	1660
1719	1650
1716	1640
1713	1630
1711	1620
1709	1610
1707	1600

This schedule can be used to obtain by interpolation estimates of the height of the French male

population from which the soldiers were drawn. 1716.2 mm converts approximately to 1640.1

mm while 1711.8 mm converts to 1623.2 mm. One can also run a linear regression with

$\hat{m}_{TOLS(converted)}$ as the dependent variable and $\hat{m}_{TOLS}$ as the independent variable in order to

obtain a conversion formula. This procedure can also be used to convert the estimates

obtained by the K&K method into population height estimates.

[40] In addition, it estimates σ of the height distribution (provided that the truncation point is to

the left of the mean), rather than assume it as with TOLS.

[41] For instance, in the French example above, one would discard all observations smaller than

60 French inches (F.i.) during the period 1740-1763, and those smaller than 62 F.i. for the

remainder of the sample. In this case, we can use two different truncation points.

[42] Anthropometric results should be treated cautiously until corroborated with an independent data set, or with collateral evidence. After all, the socio-economic composition of the institution studied might have varied over time, even in the absence of explicit changes in the admission criteria. This might be due both to supply and demand considerations. The willingness of individuals to enter the military, for instance, might have varied over time. Similarly, the size of the military might have expanded sufficiently so that individuals were accepted who would have been rejected at an earlier time. In either case, the underlying population which provided members of the organization might have changed over time. This problem is quite intractable. How the supply of, and demand for, potential entrants into an institution fluctuated over time might not be ascertainable at all. Yet, anthropometric history is frequently concerned with the *longue durée*, so that even social processes which move at a glacial pace might have an impact on the institutions in question. Such issues would require considerable research to answer, and in such circumstances, it is advisable to leave the issue open, and look for other data sets from other types of institutions to corroborate or refute the findings.

[43] For extremely rounded data a modification of the bandwidth might be useful (Heintel and Baten, 1998, footnote 6).