Height, income, nutrition, and smallpox in the Netherlands: the (second half of the) 19th century

Jan Jacobs* and Vincent Tassenaar†

December 2002

Abstract

This paper explores the relationship between height and three of its determinants, GDP per capita, nutrition and smallpox, explicitly paying attention to dynamics involved in the curve of growth. The relation is investigated using recently constructed data on the nineteenth century for the Netherlands.

Keywords: height, GDP per capita, nutrition, smallpox, 19th century, the Netherlands

JEL-code: N13

Acknowledgements

Earlier versions of this paper were presented at the joint Economic History colloquium of the University of Groningen, October 2001, the European Social Science History Conference (ESSHC), The Hague, The Netherlands, February 2002, the first International Conference on Economics & Human Biology, Tübingen, Germany, July 2002, and the NAKE Research Day, Amsterdam, the Netherlands, October 2002. We have also benefited greatly from comments provided by J.W. Drukker, Bernard Harris and in particular John Komlos.

* Department of Economics, University of Groningen, P.O. Box 800, NL-9700 AV Groningen, the Netherlands [j.p.a.m.jacobs@eco.rug.nl].

† Department of Economic and Social History, Faculty of Arts, University of Groningen, P.O. Box 716, 9700 AS Groningen, the Netherlands [p.g.tassenaar@let.rug.nl].
1. Introduction

In the field of anthropometric history stature as a measure of standard of living attracts a lot of attention, see for example the overview of Steckel (1995) and the references therein. It is well-understood that heights are influenced by labour intensity, morbidity and nutrition; improvements in dietary intake, public sanitation, and medical technology play a role too (see e.g. Tanner 1981, Komlos 1989, and Floud et al. 1990).

This paper explores the relation between height and three of its determinants, GDP per capita, nutrition and smallpox. The relationship between height and the first two determinants has been studied quite extensively. An early contribution is Brinkman et al. (1988). Other examples are Coll (1998), Craig and Weiss (1998) and Haines (1998) in the collection of Komlos and Baten (1998), Mosk (1996) and the recent papers of Baten (2000) and Baten and Murray (2000). The relationship between smallpox and height received a lot of attention too. The conclusion of Voth and Leunig (1996) that smallpox had a significant negative impact on height in Britain was questioned by e.g. Razzell (1998), Baten and Heintel (1998) and Oxley (2002).

Attained height is a function of its determinants during the years of growth. So, lags enter the empirical model. However, in our opinion the issue of dynamics has not yet been dealt with convincingly. All empirical studies adopt a fixed lag scheme for the determinants of height, i.e. assume that the curve of growth does not shift over time. However there is ample evidence from medicine that the shape of the curve of growth shifts over time (Ljung et al. 1974; Eveleth and Tanner 1990; Wit et al. 1999; Fredriks et al. 2000).

We will derive a dynamic model for height and its determinants starting from the identity that the measured heights are the sum of the height increments of the years from birth onwards. The only assumption we adopt is that height increments of a given year are influenced by the environmental circumstances of that year. Unfortunately, the empirical model we end up with cannot be tested with aggregate data in which we have only one observation on height for each period. Ideally, we need information on the complete growth profile, like e.g. the Stuttgart schoolboys sample described in Komlos et al. (1992).

Therefore we try to illustrate the shifting lag pattern in an indirect way. Following the lead of Woitek (1998) we calculate cyclical components of 19th century series on height, GDP per capita and nutrition in the Netherlands. Whereas Woitek uses spectral analysis to investigate the cyclical components, we apply more conventional correlation statistics to bring to the fore the evolving dynamic relations between them. 3D-plots of correlations coefficients between the cyclical components of height and the (lagged) cyclical components in GDP per capita and nutrition over a moving thirty-year window make the shifting lag patterns visible to some extent.

The height-smallpox relationship is treated in a different manner, because of the typical pattern of the smallpox series in the Netherlands. The Dutch data do not support the conclusion of Voth and Leunig (1996) that smallpox had a serious negative effect on height. The smallpox outbreak in the beginning of the 1870s may have led though to a stagnation in height increases in the late 1880s till the mid-1890s.
The remainder of the paper is structured as follows. The next section presents the curve of growth and cites some evidence from the literature why this curve is not time-invariant. Section 3 derives our empirical model for height. Section 4 illustrates the shifting lag pattern with recently constructed time series on the 19th century in the Netherlands. Section 5 concludes.

