Please Note:

This is the last working paper version before this study was submitted and accepted. Please cite as Heintel, Markus and Baten, Joerg. "Smallpox and Nutritional Status in England, 1770-1873: On the Difficulties of Estimating Historical Heights", *Economic History Review* 51-2 (1998), pp. 360-371. DOI: 10.1111/1468-0289.00096

Economic History Review, LI, 2(1998), pp. 360–371

Smallpox and nutritional status in England, 1770-1873: on the difficulties of estimating historical heights[1]

By MARKUS HEINTEL and JOERG BATEN

In a recent issue of this journal, Voth and Leunig (V&L) analysed the impact of smallpox on physical stature in the eighteenth and nineteenth centuries and drew far-reaching conclusions about the correlation between the incidence of smallpox and the development of human stature in England.[2] Using a sample of lower-class boys recruited into the Marine Society,[3] they estimated that the disease had a strong negative impact of about 1 inch on height. We argue below that their methodology is inappropriate and that their results are flawed for at least three reasons.

I

The Marine Society was a charitable institution that took poor boys from the streets of London and educated them as servants and apprentices for the navy. However, the admission of the boys was controlled by minimum height requirements, which changed frequently over time, and increased substantially after the Napoleonic wars. In practice they were enforced with varying degrees of stringency. Hence, some observations below the height standard are also included in the sample, a phenomenon called *shortfall*.[4]

Due to these sampling biases the first and probably most important analytical step is a visual inspection of the height distribution which in our case reveals sample deficiencies such as shortfall, right-hand truncation, and other deviations from normality.[5] Here we applied a continuous

[1] We thank the editors, John Komlos, the participants in the IEHA conference on 'The biological standard of living and economic development', Munich, January 1997, and an anonymous referee for helpful comments. The research was supported by the DFG (German Science Foundation).

[2] Voth and Leunig, 'Did smallpox reduce height?'.

[3] This sample was first analysed by Floud and Wachter, 'Poverty and physical stature'.

[4] See Wachter and Trussell, 'Estimating historical heights'; Heintel, 'Estimating means and trends'.

[5] Komlos found that 16- and 15-year-old boys in the late eighteenth century were also truncated on the right side. Boys of this age who were tall enough might have had the opportunity to go directly into the navy, as they were able to mislead recruiting officers about their age; see Komlos, 'Secular trend', p. 122.

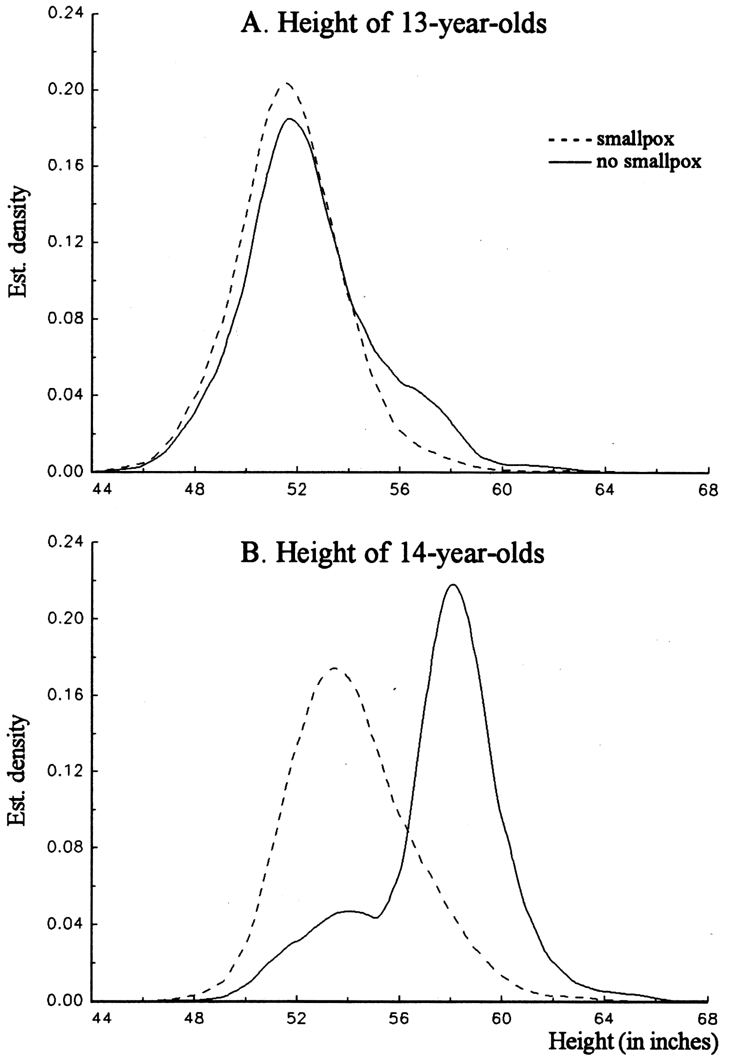

Figure 1. *Height distributions of boys with and without smallpox, full sample, (1770-1873)*
Source: see text

