Please Note:

This is the last working paper version before this study was submitted and accepted. Please cite as Guntupalli, Aravinda and Baten, Joerg. "The Biological Standard of Living in Europe During the Last Two Millennia", *European Review of Economic History* 9-1 (2005), pp. 61-95.

Wirtschaftswissenschaftliche Fakultät
der Eberhard-Karls-Universität Tuebingen

The Biological Standard of Living in Europe During the Last Two Millennia

Nikola Koepke
Joerg Baten

Tuebinger Diskussionsbeitrag Nr. 265
Tuebingen Economics Department Working Paper No. 265
September 2003

Wirtschaftswissenschaftliches Seminar
Mohlstraße 36, D-72074 Tuebingen

Nikola Koepke and Joerg Baten, Univ. Tuebingen and CESifo

The Biological Standard of Living in Europe During the Last Two Millennia

ABSTRACT

This paper offers the first anthropometric estimates on the biological standard of living in central Europe in the first millennium, and expands the literature on the second millenium. The overall picture is one of stagnant heights. There was not much progress in European nutritional status, not even between 1000 and 1800, when recent GDP *per capita* estimates arrive at growing figures. We find that heights stagnated during the Roman imperial period in Central, Western and Southern Europe. One astonishing result is the height increase in the fifth and sixth centuries. Noteworthy is the synchronicity of the height development in three large regions of Europe. In a regression analysis of height determinants, population density was clearly economically (but not statistically) significant. Decreasing marginal product theories and Malthusian thought cannot be denied for the pre-1800 period. Of marginal significance were climate (warmer temperatures were good for nutritional status), social inequality and gender inequality (both reduce average height).

MAIN QUESTIONS

How did the standard of living develop in the very long run? We distinguish the biological standard of living (that includes important elements of the human utility function such as health, longevity, and quality of nutrition) from purchasing power oriented components of the standard of living.[1] In this study, we use adult stature as a proxy for the biological components and analyze 9477 height measurements that are mostly based on skeletons. Some of the heights were published only in an aggregated way by previous investigators and excavators: thus we have 2974 height values that refer to either one or two to 360 individuals. Our data stem from 314 sites all over Europe. There are important deviations between the biological and purchasing power related components of living standards during the past 2000 years: for example, during the time period of the Roman Empire mean stature did not increase. This stands in contrast to the widely accepted view that purchasing power should have increased due to the "Smithian Growth"[2] during the Roman Empire. It is also remarkable that heights increased in the fifth and sixth centuries, after the breakdown of the Empire in the West.

A related study by Steckel (forthcoming) found a substantial decline in Scandinavian heights since the Middle Ages.[3] In a similar vein, Cohen and Armelagos (Cohen [1989], Cohen and Armelagos [1984]) and the research group around them stressed the detrimental influence of civilization on human health and nutritional status over the millennia (until the late 19th century). Maat (2003) confirmed this view recently for the Netherlands (see also Brothwell [2003]). However, the number of observations is very limited in these studies, especially for the periods before the Middle Ages. Our study will use a substantial number of sites and observations to trace the biological standard of living over the last two millennia.

After describing the sample characteristics, we will firstly estimate a general height trend and more disaggregated trends in three European regions. The time series of the latter estimation will be further analyzed, regressing it on population density, climate, social and gender inequality, and similar variables (that are measured with substantial measurement error). We finally compare our height trends with purchasing power estimates and speculate about implications for other debates such as the "Great Divergence" issue.

DATA AND REGIONAL CHARACTERISTICS

Our height data stem mostly from archaeological excavations. We have the largest number of observations on Central and Western Europe (see Table 1): today's Germany, Benelux, Austria, Northern France, Switzerland, and the UK.

Table 1

AREAS COVERED BY THE DATA SET (NUMBER OF INDIVIDUALS)

Century	Central /Western Europe				Eastern/ Northern Europe		Mediterranean	Total
	Bavarian/ Austrian region	Northern Rhine region	Southern Rhine region	UK	Eastern Europe	Northern Europe	Mediterranean region	
1	21	21	14	2	36	72	95	261
2	56	6	210	711		14	11	1008
3	16	95	24	184	37	46	50	452
4	361	1	26	227	65		4	684
5	3		113	9	68	5	164	362
6	338	208	380	2	99	7	35	1069
7	146	35	456	2	218	5	7	869
8	225		20		179			424
9	78	5	36	164	135	251	12	681
10		49	24		1	133		207
11	1		14	4	229	823		1071
12	1		18		20	440	125	604
13	3	12	9	174	29	53		280
14		113	3	4	6	547	7	680
15		55			6	11	29	101
16	73	55	314		17		3	462
17	21	19	39		39	41		159
18		103						103
	1343	777	1700	1483	1184	2448	542	9477

Sources: see www.uni-tuebingen.de/uni/wwl/twomillennia.html

The number of cases is sufficient to subdivide this Central/Western region further: The regions along the Northern Rhine area (Benelux, Northern France, Western Germany), the regions around the Southern Rhine (South-Western Germany, Switzerland, North-Eastern France), Bavaria/Austria, and the UK. "Northern and Eastern Europe" denotes those regions that had only little contact with the Roman Empire and its provincial economy: Scandinavia, Poland, Northern-, Eastern Germany and Hungary. Hungary is included here, because the large region East of the Danube was not fully integrated into the Roman imperial economy. Northern and Eastern Europe was the least densely settled region in our sample, and it probably had the highest milk consumption per capita, because a high land ratio allows *ceteris paribus* a large number of cows and a nutrition with high-quality proteins (on the role of high quality proteins (such as milk) for pre-industrial nutrition and health, see Baten [1999], Baten and Murray [2000], Craig et al. [2002], and many others).

Even today, Scandinavians (and Dutch) have the highest milk and dairy product consumption per capita, while East Asians have the lowest. The West-Mediterranean dairy consumption was extremely low during the years 0 – ca. 1950 (parallel to heights) but has strongly increased thereafter (Quiroga [1998]. Bisel [1988]. Ancient sources such as Tacitus and Plinius [Tac. Germ. 23; Plin. nat. VIII 179] emphasized this as well). The "Mediterranean" region of our sample includes Italy, Spain, Portugal and the Balkans. Obviously, the small number of observations in Table 1 indicates that more studies on the Mediterranean region are welcome and we are open for future co-operations.

For the early Middle Ages, data are quite abundant.[4] After the 12th century, height data are scarcer because bones in cemeteries were more often lost or mixed with bones of later epochs. Beginning in the 17th and 18th century, archival (written) sources are available providing much larger sample sizes but also additional selectivity and truncation problems (see Komlos [2003], forthcoming). Because the period of the 18th to 20th centuries is relatively well studied, we focus in this paper mainly on the earlier centuries.

The sample consists of 2938 female and 6539 male heights that are quite well distributed: all major time periods are more or less covered. Only the 17th and 18th centuries have an insufficient number of cases for women (12 and 0, respectively).[5] The excavators and original investigators aggregated a large proportion of the overall 9477 height measurements. Wherever possible, we collected disaggregated figures. 2974 separate height numbers are in the final database, after discarding extreme heights (<145 cm, > 200 cm). When the dating was imprecise, we used the average between the earliest and the latest date that the principal investigators mentioned because the true date is as likely to be before as after the average. We experimented with estimation techniques that gave smaller weight to imprecisely dated observations or discarded them completely, but the main results are robust.[6] Because of these data limitations our unit of analysis is restricted to the century. We organized all heights by birth year (or birth century) and discarded still growing and old individuals (< 23 years, > 59 years).