2. The curve of growth

Two periods of intense activity characterize the growth process following birth (Tanner 1989). The change of height is greatest during infancy, falls sharply, and then declines irregularly into the childhood years. During adolescence velocity rises sharply to a peak that equals approximately one-half of the velocity during infancy, then declines rapidly and reaches zero in adulthood. The thick line in Figure 1 represents a stylized growth curve.

The height of an individual reflects the interaction of genetic and environmental influences during the process of growth. If a particular stimulus is lacking at a time when it is essential for the child (times known as “sensitive periods”) then the child’s development may be stunted as it were, from one line to another (Eveleth and Tanner 1990). Young children and adolescents are particularly susceptible to environmental conditions. The return of adequate nutrition following a period of deprivation may restore normal length through catch-up growth. If conditions are inadequate for catch-up, individuals may approach normal adult height by an extension of the growing period by as long as several years. Prolonged and severe deprivation results in stunting, or a reduction in adult size (Steckel 1995). The solid line in Figure 1 illustrates this. Suppose that environmental circumstances are bad in the years immediately after birth, then the height increase in that period is stunted. The reduction in height can be recovered if conditions become more favourable in the next couple of years, but not completely: the surface below the solid line at the left is smaller than the surface below the thick solid line in the same period. The same applies for an adverse shock in environmental circumstances in the sensitive period of adolescence. Here height increments may become smaller too, and the reduction can be regained afterwards if things turn to the better by an extension of the period of growth.

The figure further illustrates the importance and the sensitivity of the age of measurement. Individuals measured at age of twenty instead of nineteen may easily gain an additional two to three cm in height, cf. Section 4.1 below. This observation explains in part the finding of Baten (2000) that if the individuals are measured after having reached their final height, then the environmental circumstances of the first years after birth have the stronger influence.
3. The empirical model

Let $H_{\tau}^t$ be the average height of conscripts at age $\tau$ measured in year $t$, which is observed from $t = 1, \ldots, T$. The attained height at age $\tau$ is by definition equal to the increments in stature from the year of birth

$$H_{\tau}^t \equiv \nabla H_{\tau}^t + \nabla H_{\tau}^{t-1} + \ldots + \nabla H_{\tau}^{1} + \nabla H_{\tau}^{0},$$

where $\nabla H_{\tau}^{\tau-i} \equiv H_{\tau}^{\tau-i} - H_{\tau}^{i-1}$ is the increment in height of conscripts measured in year $t$ between age $\tau - i$ and $\tau - i - 1, i = 1, \ldots, \tau$ and $\nabla H_{\tau}^{0}(\equiv H_{\tau}^{0})$ is the height at birth. We assume that the (unobserved) increments in height depend linearly on income $Y$, or more precisely

$$\nabla H_{\tau}^{\tau-i} = \alpha_{ti} Y_{t-i} + \epsilon_{t-i},$$

where $\epsilon_{t-i}$ is an error term. Of course, the framework can easily be extended to allow for more than one explanatory variable. Substituting Equation (2) into (1) gives

$$H_{\tau}^t = \sum_{i=0}^{\tau} \alpha_{ti} Y_{t-i} + \sum_{i=0}^{\tau} \epsilon_{t-i} \equiv \alpha_t(L) Y_t + \epsilon_t,$$

where $L$ is the lag operator $LY_t = Y_{t-1}$ and $\epsilon_t \equiv \sum_{i=0}^{\tau} \epsilon_{t-i}$ is a moving average error expression. The matrix polynomial $\alpha_t(L)$ captures the curve of growth: $\alpha_{ti}$ are relatively higher in the sensitive periods of the process of growth. As the subscript $t$ indicates, we do not assume that the curve of growth is time-invariant.

Since the height series is observed from $t = 1, \ldots, T$, the height-income relation of Equation (3) cannot be estimated without making further assumptions; we only have $T$ observations to estimate $T \times (\tau + 1)$ parameters. Brinkman et al. (1988) make two additional assumptions: (i) the curve of growth does not shift in time: $\alpha_t(L) = \alpha(L)$; (ii) the growth curve can be modelled by an appropriate polynomial lag for the matrix polynomial $\alpha(L)$. Under these assumptions Equation (3) simplifies to the familiar model

$$H_{\tau}^t = \alpha(L) Y_t + \epsilon_t.$$

The first assumption allows Brinkman et al. to assign weights to different ages that might affect height: the Yearly Age- and Sex-Specific Increase in Stature (YASSIS), our $\alpha(L)$. They estimated the YASSIS terms by a third degree polynomial lag. Coll (1998) simplified their method by assigning a weight scheme with weights fixed for a number of years. By this he avoided the polynomial lag. Baten (2000) observes that the form of the second assumption does not generate statistically significant differences in outcomes. Below we look for evidence of shifting lag patterns in the height-income and the height-nutrition relationship.
4. Illustration