kernel density estimator which leads to better mode approximations than the histograms of V&L.[6]

As an example of this estimation technique consider the 'continuous versions' of V&L's figures 1 and 2, the height distributions of boys aged 13 and 14 with and without smallpox in the complete sampling period (1770-1873) given in figure 1, panel A and B.

[6] For a description of this technique see Scott, *Multivariate density estimation*. An application to height samples is given in Heintel, 'Historical height samples', and Baten and Heintel, 'Zum Problem'. As the Marine Society heights are rounded to the nearest inch, we modified the bandwidth h of the estimator via $h_{\text{mod}} = \min(h, 2)$.

Because their distributions were not standardized for recruitment period, such graphs say nothing at all about the impact of smallpox on heights. This is crucial insofar as more than 90 per cent of the boys admitted into the society before the 1820s were exposed to smallpox, while those born thereafter were not exposed at all.[7] This implies that V&L compared two widely separated birth cohorts, assuming that the only difference between the two groups was their exposure to smallpox. However, that is incorrect. Much social and economic change had taken place in the interim, and most importantly, the minimum height requirement was increased substantially after the end of the Napoleonic wars. Because the increase coincided with the eradication of smallpox, a comparison cannot be made accurately.

Table 1. *Announced and estimated height standards (inches)*

Year in which minimum came into effect	*Announced standard*[a]	*Estimated standard*[b]
1770	51	51
1786	54	52
1792	52	52
1798	54 or 52	51
1809	50 or 51	48
1812	51	51
1814	54	54
1818	55	54
1821	56	56
1824	57	57
1854	59	59
1857	57	58

a Floud and Wachter, 'Poverty and physical stature'
b see section 1

As the announced minimum height requirement could differ from the actual height standard enforced by the admission official, we estimated the real height standard for each of the mustering periods using the truncation point estimator (TPE)[8] (table 1).

The estimates confirm that the height standards were sometimes lowered in practice, but—apart from the second and fifth periods—the results are similar to the announced height requirements.[9] Note that the estimates corroborate the dramatic increase of the height standard in 1814.

As a result, our figure 1 and most of the figures of V&L (figs. 1-5) report not only the correlation between smallpox and time but also the correlation between smallpox and the minimum height requirement

[7] Our main concern here is to address the issues raised by V&L within the context of the Marine Society sample. Hence, we ignore the possible under-registration of the incidence of smallpox in the subsamples after the 1820s (personal communication with Bernard Harris and Jane Humphries).

[8] The TPE is deduced in Heintel, 'Historical height samples'.

[9] In periods with observations numerous enough for smallpox/no smallpox cases (at least 30 for the respective sub-group), we also estimated the minimum height requirement separately for the two groups. The results differed only in the third period: smallpox: 51 inches; no smallpox: 50 inches.

(which is correlated with time).[10] The very small difference in the smallpox/no smallpox distributions for the 13-year-olds is also explained by this fact, because the society stopped recruiting 13-year-olds after 1832.

For a visual analysis we divided the complete sample into seven sub-periods each lasting about 14 years with starting points 1770, 1785, 1800, 1814, 1828, 1842, and 1857.[11] We then estimated the height distributions of the children with and without smallpox separately for ages 13 to 16 in each sub-period.[12] As we investigate the effect of smallpox on height, the first three sub-periods (when both smallpox and no-smallpox cases are available in large enough numbers) are of primary interest (figures 2-5).[13]