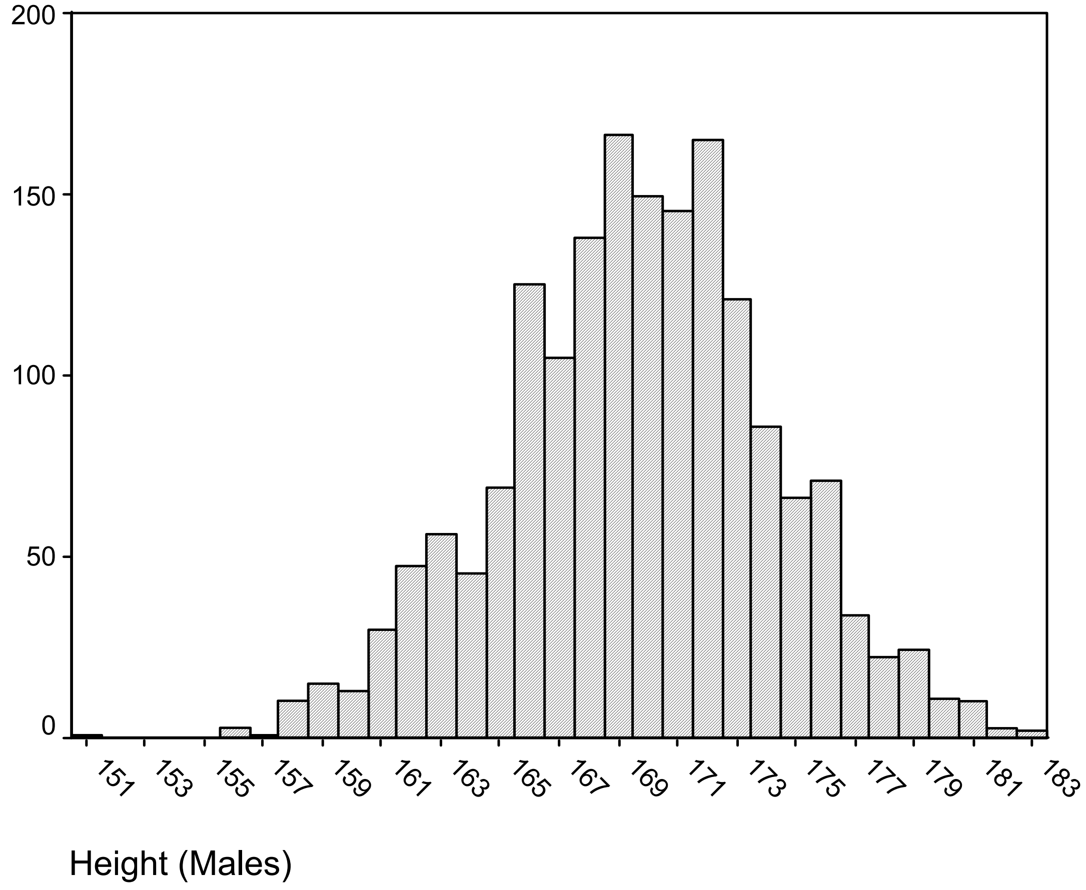

Figure 1a

MALE HEIGHT DISTRIBUTION, ALL CENTURIES

Source: see Table 1

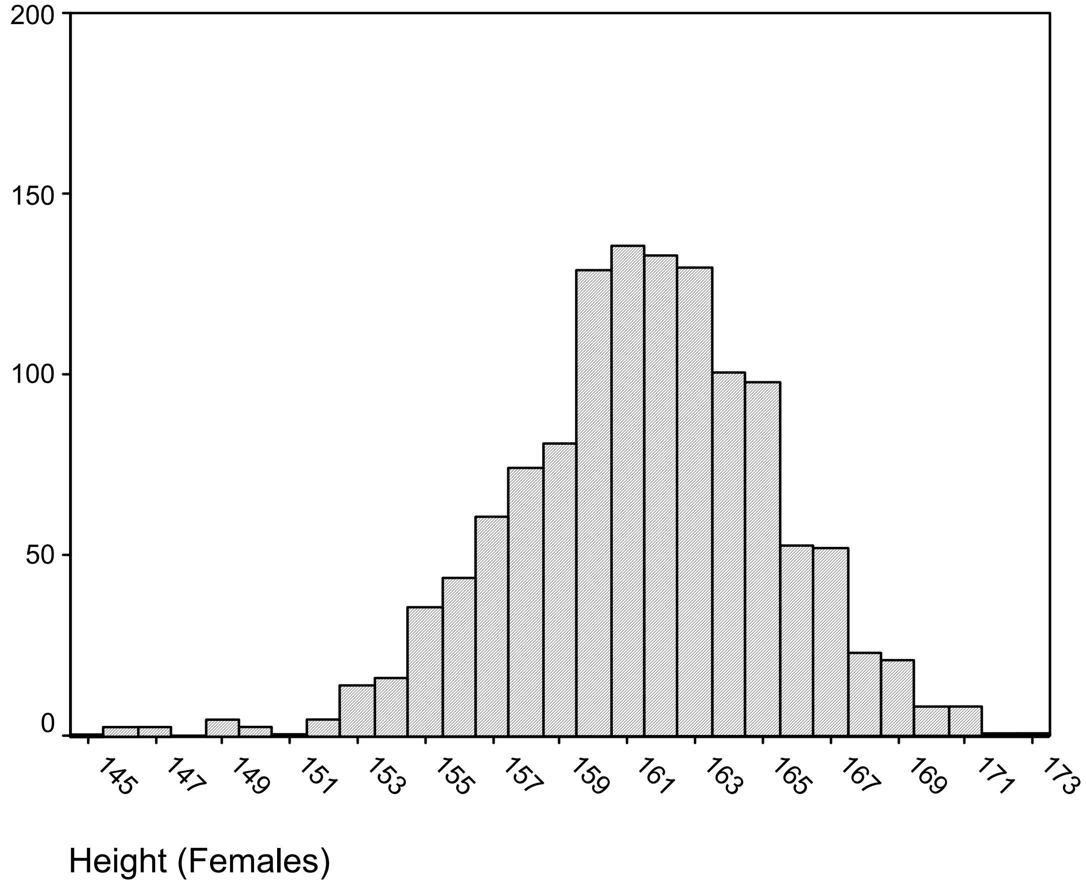

Figure 1b

FEMALE HEIGHT DISTRIBUTION, ALL CENTURIES

Source: see Table 1

Heaping and truncation did not play a large role, as illustrated by the approximately normally distributed heights (Figure 1a and 1b).

We intended to collect as much height data as possible. This also had the consequence that there are different types of height information. The majority of measurements are based on excavated long bones (see next section), but some information was achieved from complete skeletons that were measured lying in their original grave (*in situ* measurements). For the *in situ* measurements, we relied on the original author's judgement and adjustments (typically 2 cm is added in order to adjust for non-bone parts of the body that disappeared in the case of cadaveric length, none for *in situ* measurement as the *post mortem* stretch is compensated by missing skin see Maat [2003]).

We also used heights that were estimated from "knight's" armours of the 16th and 17th century Central and Eastern Europe. One could imagine that the armour (harness) did not fit exactly to the individuals wearing them, but were slightly larger to allow for some flexibility.[8] Fortunately, our data set contains for those centuries a sufficient number of archaeological height measurements, so that we can compare them with the armour data (that relate to 12 height numbers and 198 individuals in the 16th century and 4 numbers respectively 105 individuals in the 17th century). The simple average difference between armour height data and other height data was only about 0.3 cm (insignificant). Once we controlled for social, regional, and inter-temporal influences in a multiple regression we found that the difference was only 0.17 cm (statistically insignificant, results not shown here). We therefore decided that any adjustment for armours should be omitted because it might introduce an artificial measurement error.

HOW TO ESTIMATE HUMAN STATURE FROM LONG-BONES AND CREMATED BONES

Archaeological excavators typically do not find complete skeletons in cemeteries of early settlements, but fortunately methods have been developed to estimate human stature, even if only one long bone survived, or a part of the long bone. Anthropologists did not yet reach full agreement about the optimal estimation procedure. Naturally all the formulas only give an estimate of height. The length of a long bone represents an approximately constant proportion of height. The femur gives the best approximation to reconstruct the height of inhumated skeletons; luckily it is the most frequently survived bone. If only a part of the long bone is available, the length of the long bone has to be reconstructed first before height can be reconstructed. Because two regressions are needed in this case, the measurement error of the result is larger. Fortunately, estimating long bone length with long bone heads (e.g. *caput* *femoris*) gives a substantial R^2 of 0.6-0.8.[9] Completely preserved long-bones provide more exact results than cremated bones.

A large number of various regression models exists to reconstruct adult stature (see Wahl [1982], Heussner [1987], Hermann et al. [1990], Rösing [1988], Wurm [1985], Wurm [1986] and Wurm and Leimeister [1986])[10]: in order to find the best model anthropologists evaluated various reconstruction models for each long bone for both genders (Figure 1), which provide different results (e.g. Bach [1965]; Breitinger [1938]; Dupertius/Hadden [1951]; Manouvrier [1893]; Rösing [1977]; Trotter und Gleser [1952]; Trotter [1970]; Olivier et al. [1978]; Pearson [1899]; Wolanski [1953]).

No general consensus exists yet: anthropological analyses of the various excavation sites are based on these different reconstruction methods. Therefore it makes no sense to report the stature estimates of excavation reports directly. What could we do to make our data base homogenous? It was necessary to construct algorithms in order to standardize all height estimations to the most appropriate method. Most anthropologists and archaeologists used the following four methods to reconstruct heights of inhumated individuals: Of 2712 height estimations, 47.9 % were reconstructed with the model by Breitinger (for males) and Bach (for females), for 18.1 % the Trotter/Gleser-, for 11.3 % Pearson (respectively Wolanski)-model was used, and less often, for 5.9 % of the observations the method by Manouvrier was applied.[11] We used for cremation analysis - ca. 5.1 % observations - the model by Rösing exclusively.