4.1 Data

*Height.* We use the median height of Dutch conscripts series as estimated by Drukker and Tassenaar (1997). Shifts in the year of measurement complicate the analysis of height of conscripts data. The Dutch government changed the military laws in 1861 because of the rising percentage of undersized conscripts. From 1863 on, the age of recruitment was raised from 19 to 20. The top panel of Figure 2 shows the median height series. The change in the recruitment age results in a height increase of two to three centimeters. In our computations below we only use the second part of the sample, the 1863-1913 period.

![Figure 2 about here.]

The median stature of Dutch conscripts increased substantially and almost uniformly after 1860, even if the gains during the initial decade (1863-1870) were only a recovery to previous levels. Average height reached its nadir in 1857 (Tassenaar 2000) after a period of decline starting in the 1840s. The phenomenon of decreasing heights in European countries in the early-industrial revolution era is known as the “shrinking paradox” (Komlos 1998). In the 1860s height recovered somewhat in comparison to the late fifties, but stayed well below previous levels. The upward trend started with the cohorts of the 1870s born in the 1850s. In the early 1880s median height surpassed levels reached during the first half of the nineteenth century. Thereafter it rose continuously. Until WW I only very short cyclical setbacks can be observed.

*GDP per capita.* Our real GDP per capita series is constructed by Smits et al. (2000), see Figure 3. The discontinuity around 1840 is probably caused by improper accounting for the separation of Belgium. Apart from that their estimates are quite plausible and indicate that real GDP per capita had an almost constant growth rate from the end of the Napoleonic era until World War I.

![Figure 3 about here.]

*Nutrition.* Our nutrition series originates from Knibbe (2001) and covers the period 1852-1913. Daily caloric intake started to rise from 1855 until the end of the 1880s, followed by a levelling-off for almost a decade. From the midst of the nineties caloric intake rose again until the First World War, with a short dip in the years 1907-1909. During the whole period the nutrition situation improved even in some of the stagnation years.

![Figure 4 about here.]

*Smallpox.* Our smallpox data come from Rutten (1997). Smallpox were prevalent during most of the 19th century. Vaccination began in the beginning of that century, so smallpox was more or less under control. The exception to the rule is the smallpox pandemic of 1870-1872, with over 20,000 victims, for the greater part young children in the age of one
to three years. A striking fact is that the three largest cities in the Netherlands, Amsterdam, Rotterdam and the Hague were not severely hit by the smallpox outburst. Mortality was highest in the provinces Utrecht, North-Holland south of the river Y and South-Holland north of the river Meuse, the area now known as Randstad and the Green Heart. Rutten (1997, Section 11.3) lists a number of explanations. First, mortality of smallpox was highest in areas where vaccination was optional and lowest where vaccination was compulsory. Secondly, population density plays an important role and thirdly, the distance to smallpox centres at home and abroad. Finally, factors in the biological environment were important.

[Figure 5 about here.]

Figure 5 shows the number of smallpox victims in the Netherlands and in the three cities Amsterdam, Rotterdam and the Hague. The smallpox pandemic in 1870-1872 may have resulted in the stagnation in height increases starting in the late 1880s until the mid-1890s. Smallpox are excluded from our analysis of cyclical components to which we turn now.

4.2 Cyclical components

The Hodrick-Prescott (HP) filter. To analyse the existence of shifting lag patterns in the height-income and the height-nutrition relationship we calculate cyclical components of these series. To that purpose we apply the Hodrick-Prescott (HP) filter, which works as follows. Assume that an observed time series \( y_i \) can be represented as the sum of a cyclical component \( c_i \) and a trend or growth component \( t_i \)

\[
y_i = c_i + t_i.
\]

Let \( \lambda \) be a parameter that reflects the relative variance of the trend component to the cyclical component. The parameter \( \lambda \) determines the penalty for adjusting the trend, in other words \( \lambda \) controls the smoothness of the trend. Given a value for \( \lambda \), the HP filtering problem is to choose the trend component \( t_i \) to minimize for \( c_i \) the following loss function

\[
\sum_{i=1}^{T} c_i^2 + \lambda \sum_{i=1}^{T} [(t_{i+1} - t_i) - (t_{i} - t_{i-1})]^2.
\]