Contrary to V&L's assertions, the mode of the smallpox sample is as likely to be below the no-smallpox sample (figures 2A, 2C, 3A, 4A), as above (2B, 3B, 4B, 5B). Sometimes the modes are nearly identical (3C, 4C, 5A).[14] This simple graphical analysis implies that there is no evidence that exposure to smallpox reduced the physical stature of boys of the Marine Society. The contrary finding of V&L is based primarily on three sub-samples recruited between 1770 and 1775 (figures 7-9 of V&L, similar to our figures 2A, 3A, 4A). These sub-samples are by no means representative for the entire period. In only four of the 11 sub-samples were children who had exposure to smallpox smaller than those who had not. In other words, V&L disregarded all periods in which their hypothesis is not confirmed. In a similar vein, their sub-sample regression analysis (table 2) is unrepresentative, inasmuch as it just happens to pertain to those dates in which the no-smallpox group was taller.[15]

II

V&L's regression using the entire sample (their table 1) suffers from two serious mistakes. First, as a result of shortfall, the residual of their regression model is correlated with the co-variates. This leads to biased estimates.[16] Secondly, with changes in the height standard, the amount of shortfall and the bias caused by shortfall usually changed as well.

[10] For example, the correlation coefficient between minimum height requirement and percentage of smallpox is −0.70.

[11] With smaller periods it is impossible to separate the respective subsamples according to age and smallpox/no smallpox.

[12] Throughout the paper we report only results based on sample sizes above 30 observations. For this reason the 17-year-olds are omitted entirely.

[13] In the period between 1814 and 1819 nearly all observations are afflicted with smallpox, while from the 1820s onwards smallpox disappears in the sample.

[14] While in figure 2A the smallpox mode is lower, the smallpox distribution itself is almost identical to the no-smallpox distribution.

[15] V&L regressed height on smallpox, age, and the ability to read and write. We continued these regressions for the periods between 1776 and the 1820s, but we found no significant impact of smallpox on height (with 5% level of significance). The deficiencies of the data mean that such regressions (including the *t*-statistics) are hardly interpretable because of violations of necessary assumptions such as independence of the error term; see section II.

[16] See also Heintel, 'Estimating means and trends'.