We have to calculate conversion factors to get consistency of the data set. Now, which is the best regression model to use, into which the other estimations should be converted? We decided to use the reconstruction formulas after Breitinger/Bach as the 'basic' model, because it is the most commonly used model in the data we collected by now.[12] Furthermore it gives the most exact results (lowest mean of differences in stature) for the relevant spectrum of centimetres (Formicola [1993])[13]: For the height range 164 – 178.9 cm for males Formicola arrives at the lowest estimation error for Breitinger, whereas the Trotter/Gleser and Pearson methods are less efficient (see also Søjvold [1990]). In our case 92.2% of the female observations and 89.3% of the male observations fall into the range 164-178.9 cm mentioned above.

A third argument for Breitinger is that Trotter/Gleser arrive at higher, and Pearson at lower height estimates, whereas Breitinger/Bach estimates fall between them. Thus Breitinger/Bach estimates are normally close to a "compromise" estimate. That would "average" the three most widely used methods; for example the mean male height of our sample (converted measurements in comparison), would be 169.7 cm with Breitinger's method, 166.1 cm with Pearson and 170.7 with Trotter/Gleser[w].

And last but not least the Breitinger and Bach measurements are the best to compare with measurements after Rösing (for cremated bones), which cannot be converted, because the reconstruction formulas are based on different parts of bones: in the case of cremated individuals these are the diameters of long bone heads (e.g. femur F18; all measurements after Martin [1928]), but in the case of inhumated skeletons these are the length of the long bones (e.g. femur F1 'largest length').[14]

DETERMINANTS OF ADULT STATURE AND ESTIMATION OF HEIGHT TRENDS

As our height data are not perfectly equally distributed over regions and over time, we perform a number of regression analyses with time (century)-, regional, and other dummy variables in a first step. In a second step, we will apply panel data analyses on the aggregated level in order to explain the height development in the regions of Europe.[15]

We discuss the variables and their potential effects on height simultaneously in the following. Firstly, social status is an important variable, as many studies on the 18th to 20th centuries found height differences of typically 2-4 cm among adults of lower versus middle and upper class (Table 2, studies see e.g. Baten [2000]).[16] In our data set, we relied mostly on the classification schemes of the original studies. If skeletons were not of higher social rank, the excavation reports often did not find this fact worth mentioning.[17] We therefore assigned dummy variables only to the cases of middle and upper class social origin (leaving a "lower or unknown" group for the constant). This also means that we should not over-interpret the coefficient of this social status variable. On the other hand, this variable is not only important by itself, but it is also necessary to control the social composition and potential social selectivity when we analyze height trends. Although the bulk of our measurements stems from burial sites that continued all social strata, we wanted to exclude the possibility of social selectivity causing height trends as far as possible. In the aggregate analysis, we find that overall middle and upper class heights were 0.6 cm higher than the residual group (col. 2 and 3 in Table 2), this was at best marginally significant (p-value 0.11).

Another factor that we want to control is migration. Two points are important in this respect. Firstly, most migrants experienced a different environment during their first years of life that influenced their stature differently than that of the autochthon population. For example, if they were born in a Northern or Eastern European agricultural environment and than migrated to the Mediterranean in their later life, we would expect them to be significantly taller. Secondly, if immigration is large enough, agricultural production techniques might be transferred to the target region, if they turn out to be sufficiently efficient in the new environment. In addition, a number of anthropologists is still convinced that genetic height potentials play a large role, whereas other anthropologists have doubts whether genetic height potentials explain any variation in average heights (Bogin 1988, 1995).[18]

We know that the most important migration streams went from the Mediterranean region into Central and Western Europe in the first to third century, and there was an important Germanic (and other) migration in the forth to sixth centuries from Northern Europe to Eastern, Central and Southern Europe and later to the British Isles.

Migrants from the Mediterranean to Central Europe (especially Roman soldiers and officers, as well as administrative staff) turned out to be 4 cm shorter than the rest of the population (Table 2, col.4). Again, this must be regarded as a core group that did put very much emphasis on Roman customs, such as being buried with a *balsamarium* (flacon) and/or a lamp (both necessary instruments for traditional Roman burial ceremonies), because that was our only available information to identify this group. Skeletons that could be identified as "Germanic migrants" were not significantly different from Eastern Europeans. Also not statistically significant, but economically meaningful was their coefficient in the "Mediterranean" regression: 1.63 cm taller were Germanic migrants that died in the Mediterranean region.

In contrast to data on living heights, using long-bones we do not have to take into account that older individuals experience a biologically determined shrinking process: because only the body shrinks due to compression of the disks between the vertebrae (as well as poor posture). But the femur does not change significantly with age.[19] Still we include a "age 51-59"–variable, because poorly nourished people, which are also likely to be shorter, bear a higher risk to die earlier (from a nutrition-related disease). Therefore, disadvantaged groups underlie stronger selection mechanisms.[20] But the "age 51-59"–variable turned out to be insignificant.

As we want to use all available data points for the height trend estimation, we pooled male and female heights and controlled the difference with a dummy variable. This requires the assumption that the secular height trends of both genders move more or less together (we will check that below). It is interesting to note that the largest difference can be found in the least densely populated North/Eastern regions, the smallest difference in the Mediterranean region. This might be partially explained if gender dimorphism (= gender differential) increases ceteris paribus with average height, a hypothesis that needs further research (Koepke 2002).

Table 2

FOUR REGRESSIONS: DETERMINANTS OF MALE AND FEMALE HEIGHTS IN THE THREE PARTS OF EUROPE

1 Region	2 All	3	4 Central/Western Europe	5	6 Mediterranean	7	8 North./Eastern Europe	9
Constant	159.51	*0.00*	159.78	*0.00*	161.13	*0.00*	160.13	*0.00*
Status mid/high	0.58	*0.11*	0.45	*0.35*			0.32	*0.58*
Migr. Mediterr.	-3.87	*0.01*	-4.00	*0.00*				
Migr. Germanic	0.75	*0.35*			1.63	*0.24*	-0.20	*0.89*
Age 51-59	-0.14	*0.57*	-0.19	*0.53*	0.47	*0.74*	0.30	*0.51*
Male	8.20	*0.00*	7.97	*0.00*	7.72	*0.00*	8.62	*0.00*
Centuries:								
1	1.00	*0.00*	1.42	*0.03*	-0.55	*0.74*	0.59	*0.60*
2								
3	1.04	*0.00*	1.12	*0.01*	-1.20	*0.39*	1.51	*0.19*
4	0.17	*0.63*	0.20	*0.60*	0.46	*0.85*	-0.95	*0.44*
5	1.44	*0.00*	1.47	*0.01*	-0.38	*0.83*	1.76	*0.20*
6	2.80	*0.00*	2.92	*0.00*	1.30	*0.47*	2.88	*0.06*
7	1.99	*0.00*	2.24	*0.00*	0.31	*0.88*	-0.43	*0.75*
8	1.05	*0.01*	0.30	*0.51*			2.20	*0.06*
9	1.24	*0.00*	1.32	*0.02*	0.78	*0.72*	1.06	*0.35*
10	0.86	*0.17*	1.21	*0.15*			0.37	*0.78*
11	1.61	*0.00*	1.57	*0.25*			1.44	*0.19*
12	1.90	*0.00*	-0.94	*0.59*	-0.49	*0.75*	1.86	*0.09*
13	-0.14	*0.78*	0.84	*0.36*			-0.59	*0.61*
14	1.22	*0.02*	2.68	*0.01*	4.54	*0.06*	0.34	*0.77*
15	0.90	*0.32*	3.61	*0.02*	0.61	*0.75*	-2.47	*0.15*
16	1.64	*0.00*	1.57	*0.01*	6.25	*0.06*	2.61	*0.21*
17	-0.51	*0.39*	0.99	*0.23*			-1.97	*0.12*
18	1.62	*0.19*	2.09	*0.10*				
Rhine, South	0.71	*0.07*	0.48	*0.10*				
Rhine, North	1.14	*0.04*	0.62	*0.16*				
UK	-0.90	*0.08*	1.09	*0.00*				
Northern Eur.	2.21	*0.00*					1.65	*0.00*
Eastern Eur.	0.78	*0.00*						
Mediterranean	-0.01	*0.99*						
Adj. Rsq	0.54		0.51		0.63		0.61	
N (original)	9477		5303		542		3632	

P-Values in columns 3, 5, 7, 9 in italics. The constants refer to a Bavarian/Austrian (col .2/3), a not further specified Mediterranean (col. 4/5), and an Eastern European (col. 6/7). The weighted number of cases (adjusting for aggregated observations using square roots) is for the three regions 1896, 86 and 990 respectively.
As we are working with grouped data, special estimation problems could arise. See, however appendix 1, in which we demonstrate that even excluding observations with N > 1 does not change the results substantially.