In this minimization problem there exists a trade-off between the extent to which the trend component tracks the actual series and the smoothness of the trend. For \( \lambda = 0 \) there is no penalty on trend adjustment, so the trend component is simply the observed series. As \( \lambda \) goes to infinity, the trend component approaches a linear trend. The value of \( \lambda \) depends on the frequency of the observations. For annual data the choice of \( \lambda = 100 \) is customary.

We prefer the HP filter to compute cyclical components despite the fact that the filter is not beyond criticism (Jacobs 1998, p.53). Bonenkamp et al. (2001) observe that the HP-filter produces more or less similar turning point chronologies in the annual 19th century GDP per capita series in the Netherlands as the theoretically more appealing band-pass filter proposed by Baxter and King (1999).
Cyclical components of height, GDP per capita and nutrition, 1863-1913. The bottom panels of Figures 2, 3 and 4 show the cyclical components in height, GDP per capita and nutrition, respectively, which are calculated by applying the HP-filter on the natural logarithms of the series. For convenience we combine the cycles into one figure, Figure 6. We clearly observe cyclical patterns; nutrition cycles have a higher frequency than the cyclical components in the other two series.

Correlation between cyclical components. To investigate the existence of shifting lag patterns between height and two of its determinants, we calculate correlation coefficients between cyclical components of heights and (lagged) cyclical components of GDP per capita, Figure 7, and nutrition, Figure 8. We take a thirty-year window. For that window we calculate the correlation between the cyclical components in height and GDP per capita (nutrition). Then we lag the height (nutrition) series by one year and calculate the correlation coefficient again. We continue taking lags up to and including 21 years (the examination age plus one year). The correlation coefficients calculated in this manner are shown as the line that is closest to the axis labelled lag belonging to sample=1 in the figures. We only show significant (1%) correlation coefficients in the figures using the statistic \((N−2)r^2/\sqrt{1−r^2}\), which follows a \(t\)-distribution with \(N−2\) degrees of freedom, where \(N\) is the number of observations and \(r\) is the correlation coefficient.

In Figure 7 the first window refers to the 1864–1893 period. The next step is to move the window by one year, i.e., we look at 1865–1894, and repeat the analysis calculating the (significant) correlation coefficients between the cyclical components in heights and the (lagged) cyclical components in GDP per capita (sample=2). We keep on moving the window until the final year of the window coincides with the final observation in our sample. The final sample, the 21th, is the 1884–1913 period. Since our nutrition series is shorter, Figure 8 shows less samples for the correlation coefficients. The first sample refers to the 1874–1902 period; the 11th and last the 1884–1913 period.

A puzzling property of the correlation surfaces of Figures 7 and 8 is the occurrence of negative correlation values at certain lags. The canyon in the height-income graph, Figure 7, around lag 10 is a typical example. We haven’t yet found a satisfactory explanation for this phenomenon. The straightforward interpretation of positive economic conditions ten years before examination having a negative influence on height is clearly not true. A candidate explanation might be the following one. The correlation coefficient between two series that are perfectly synchronised is equal to one, but if we compute correlations between one series and another series that is lagged up to the duration of the full cycle, the correlations fluctuate between 1 and -1. Taking into account that in the process of
growth a person cannot become smaller, we do not feel too uncomfortable focussing on
the positive values of the correlations and ignoring the negative ones. But of course, any
real explanation for this intriguing phenomenon is appreciated.

With these caveats the correlation surfaces of Figures 7 and 8 allow the following observ-
ations.

(i) All samples have positive correlation coefficients between cyclical components in height
and GDP per capita at the second lag, i.e. at age 18. There is some evidence of positive
correlation at lag 15 (age 4-5) which a shift to lag 16 (age 3-4) in later samples. Further-
more, we find positive correlations at lag 21, which corresponds to the year before birth.
(ii) We find positive correlation coefficients between cyclical components in height and
nutrition at lag 13, which corresponds to nutrition conditions experienced at the age of
6-7. This outcome is robust; it is observed in all samples. In addition, the correlation is
positive at lag 4 in the first six samples; but we observe a shift to the seventh lag in the last
two samples.