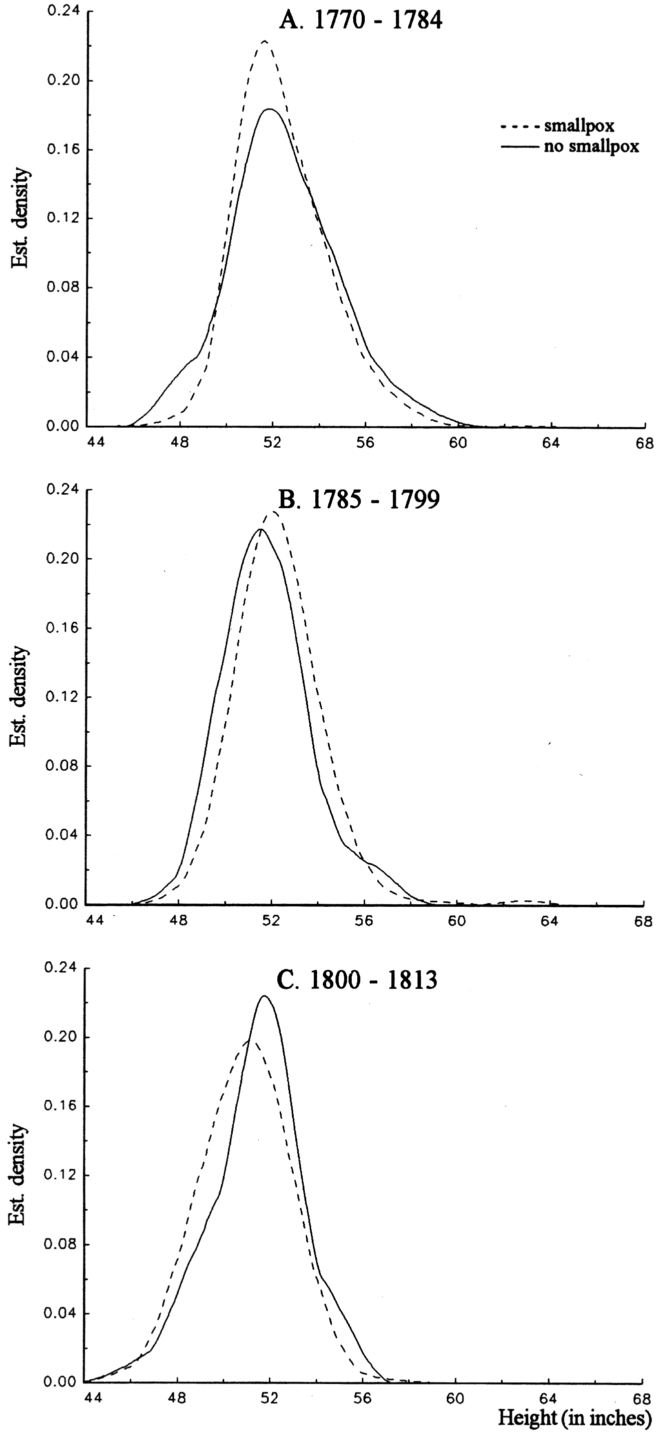

Figure 2. *Height of 13-year-olds, 1770-1813*
Source: see text

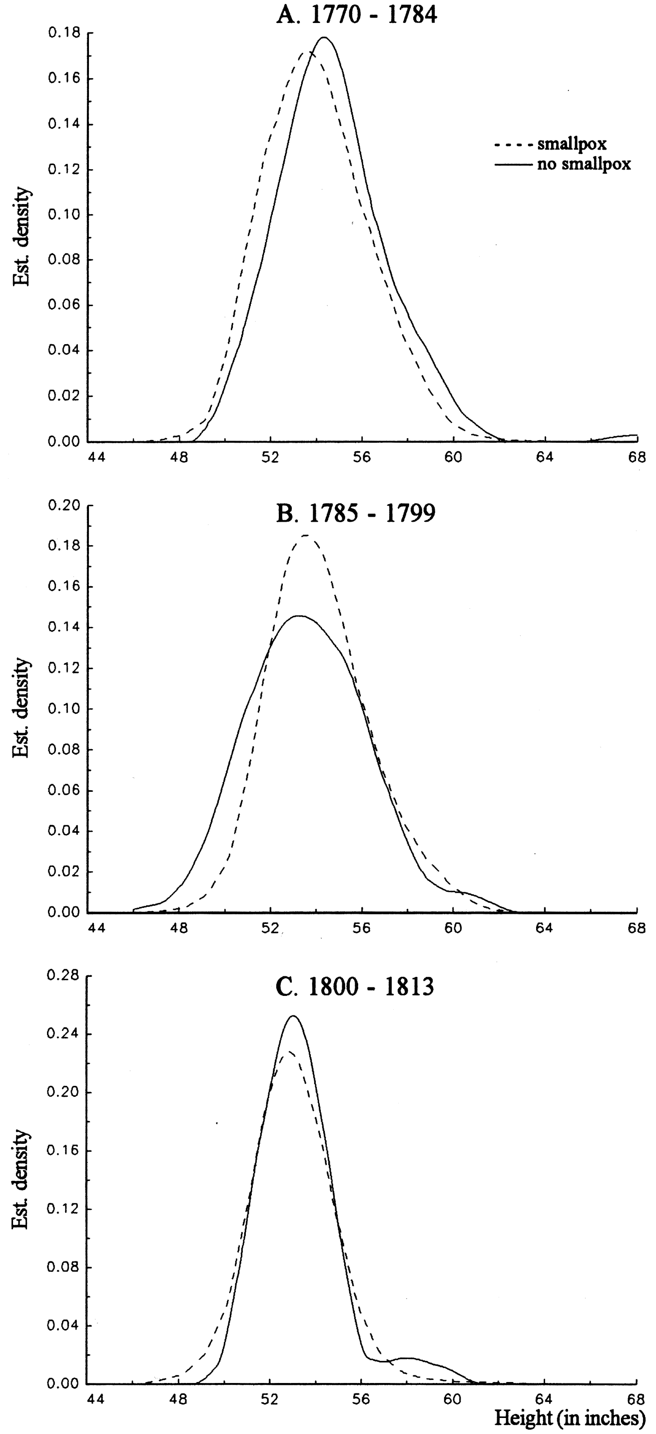

Figure 3. *Height of 14-year-olds, 1770-1813*
Source: see text

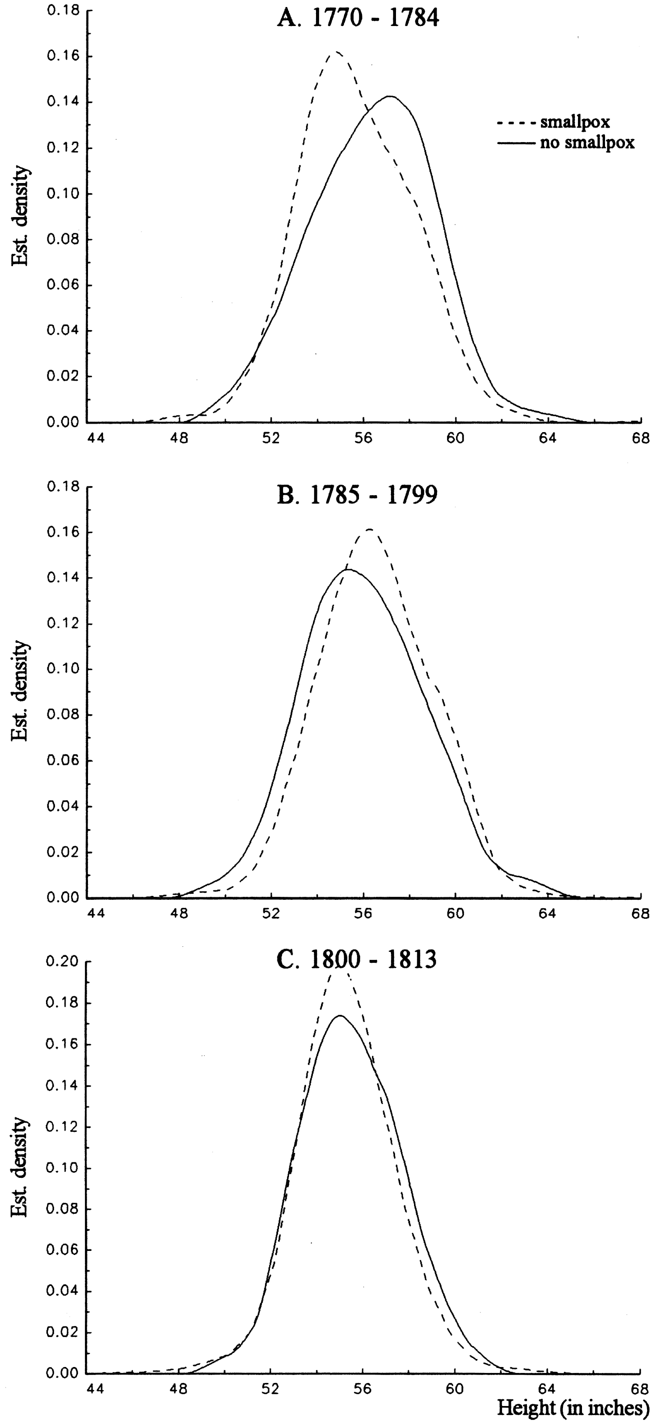

Figure 4. *Height of 15-year-olds, 1770-1813*
Source: see text

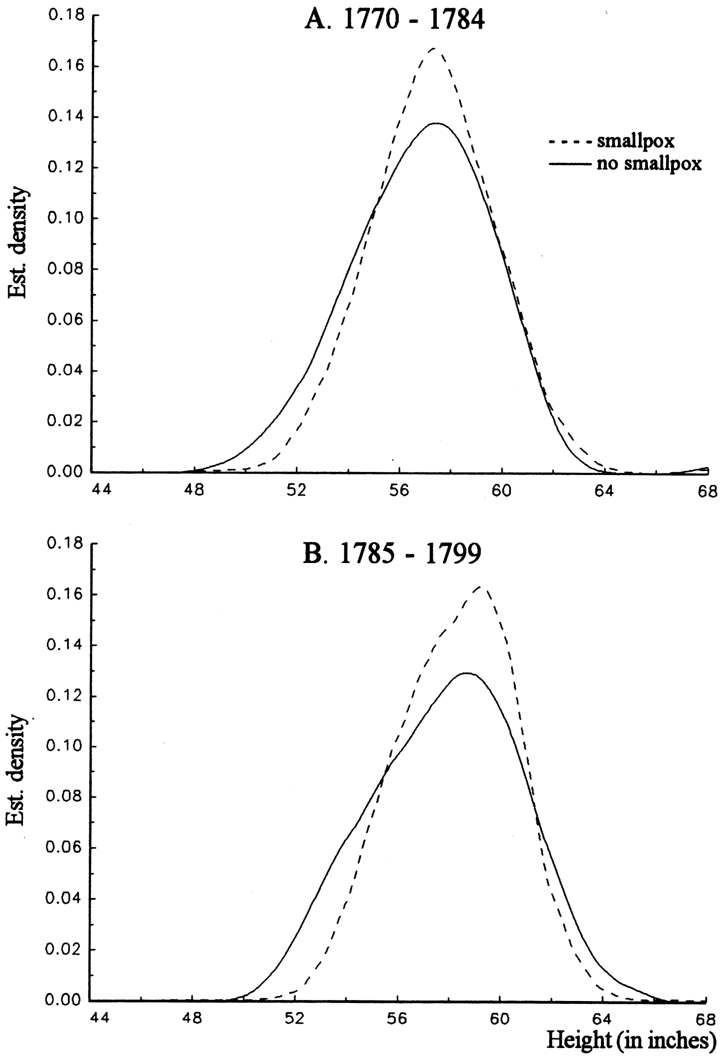

Figure 5. *Height of 16-year-olds, 1770-1799*
Source: see text

Because V&L failed to 'equalize' this bias, they introduced an *artificial* impact of smallpox on height.[17]

Unfortunately, with the Marine Society dataset the greatest minimum height requirement in the nineteenth century is above the mode of heights

[17] As already noted, the incidence of smallpox is correlated with time (the Cramer contingency measure between smallpox/no smallpox and our seven mustering periods is 0.83). To illustrate the importance of, at least, equalizing the shortfall bias, consider the following experiment: We simulated 100 observations normally distributed with mean 60 and standard deviation 2.6 afflicted with 10% shortfall below 59. We assumed that 80% of the 100 observations were *randomly* afflicted with smallpox (which has no impact on the mean). We also simulated 100 observations distributed as above but with 30% shortfall beginning below 59.5 and afflicted with only 20% smallpox (which again has no impact on the mean). Estimating an uncorrected regression with these 200 data points leads to *height* = 61.2 – 0.9*smallpox*, with a *t*-value of –2.7 for the smallpox coefficient (*p*-value 0.00).