Source: see Table 1

The regional dummy variable coefficients confirm our expectations. Northern Europe with its low population density and traditionally high protein production *per capita* had the tallest heights, Eastern Europe with its also lower population density but probably lower number of cows per capita (given the soil structure) ranked below that, on a similar level as the "North Rhine" region around the Netherlands, Western Germany etc. Short people were dominant in the Bavarian/Austrian, the Mediterranean and the British regions, the latter especially in the Celto-Roman period.[21] The time dummies allow the description of the secular height trend, after controlling for regional, social, age- and migration related composition.

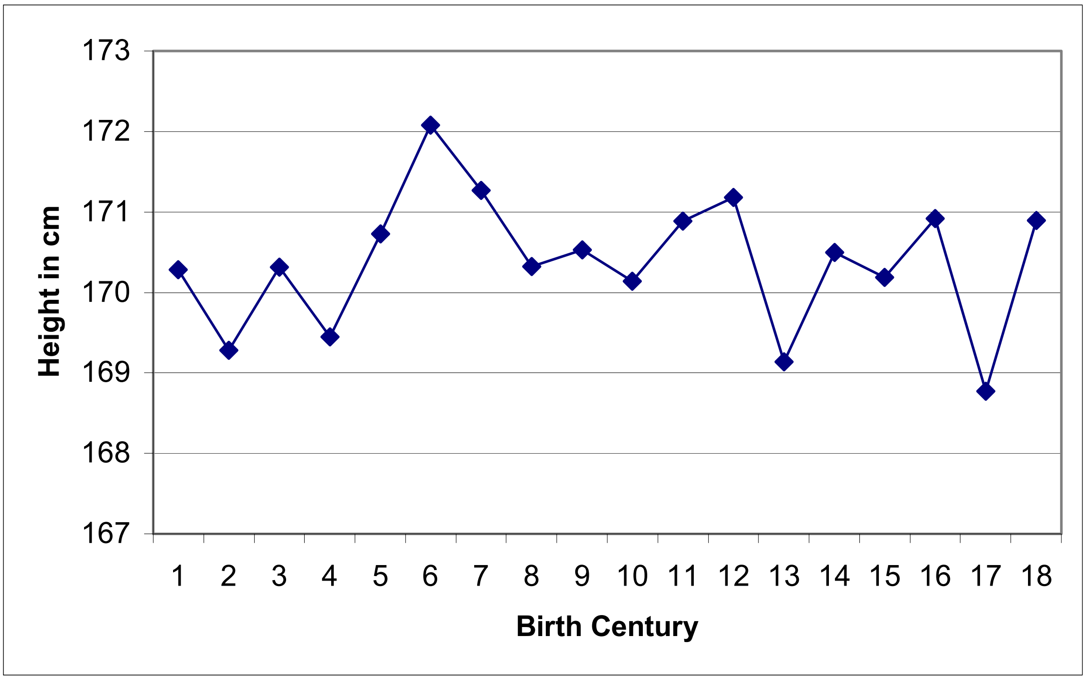

Figure 2

HEIGHT DEVELOPMENT, 1st TO 18th CENTURIES

(IN CM, MALE AND FEMALE)

Source: see Table 1. The level of heights was adjusted to male heights of an average European (using the regional coefficients and weighting them with sample weights).

Which overall height trend evolves when we consider the century dummies of column 2 of the previous regression table? During the time period of the Roman Empire heights did not increase (Figure 2). This stands in contrast to the common view that suggests an increase in living standards under the protection of *pax Romana* and especially in purchasing power due to the growth during the Roman Empire. It is also remarkable that heights increased in the fifth and sixth centuries, even after the breakdown of the Empire in the West.

But with increasing population and adverse climate during the following period, adult stature declined again until the 10th century AD. A favourable height level characterized the medieval warm period of the 11th and 12th. During the 13th century, heights collapsed, whereas the 14th to 16th centuries might have had good nutritional conditions. The 17th century was a period of nutritional crisis again. The estimates on the 16th to 18th centuries are less reliable as the number of cases is small.

If we compare our skeletal measurements with height estimates based on military archival sources, we find that skeleton male heights during the 18th century were in the same range, but on average slightly higher than military heights: Our values of 168 – 170 cm (17th/18th c.) correspond with the 164 and 172 cm estimates of the military samples. 17th century military heights were even below that. Thus this increase between the 17th and 18th century is supported by our skeleton height trend (see Komlos 1989, 2003).

How can we further assess whether this series reflects the true height development? One strategy of counter-checking is to look at disaggregated data by region and gender. If the disaggregated series move in a similar way when we would expect them to correspond, and only deviate where it makes sense from a theoretical perspective, then this would support the validity of the overall height trend. The development in the regions Mediterranean, Central/Western Europe and North/Eastern Europe is in general quite similar (Figure 3): the decline in the fourth century in Central Western and North-Eastern Europe, the astonishing increase in the fifth and sixth centuries and the low points in the 13th and 17th centuries. The increase in the 14th century and the high value during the 15th century can also be observed in more than one region. Deviations appear in the seventh and eighth centuries. The North/Eastern height series remains constantly above the other ones, but from the 13th century onwards it loses its leading role. In the 13th, 14th and 17th centuries Scandinavians and Eastern Europeans became shorter than British, Dutch and other Central/Western Europeans![22] It might have been that Northern and Eastern populations suffered particularly during the Little Ice Age (14th – 18th c.) whereas they benefited from the climatic maximum of the 11th and 12th century. In contrast the maritime climate in the Netherlands, the British Isles and Western Germany allowed a relatively more favourable nutrition during this Ice Age period.

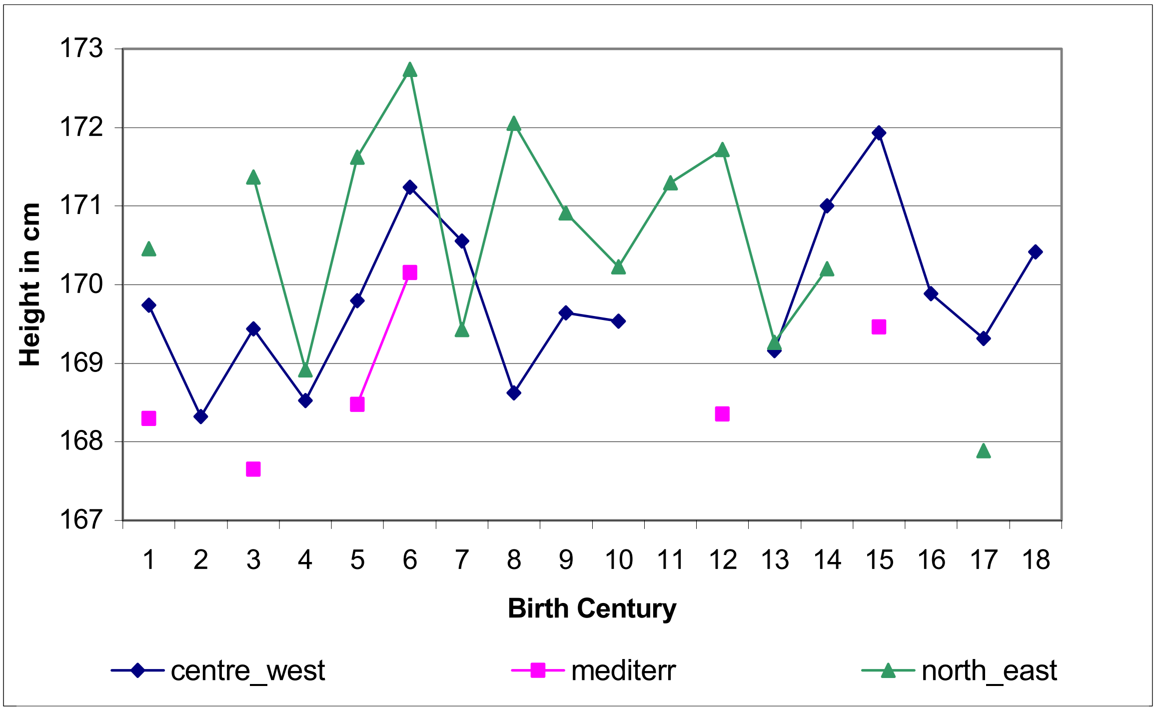

Figure 3

HEIGHT DEVELOPMENT BY MAJOR REGIONS (IN CM)

Source: see Table 1

In a similar vein, male and female heights moved in a more or less similar way (Figure 4). The common maximum during the sixth century might have even been caused predominantly by the female height series but both series reached a maximum at this point, and a further high value during the 11th and 12th centuries. During the high Middle Ages females lost some ground compared with males. The increase of female heights during the 15th century is supported by only 18 observations but the positive situation of the 16th century relies on 118 cases. Women might have benefited from the Renaissance period, whereas gender discrimination was particularly severe during the "Dark Ages": in the course of the Medieval Ages women lost social position as archaeologists demonstrated (see Ulrich-Bochsler [1996]).