How do these observations correspond to the curve of growth as sketched in Section 2?
The lags observed in the correlation between (cyclical components in) height and GDP
per capita fit in nicely: the lag of 2 years corresponds to the adolescence growth spurt
and the lags of 15 and 16 years to a childhood growth period. This observation is in line
with Baten (2000). We do not find a significant birth-year effect. The height-nutrition
correlation outcomes are more difficult to interpret. The positive correlation observed at
lag 13, corresponding to nutritional conditions at the age of 6-7, is hard to reconcile with
the curve of growth. This might be caused by the nature of the nutrition series that we used
here. Nutrition is taken to be the average daily caloric intake, perhaps too rough a measure.
We experimented with protein, a second series provided by Knibbe (2001). This did not
lead to qualitatively different conclusions.

In our opinion, the figures provide some evidence of shifting lag patterns, but the evidence
is poor. Estimation of height-income relationships assuming that the curve of growth is
time-invariant does not lead to serious errors for this period. However we expect to find a
clearer shift when analysing an extended period.

5. Conclusion

The aim of this paper was to explore the dynamic relationship between height and three
of its determinants, GDP per capita, nutrition and smallpox. Empirical height studies typi-
cally assume that the curve of growth is fixed in time. We investigated the consequences of
relaxing the assumption of time-invariant lag patterns and tried to find evidence for shift-
ing lag patterns using a recently constructed data set on 19th century conscript heights in
the Netherlands.

The impact of smallpox on conscript heights was analysed separately. Because vaccination
was well-spread, smallpox hardly existed in the Netherlands in the nineteenth century.
There was one exception, the outbreak in the early 1870s. Eye-balling graphs suggests a
relation between this outbreak and the slowdown of height increases some fifteen to twenty years later.

To assess the significance of shifting lags we calculated correlation coefficients between cyclical components in height and (lagged) GDP per capita and nutrition cycles. The correlation between cyclical components in height and lagged GDP per capita was strongest at lags corresponding to early childhood and adolescence, two well-known sensitive growth periods. Perhaps surprisingly, the height-nutrition correlation outcomes did not fit nicely in a growth curve pattern.

We found some evidence of a shift in the lag pattern between height cycles and GDP per capita and nutrition cycles. This evidence is not very strong, though. So, for the period we investigated the relation between Dutch conscripts heights and two of its determinants, the second half of the nineteenth century, the ‘standard’ estimation method assuming a time-invariant growth curve seems justified. However, more research is needed, covering longer periods and other countries, to dismiss our hypothesis of shifting lag patterns in the relation between height and standard of living indicators completely.
References


List of Figures

1. A stylized growth curve (thick line) and catch-up growth (solid line)  13
2. Average heights of conscripts at age 19 (before 1862) and at age 20 (after 1862); levels (top panel) and cyclical components (bottom panel)  14
3. GDP per capita; index numbers 1815=100 (top panel) and cyclical components (bottom panel)  15
4. Nutrition; kcalories per day (top panel) and cyclical components (bottom panel)  16
5. Smallpox victims in the Netherlands and three cities (Amsterdam/Rotterdam/the Hague)  17
6. Cyclical components in height, GDP per capita, and nutrition  18
7. Correlation between height cycles and (lagged) GDP per capita cycles (1% level)  19
8. Correlation between height cycles and (lagged) nutrition cycles (1% level)  20
Figure 1: A stylized growth curve (thick line) and catch-up growth (solid line)

Source: Tanner (1989)
Figure 2: Average heights of conscripts at age 19 (before 1862) and at age 20 (after 1862); levels (top panel) and cyclical components (bottom panel)

Source: top panel Drukker and Tassenaar (1997); bottom panel see Section 4.2
Figure 3: GDP per capita; index numbers 1815=100 (top panel) and cyclical components (bottom panel)

Source: top panel Smits et al. (2000); bottom panel see Section 4.2
Figure 4: Nutrition; kcalories per day (top panel) and cyclical components (bottom panel)

Source: top panel Knibbe (2001); bottom panel see Section 4.2
Figure 5: Smallpox victims in the Netherlands and three cities (Amsterdam/Rotterdam/the Hague)

Source: Rutten (1997)
Figure 6: Cyclical components in height, GDP per capita, and nutrition
Figure 7: Correlation between height cycles and (lagged) GDP per capita cycles (1% level)
Figure 8: Correlation between height cycles and (lagged) nutrition cycles (1% level)