in the eighteenth century. This implies that even sophisticated statistical procedures such as truncated OLS[18] or truncated Maximum Likelihood[19] are unreliable. Therefore, regression analysis over the entire period should not be undertaken.

III

V&L suggest that much of the history of the secular trend of nutritional status in England can be explained by inoculation against smallpox. In this regard they invoke the study of Floud and Wachter who estimated that the heights of the Marine Society boys increased by about 4 inches within the birth decade after 1800.[20] In order to verify this result, we re-estimated the height trend using the modes of the kernel density estimator for five-year birth cohorts separately by age.[21] In contrast to V&L's assertion that heights improved 'steadily', the well-known deterioration of nutritional status in the late eighteenth century becomes obvious (figure 6A). An application of the more traditional method of Komlos and Kim[22] confirmed the negative trend, which is also in accordance with a decline in real wages in late eighteenth-century London (figure 6B).

The estimation of the height trend in the early nineteenth century is more complicated because of the structural break in the minimum height requirement in 1814 (see table 1). The Marine Society obviously became more selective in its admission policy. Figure 7 gives an example of the enormous differences in height distributions before and after the structural break.

It is important to note that at the time of this structural break, the decade of the 1810s, nearly all boys were afflicted with smallpox. This again implies that the increase in height in this period was not caused by the introduction of inoculation, but by the dramatic change in the minimum height requirement.

Mode estimations are unreliable after the structural break, because the true mode might have been within the shortfall region. Therefore we analyse the height trend for the nineteenth century only with the Komlos and Kim method.[23] The results suggest that there was a recovery in heights between the 1800s and the 1820s which can be explained by the

---

Although smallpox has no impact on height, a strong negative impact is estimated. This is because the smallpox variable is a proxy for a change in the sampling bias. If we equalize this bias by truncating the sample below 59.5, the artificial correlation of height and smallpox is avoided: $height = 61.9 - 0.2 smallpox$, with a t-value reducing to -0.8 (p-value 0.43).

[18] Cf. n. 17; see also Greene, *Econometric analysis*, sec. 22.3a.

[19] Greene, *Econometric analysis*, sec. 22.3b.

[20] Floud and Wachter, 'Poverty and physical stature'.

[21] We omitted estimations based on subsamples with a sample size below 30. We also omitted mode estimations with very skewed and bimodal samples such as the first three birth cohorts of the 15-year-olds. Detailed graphs are available from the authors.

[22] Komlos and Kim, 'Estimating trends'. As mentioned above, n. 5, there is some evidence of right-hand shortfall with the 15- and 16-year-olds (see also figure 5B). To avoid a determination of an additional right-hand truncation point, we did not analyse these sub-samples.