Do female heights have a higher variance in our sample? The overall coefficient of variation is 0.68 for females and 0.49 for males. Biologists would expect that female bodies are more robust under adverse conditions (see e.g. Schweich/Knüsel [2002]; Ortner [1998]), so the variability should be lower. But our results indicate that gender discrimination might have been stronger than biological factors. This fits to the argument that the position of women deteriorates in relative terms when times are getting worse (Klasen [1996]). We have to admit though that the higher variability could also be influenced by the lower number of female observations.

How can we be sure that that there is no bias due to various burial customs? In general we considered excavation reports that included the whole population and not just some noble men's graves. We surveyed the archaeological literature and found no hints on burial customs that could have biased our results significantly: it seems to be not the case that rich and poor graves are exposed to different preservation conditions on average. Another strategy to test this important aspect is to compare different regions, because in this different regions one might expect that burial customs were not perfectly correlated: as we found similar trends for the different regions – except for the plausible decline in heights in North-Eastern Europe (Little Ice Age) – we conclude that time trends were not caused by local burial customs. We also looked at single sites which were occupied over more than one century, but with homogenous culture and burial custom; we found that the observed trend is confirmed in most cases.

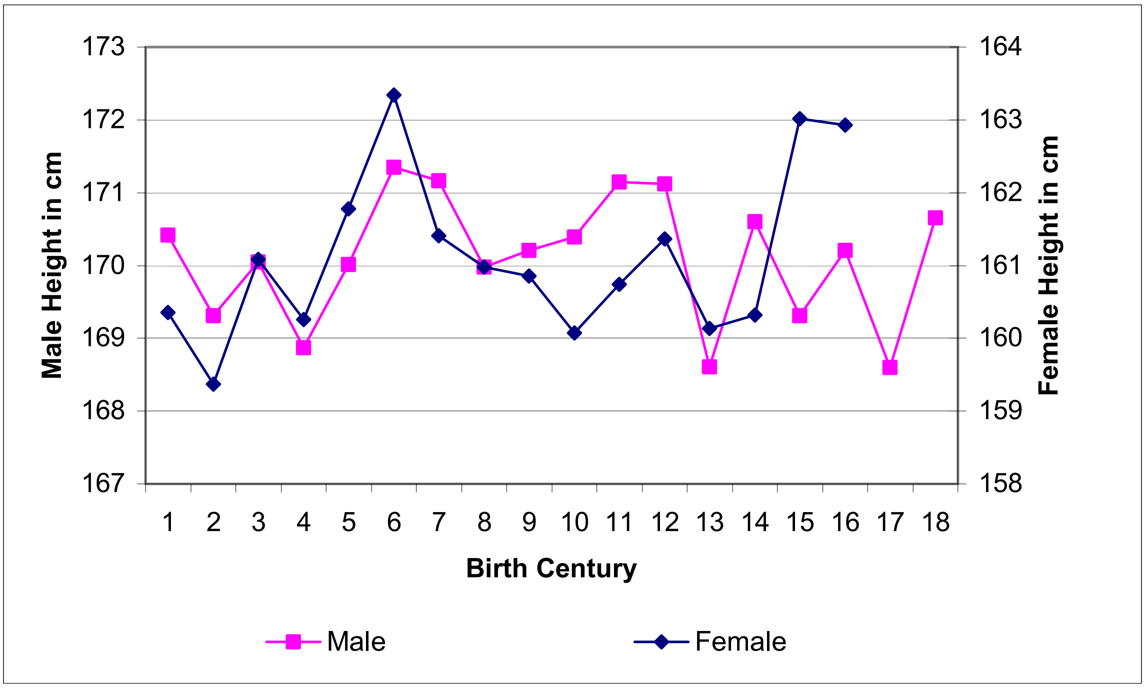

Figure 4

HEIGHT DEVELOPMENT BY GENDER, 1st TO 18th CENTURIES (IN CM)

Source: see Table 1

POTENTIAL ECONOMIC DETERMINANTS OF HEIGHTS IN THE LONG RUN

In the following we assess the potential influence of a number of variables on stature by century and region:

(1) Land *per capita* and urbanization

(2) Climate

(3) Income

(4) Social Inequality

(5) Public Health

(6) Gender Inequality

(1) Land per capita and urbanization

Was Malthus right with his argument that land is the limiting factor of human development? Did population growth tend to outpace food production because of the decreasing marginal product, until a major demographic catastrophe took place? Before mortality crises took place, Malthus would expect a decline in the nutritional status of populations under risk, and this relationship is tested in the following. On the contrary, after major mortality crises nutritional quality might be quite good, because more land was available. Cows and other farm capital were not reduced as much as population during, for example, the major plague epidemics of the 6th, 14th and 17th centuries.

Apart from the effect of land availability, increasing urbanization itself might have separated urban dwellers from de facto untradeable goods such as milk. And infectious disease spreads more easily in urban centres.

We measure population density as a weighted average of the country estimates of McEvedy and Jones (1978), wherever possible improved with newer and more detailed country studies (such as Wrigley/Schofield (1981); Dupâquier et al.1988). For further details, see the appendix under www.uni-tuebingen.de/uni/wwl/twomillennia.html. We estimate urbanization levels as described in the appendix (based on Bairoch (1976); Federico/ Malanima (2002)[23]).

(2) Climate

One of the fascinating topics of long-run economic history is the relationship between climate and human living standards. For example, did climatic change cause the demographic catastrophes of the 14th and 17th centuries, as Galloway (1986) argued (see e.g. Kelly [n.y.])? Colder winters tend to make food production (especially protein production) more difficult in Central Europe (on the 18th century climate-height effect see Baten [2002]). The impact on human history was immense. Grove (2002) demonstrated how the switch from the medieval warm period (900 early 13th century) to the little Ice Age starting in the late 13th century has decreased harvests and protein production from cattle and sheep.[24] Not only were temperatures declining, but, as colder winters tended to be generally correlated with more frequent weather extremes, also other climatic problems created a deadly synergy. For example, cattle epidemics spread rapidly in Northern and Western Europe already in the 13th and early 14th century, killing a large part of the cattle stock. Grove therefore argued that the agricultural production decline took place before and parallel to the Black Death of the mid-14th century.

Although plague is a highly infectious disease that is only mildly influenced by malnutrition, the lower nutritional status might have weakened the immune system of the European population, contributing strongly to the large population loss of the 14th century. In addition, during famines people often leave their households and start to move around in search of other possibilities of subsistence (Mokyr/O'Grada [2002]). The most Northern, cattle- and fishery-based economies of Europe suffered most. Iceland lost most of its population, and the European population of Greenland completely disappeared.

The 15th and the first two thirds of the 16th century were warmer again, but the 17th century represents the next climatic catastrophe. Pfister (1988) has described how climate reduced Swiss nutritional status in the last decades of the 16th and most of the 17th century. While traditionally most of the population decline of the 17th century is traditionally attributed to the Thirty Years' war as well as to the hunger and the infectious disease that accompanied it, the rapid climatic deterioration could have contributed to the large number of deaths from (at least partially) nutrition-related diseases during this devastating war. The synergy between protein malnutrition and death from a large array of diseases can also explain why the population stagnated or declined even in countries that did not directly participate in the Thirty Years' War. Milder episodes of climatic deterioration in the late 18th and mid-19th centuries coincided with milder demographic effects on average, even if some regions and countries were severely affected (Grove 2002).

Recent research has created new estimates from Alpine and Scandinavian glacier movements, from Greenland ice kernels, from oak tree rings and lake sediments to quantify climatic change over the centuries. All of those series appear to be correlated in general.