[23] Like Komlos, we found no evidence of right-hand shortfall in the nineteenth-century sample.

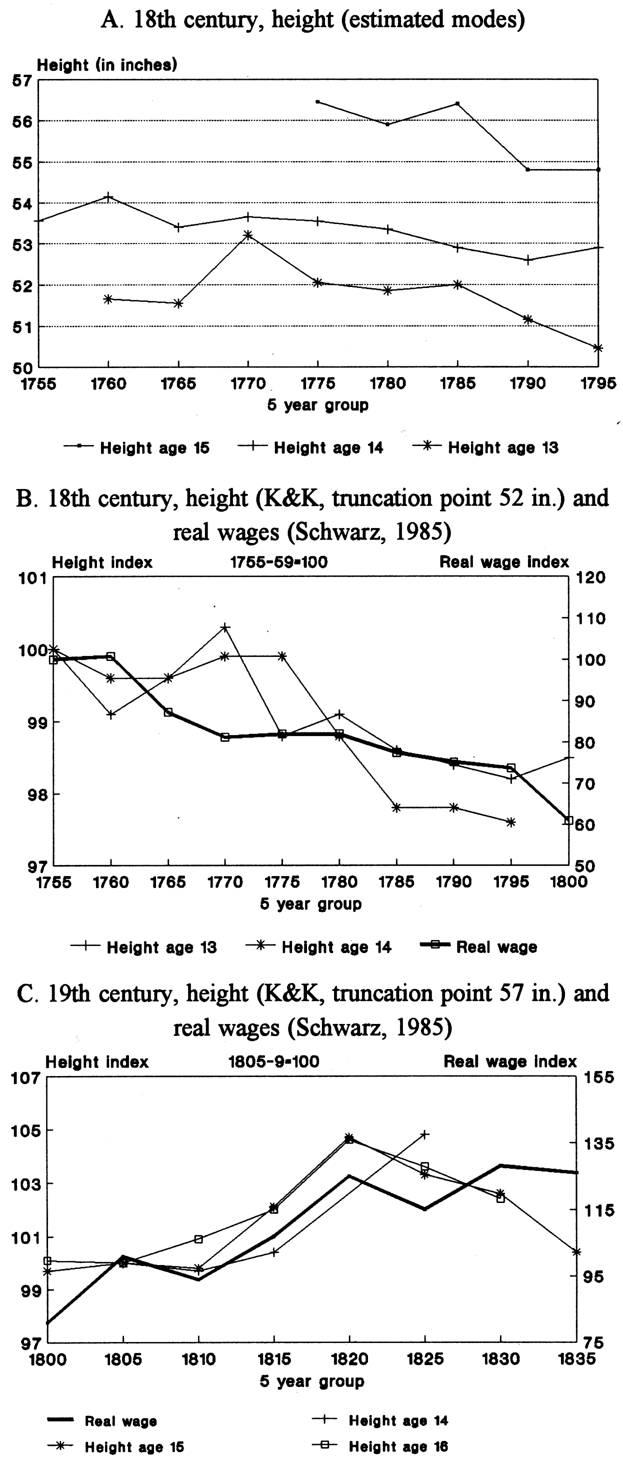

Figure 6. *Trends in heights in London, 1755-1835*

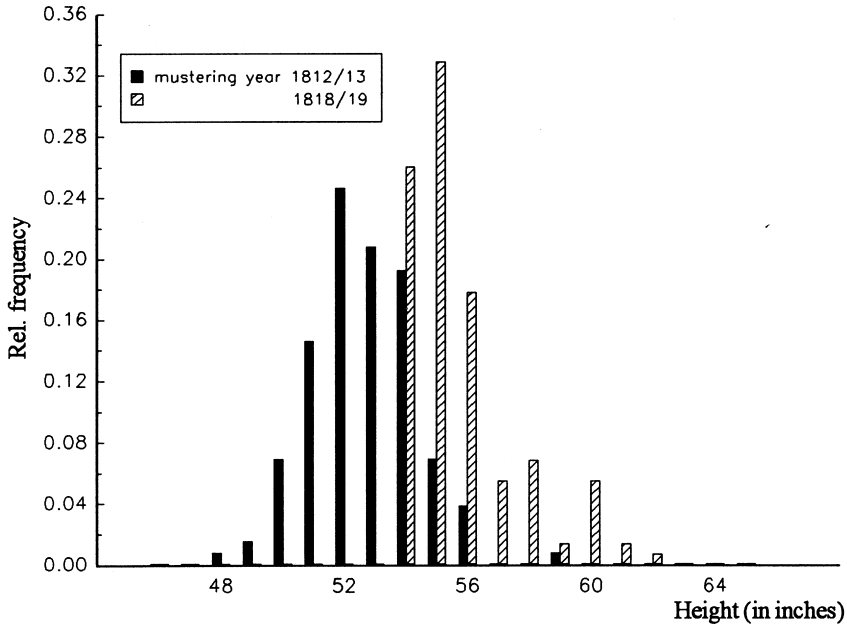

Figure 7. *Heights of 14-year-olds before and after the structural break (smallpox cases)*
Source: see text

movement of real wages (figure 6C).[24] Might the higher nutritional status in the 1820s contribute to the contraction of smallpox, in addition to the inoculation campaigns? Most physiologists suggest that smallpox is one of the few infectious diseases which is not strongly influenced by nutritional status.[25]

In sum—contrary to V&L's assertion—the conclusion of a negative impact of smallpox on height cannot be sustained using the Marine Society sample: either in a cross-sectional analysis or in a time series context. The regression analysis performed by V&L suffers from serious mistakes such as misspecification and (time dependent) distortion. Therefore, their far-reaching conclusions are without empirical or statistical foundation. Furthermore, we find that, first, as a result of enormous changes in the height standards the Marine Society dataset is not a reliable source from which to estimate regressions over the entire period 1770-1873; secondly, heights of the Marine Society boys declined in the late eighteenth-century birth cohorts; and finally, between the birth

[24] For an examination of height trends across the structural break other datasets are more useful: Komlos, for example, estimated a modest recovery using heights of army soldiers; see Komlos, 'Secular trend'. Comparing the height of adult recruits in the East India Company born around 1800 with those born around 1820 (i.e. those recruited in the 1820s and 1840s), a weighted average height increase of 1.6 cm (Ireland) and 1.5 cm (England) is obtained from Mokyr and O Grada, 'Height and health', p. 159.