We used mainly the glacier movements as explanatory variables, because they are available for the ancient time period and the evidence might be less indirect compared with, for example, Greenland oxygen isotope ratios (see Heide [1997]; Grove [2002], p. 316). However, the literature emphasizes that glacier movements reflect temperature changes with a certain time lag. We therefore calculated the average of the previous and the current century glacier movement. We corroborated our glacier series with a tree-ring series from North Sweden that also stretches back to the ancient period (and compared both with a shorter tree-ring series on the Alpine area: they moved in accordance, see Huntley et al. [2002], p. 278.).[25]

There are some similarities and many differences between the height and the temperature series (Figure 5). The well-documented climatic optimum of the 11th/12th centuries is visible in the height series, and lower values before and after. The low values of the seventh and eighth centuries and the crisis of the 17th century could have been caused by adverse climatic conditions. Important deviations relate to the first to sixth and to the 13th centuries. Either the relation is truly weak, or there is measurement error, especially for the early period, for which the temperature estimates are known to be particularly imprecise. The most likely interpretation in our view is that after the breakdown of the Roman Empire several interesting phenomena increased average height and nutritional status: (1) population density and urbanization decreased after invasions and plague epidemics. The consumers moved back to the proximity of nutrient production. Infectious disease might have appeared less frequently (although the latter factor might be of small importance, because the famous Roman Public Health institutions disappeared and the first occurrence of the plague in the sixth century contradicts it). (2) Germanic invaders brought their agricultural methods that emphasized protein production. Even if they were inefficient in the Mediterranean, they might have kept them for a transitory period. In Central and Western Europe, they were efficient as long as population density was low.

These two developments might explain why the climate-height relationship is not visible for the first six centuries. The low height value of the 13th century is particularly interesting and deserves further study. Was it the rapid urbanization of this period (more infectious disease, less milk for rural-to-urban migrants)? More social or gender inequality? Exists a measurement error in the height variable?

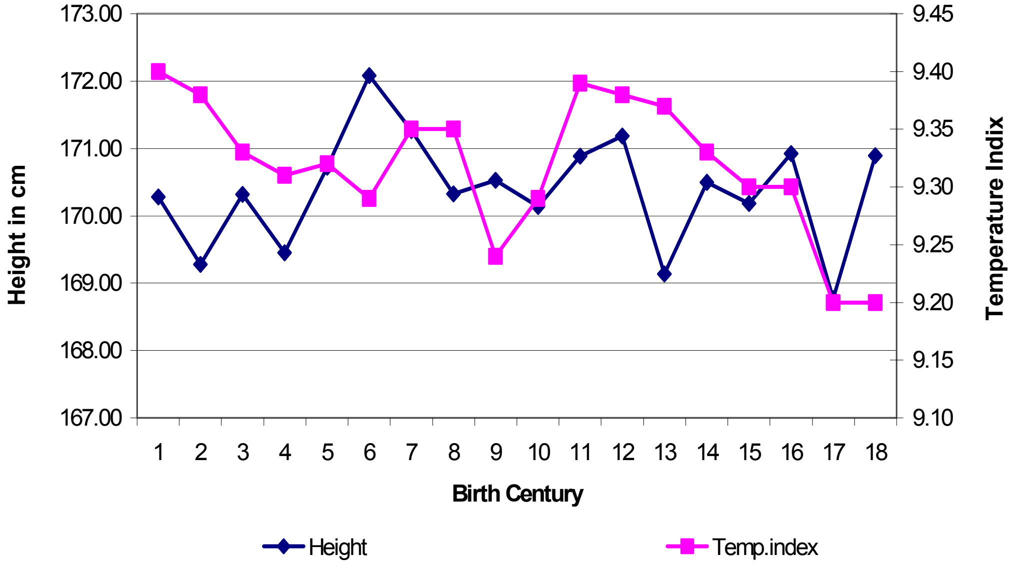

Figure 5

HEIGHT AND TEMPERATURE DEVELOPMENT (1st TO 19th C), BASED ON GLACIER MOVEMENTS AND TREERINGS

Source: see Table 1

(3) Income

Agricultural income was to a large extent a function of land per capita and climate, but the level of income might have been augmented by industrial and service sector production particularly during Roman times, the High Middle Ages, the 16th century and after the 18th century. However, we cannot test this effect, because income estimates for the first millennium are especially unreliable or are only based on urbanization rates and population growth (such as Maddison's [2001] estimates).

(4) Social Inequality

Inequality was identified by previous research as an important determinant of average height (Steckel [1995]). Growing income inequality of purchasing power without changes in aggregate real GDP per capita might make the rich richer and the poor poorer to the same extent, but as the rich will spend less on additional food, and the poor will lose decisive nutrients at their low level, average height will decline even if average purchasing power does not.

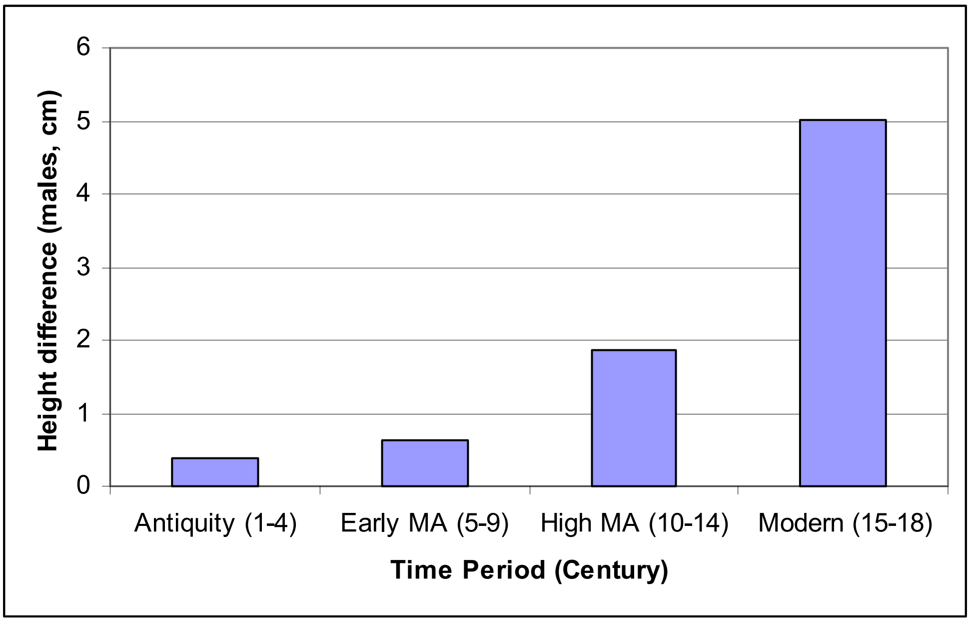

Figure 6

DEVELOPMENT OF INEQUALITY

Source: see Table 1

Height data sets allow rough estimates of health inequality (see Baten [2000], Pradhan/Sahn/Younger [2002]). Wurm (1985) argued that inequality was particularly low in the early Middle Ages. In fact, social inequality increased dramatically between the early and the high Middle Ages (Figure 6), and again in the 15th – 18th centuries. The overall trend towards inequality corresponds well with studies on income inequality. O'Rourke/Williamson (2002) have confirmed this for longer time periods using rent-wage-ratios which assumed reasonably that land-owners were relatively rich and wage earners were relatively poor (van Zanden [1995]). We controlled for social inequality by calculating the difference between the middle/higher status individuals and the lower/unknown category by major time period.

(5) Public Health

How did Public Health develop over the last two millennia? Is our image of the impressive Roman water supply technology and especially the Roman institution of bathing facilities correct and did a large part of the population benefit from this? [26] To what extent did the hygienic conditions decline after the breakdown of the Roman bath-system (see Grewe [1986]; Hermann [1985])? Were Public Health investments perhaps endogenous, and took place when urbanized and poorly nourished populations suffered more often than usual from infectious diseases? Or did the contacts of Romans with distant populations (such as Chinese, Indians, Parthians) spread new infectious diseases? In order to capture the potentially beneficial of effects of the Roman Public Health system, we coded a "Roman Bath" dummy variable as 1 for the Mediterranean for the centuries 1 to 4, and for Central/Western Europe for the centuries 2 to 4.[27] This specification means that apart from the Roman hygienic system (that might be very important for height and health) we might capture also other aspects of Roman technology and the imperial economic system. We will discuss this variable therefore as "Roman bath or other technology".

(6) Gender Inequality

We assumed in our estimation of height by (birth) century that gender differentials were constant over time. In the following we will relax this assumption and control explicitly for higher or lower gender inequality. Our expectation is that higher gender inequality *ceteris paribus* reduces average height, because Osmani and Sen (2003) have convincingly argued that female discrimination hurts both girl's and boy's height via the low nutritional status of their mother (see also Klasen [2002]). We measure the development of gender differentials over time with the dimorphism estimates that we had graphed in Figure 4. However, we calculate the percentage height difference relative to the average male height in order to adjust for possible level effects (Koepke 2002).

RESULTS

We estimated a large number of regression models: fixed effects and random effects, weighted and unweighted least square, generalized least square with adjustment for autocorrelation and heteroscedasticity (not shown in tables). The following results are robust across these various specifications. Unit root tests or cointegration might not have enough power with 18 observations over time, but if one nonetheless performs the ADF test, it turns out that the height series for North/Eastern and Central/Western Europe are stationary: ADF North/East: even nesting the trend (which is insignificant) -5.12 (critical value 10%: -3.24); ADF Central-West: -3.43 (critical value 10%: -3.24) without nested trend a unit root is rejected at the 5 % level; Mediterranean: insufficient number of cases. Panel estimation techniques exploit the variation both over time and between cross-sectional units. We present two WLS estimates with regional dummies (equivalent to fixed effects), one of them with time (period) dummies and the other without controls for unobserved inter-temporal heterogeneity (Table 3a and b). The results are also robust if time trends are included (insignificant).

Given that we have a lot of measurement error in our rough proxies for social and gender inequality, population density and urbanization, it is not astonishing that many coefficients are statistically insignificant. In addition, technological development cannot be appropriately captured over these two millennia: A certain population density in 1800 might have resulted in a different height compared with 800 (simple time trends were insignificant though). Following McCloskey and Ziliak (1996) we will not only focus on statistical, but also on economic significance, comparing high and low predicted height values of our estimates with results from the 18th and 19th centuries that often interpret differences between 1 and 3 cm as economically significant phenomena (Komlos and Baten [1998] estimated that 1 cm additional height means about 1.2 years of life expectancy).

Significant are the regional dummy variables and the time period dummy that represents antiquity. In the regression without time dummies the "Roman bath"-dummy gets significant. However, Roman bath technology and especially other technology (e.g. in agricultural terms) was not able to improve height and health quality sufficiently to outweigh the negative effects of the Roman economic system: the coefficient of this variable is actually negative. We coded it as zero for North-Eastern Europe and Central-Western Europe in the first century, and this variable remains significant if we control the antiquity effect of low heights with another dummy variable (not shown). The other variables are statistically insignificant, but most bear the expected sign: warmer weather is good for harvests and protein production in the relevant range, and this is favourable for height. The difference between two standard deviations of our climatic series is 0.12, the difference between minimum and maximum is 0.20; this can be interpreted as typical "good" and "bad" climate. The coefficient of the more appropriately specified model in Table 3b, column 2 is 2.97. The difference between "good" and "bad" climate was therefore about 0.4 cm, the difference between the extremes about 0.6 cm. Both values are at the margin of being economically significant. Without controls for time period, this variable is economically unimportant. The tall stature of North-Eastern Europeans in the warm 11th/12th century and their dramatic decline lends further support for this variable.

Table 3a

DESCRIPTIVE STATISTICS

1	2 N	3 Minimum	4 Maximum	5 Mean	6 Std. Deviation
Climate warm	36	9.20	9.40	9.32	0.06
Gender inequality	36	4.14	6.11	5.37	0.5
Urban share	36	1.00	19.30	4.68	4.64
Social inequality	36	0.37	5.02	1.6	1.66
Population density	36	3.75	40.19	11.38	8.33
Valid N (list wise)	36				

Table 3b

TWO REGRESSIONS: DETERMINANTS OF HEIGHT

1	2	3	4	5
Constant	144.24	*0.00*	164.37	*0.00*
Climate warm	2.97	*0.52*	0.82	*0.84*
Gender inequality	- 0.31	*0.50*	- 0.29	*0.46*
Urban share	0.16	*0.23*	0.2	*0.14*
Population density	- 0.06	*0.37*	- 0.08	*0.20*
Roman Bath/ Roman Technology			- 2.05	*0.01*
Social inequality			- 0.17	*0.58*
Mediterranean	- 1.66	*0.05*	- 1.67	*0.04*
North-Eastern Europe	1.17	*0.03*	0.89	*0.07*
Antiquity	- 1.68	*0.01*		
Late Medieval Period	- 0.48	*0.52*		
Modern (15th to 18th c.)	- 0.76	*0.59*		
Adj. Rsq	0.33		0.38	
N	36		36	

P-Values in columns 3, 5 in italics.
Weighted Least Squares Regression: number of cases adjusted for aggregated observations using square roots.
Constant refers to a hypothetical height value for the Early Middle Ages, and Central/Western Europe.

Source: see Table 1 and text.

Population density comes closest to statistical significance, in unweighted regressions the p-value is even as low as 0.15 (not shown in table). It suggests that lower population density is good for the biological component of the standard of living that is reflected in stature in pre-industrial times, after controlling for large-region effects and inequality. The analysis of economic significance for population density yields a height effect of about 1.0 cm for the typical "high" and "low" population density of the time and 2.2 cm between the most extreme observed values. In the other specification, the economic significance of population density would even be one third greater. Interestingly, the sign of the urbanization coefficient is positive once population density is controlled for.[29] Without time dummies, it is even almost significant. The potentially large measurement errors prevent us from hypothesizing at this early stage. But one could speculate that once the detrimental influence of high population density (that means because of the decreasing marginal product: less protein *per capita*) is removed, the human capital deepening effects of urban agglomerations on the whole country overwhelm other negative effects (such as urban crowding and hygienic problems).

Gender inequality and social inequality both had negative signs.[30] Given that these results are similar to those of many other studies on the 18th to 20th centuries, we tend to attribute a fairly large credibility to them. In terms of economic significance, social inequality meant 0.63 cm between high and low and 0.74 between extremes, whereas the effect of gender inequality was about half of that. In sum population density is definitely of economic significance, but not of statistical significance. Climate and social inequality and perhaps gender inequality are at the margin of being economically significant.

CONCLUSION: THE LARGER PICTURE, SOME SPECULATIONS AND PLANS

In sum, this paper offers the first anthropometric estimates of height time series in Europe over the last two millennia (excluding the last two centuries on which much research has been done already). Height series are often related to other biological aspects of living standard, but they do not necessarily capture other important aspects that are related to purchasing power: A Northern Barbarian of the sixth century was tall and certainly lived relatively long (as Hermann [1987] demonstrated), but if he was a young entertainment and industrial-goods loving person he might still have preferred to live in Rome of the 2nd century. We acknowledge that, but we cannot measure these aspects of welfare; but we capture other aspects that are important and were often underestimated.

The overall picture is one of stagnant heights. There was not much progress in European nutritional status, not even between 1000 and 1800, when Maddison (2001) and others arrive at growing GDP *per capita* figures (Figure 7, but see Federico [2002] for a critical view.).[31]

Likely reasons for this divergence are (a) the relatively favourable nutrition during the Middle Ages, especially the climatic optimum of the 11th/12th century. (b) The bias in pre-industrial and early GDP *per capita* estimation in favour of industrial goods consumed by middle and upper class consumers.[32] This latter possibility is also supported by van Zanden's (1999) finding of a “negative link between economic development and the level of real wages” from the beginning of the 16th to the end of the 18th century. He also described a decline of consumption of meat and dairy products *per capita*. In a similar vein, Federico and Malanima (2002) estimated a downward trend of food consumption in Italy between 1300 and 1860.

Is there an impact of this “diverging” on the “Great Divergence” debate? The GDP – Height “divergence” does not necessarily mean that we support the ”Great Divergence” view of Pomeranz (2000) and the California group. They argued that Europe did not better than China and perhaps India until the 18th century. We do not yet know the height trends for Asia.

There might have been a strong and long-run height decline in China and India to the very low height levels of the 19th and early 20th century (whereas today's young Beijing adults are only about 1 cm shorter than U.S. adults, and early medieval Asians were astonishing tall).[33]