[25] See, for example, Nunn, 'Nutrition, immunity, and infection', pp. 136-40; Landers, *Death and the metropolis*, p. 301.

cohorts of the 1800s and 1820s there is a certain recovery of height, as there is a recovery of real wages.

University of Munich

Footnote references

Baten, J. and Heintel, M., 'Zum Problem der Verteilungen mit Shortfall bei der Nutzung des Indikators Durchschnittliche Koerpergroesse', *Hist. Soc. Res.*, 20 (1995), pp. 135-54.

Floud, R. and Wachter, K. W., 'Poverty and physical stature: evidence on the standard of living of London boys, 1770-1870', *Soc. Sci. Hist.*, 6 (1982), pp. 422-52.

Greene, W. H., *Econometric analysis* (New York, 1993).

Heintel, M., 'Historical height samples with shortfall: a computational approach', *Hist. & Computing*, 8 (1996), pp. 30-41.

Heintel, M., 'Estimating means and trends of normal samples with shortfall', *Münchener Wirtschaftswissenschaftliche Beitrage*, 96-20 (1996).

Komlos, J., 'The secular trend in the biological standard of living in the United Kingdom, 1730-1860', *Econ. Hist. Rev.*, XLVI (1993), pp. 115-44.

Komlos, J. and Kim, J. H., 'Estimating trends in historical heights', *Hist. Methods*, 23 (1990), pp. 116-20.

Landers, J., *Death and the metropolis: studies in the demographic history of London, 1670-1830* (Cambridge, 1993).

Mokyr, J. and O Grada, C., 'Height and health in the United Kingdom, 1815-1860: evidence from the East India Company army', *Exp. Econ. Hist.*, 33 (1996), pp. 141-68.

Nunn, P., 'Nutrition, immunity, and infection', in R. Schofield, D. Reher, and A. Bideau, eds., *The decline of mortality in Europe* (Oxford, 1991), pp. 131-45.

Scott, D. W., *Multivariate density estimation* (New York, 1992).

Voth, H. J. and Leunig, T., 'Did smallpox reduce height? Stature and the standard of living in London, 1770-1873', *Econ. Hist. Rev.*, XLIX (1996), pp. 541-60.

Wachter, K. W. and Trussell, J., 'Estimating historical heights', *J. Amer. Stat. Assoc.*, 77 (1982), pp. 279-303.