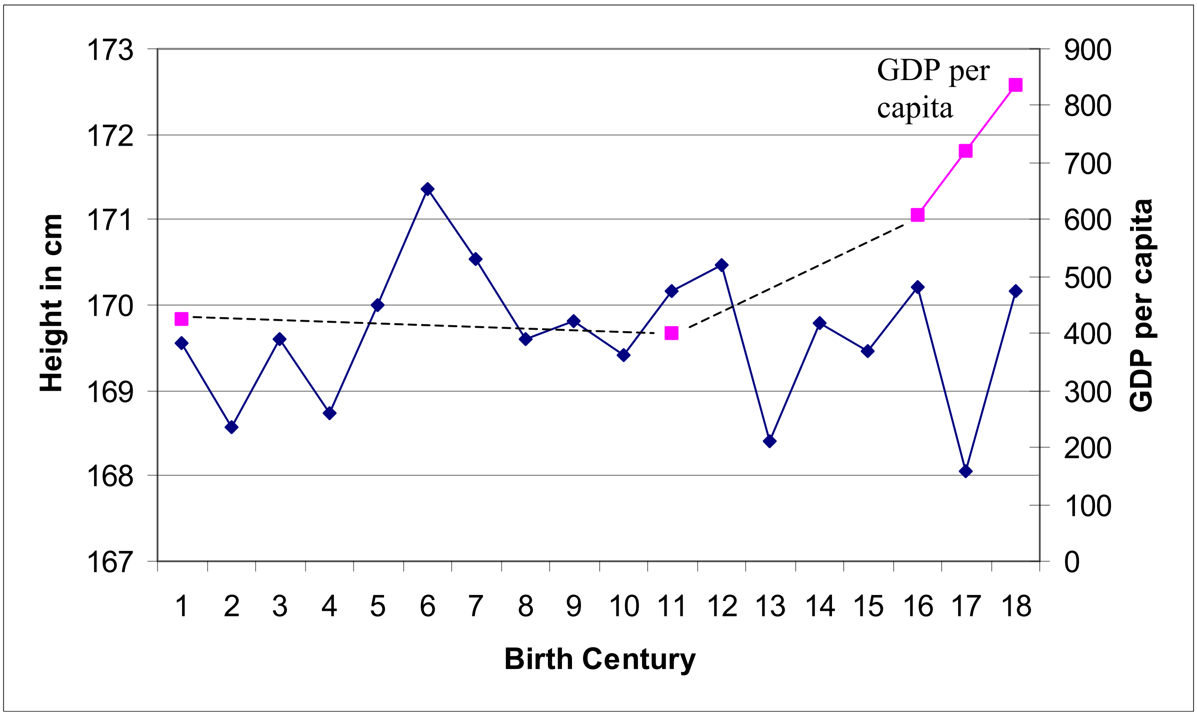

Figure 7

HEIGHT DEVELOPMENT AND GDP PER CAPITA

Source: see Maddison (2001) and Table 1. Countries excluded for which there is no height data

Our analysis stretches back to ancient times, measuring living standards during the Roman Empire and during the following “Dark Ages”. We find that heights stagnated during the Roman imperial period in Central, Western and Southern Europe. In Northern and Eastern Europe, heights might have increased between the first and third century, but fell dramatically in the fourth century. Whether this contributed to the start of the migration of peoples since the fourth century awaits further exploration.[34] One astonishing result is the height increase in the fifth and sixth centuries. This is the largest residual in our explanatory model at the moment. Declining population density in the former provinces because of the breakdown of the *imperium Romanum* and the plague of the sixth century might have contributed.[35] Noteworthy is the synchronicity of the height development in the three regions.

We constructed potential explanatory variables on the narrow basis of what we know about this early time period. Population density was clearly economically (but not statistically) significant. Decreasing marginal product theories and Malthusian thought cannot be denied for the pre-1800 period. Of marginal significance were climate (warmer temperatures was good for nutritional status), social inequality and gender inequality (both reduce average height). Controlling for population density, urbanization was positive for the whole population on average.

Questions about the social composition of the height samples over time and other potential biases can only be answered within certain confidence intervals (given the current state of research). However, we would argue that the error probability is smaller than for most other methods that can be applied to the first millennium (such as the urbanization-based GDP estimates by Maddison [2001]). If we want to study the economic history of the very long run, anthropometric techniques provide important insights on some (not all) of the central aspects of human life.

APPENDIX 1:

ESTIMATES WITH AND WITHOUT GROUP MEASUREMENTS

Our sample consists of 2774 individual heights and 200 aggregated height figures that refer to the remaining 6703 cases. Such a combination of grouped and individual data is a complicated econometric issue, because the better documented cases deserve a higher weight in the regressions, but as their distribution is unknown, the calculation of usual t-statistics requires additional assumptions. In addition, the weight should not be arbitrary, it should be greater than 1 for the reasons mentioned, but it should also be less than the number of individuals over which the average was calculated.

In a first step, we basically used the square root of the number of cases as importance weight in WLS regressions, so that the aggregated cases get a slightly smaller influence on the results than the individual measurements.

The basic econometric issue is this: As we have N people with a mean height of H and put in H as the obs and sqrt(H) as the weight, then at the very least, this observation is stochastic while the others are not (the standard errors can be both, too large and too small). So our standard errors are all not appropriately estimated because of this fact. There is also an efficiency issue, because we do not really have N observations, we just have the mean (this issue means that the standard errors might be too large). There are strong variance and covariance assumptions needed to make regression with grouped data equal to regression with individual data.

We do not directly investigate those assumptions. Since we have a large data set for most centuries, we instead simply drop the grouped (aggregated) observations and compare the estimates of those individual-only heights and the full data set (Figure A.1). The results are very robust, but the 14th to 16th centuries are poorly documented without the grouped observations. In all three centuries, there are less than 20 observations (if we would show those estimates based on N=13, N=6, and N=1, the estimates would be 1-2 cm below the estimates that included the grouped observations). For all other centuries, the height estimates based on all cases with Square Root Weighted Least Squares and with individual data only are highly correlated. If the estimates are performed by the three European regions, the correspondence is similar (except that there are more poorly documented centuries in which the deviations are naturally larger).

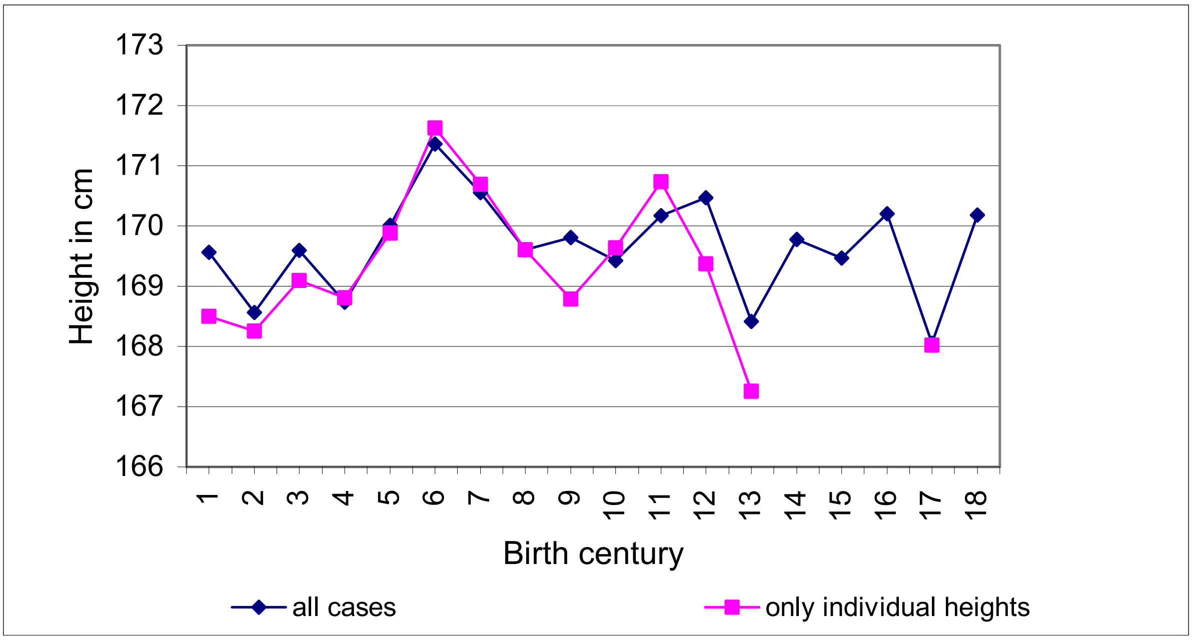

Figure A.1

HEIGHT DEVELOPMENT, 1st TO 18th CENTURIES: ALL CASES AND AGGREGATED HEIGHTS EXCLUDED (IN CM, MALE AND FEMALE)

Source: see Table 1. The level of heights was adjusted to male heights of an average European (using the regional coefficients and weighting them with sample weights).

Additional measurement issues

We excluded among the grouped data all height estimates that were not exclusively based on one gender.

FOOTNOTES:

---

[1] Komlos (1985) was the first to use this term.

[2] Growth that is mostly stimulated from trade and comparative advantage, as opposed to Schumpeterian growth that is driven by technological change, see Mokyr (1992).

[3] Richard Steckel and his co-operators have started a research project (Steckel, R. [2003]) in which they will study height and a number of diseases (as far as they can be traced with bone material). A boom in this field of research can be expected in the future.

[4] Koepke (2002) discussed intensively the estimation of height from cremated bones for the period of ancient Rome.

[5] The so-called primary deficit (smaller number of females in the case of patriarchal structured societies) of females is typical for prehistoric and ancient populations: see e.g.: Barber and Bowsher (2002); Mays (1995); Karpf and Karpf (1973); Volk et al. (1988).

[6] The same applies to age estimates.

[8] Note that most knight's armours were actually from a time period when the military technology had moved away from the horse-based knight armies that were so unsuccessful in the Hundred-Years' War. Our armour (see Wurm [1985b]) probably stem of protected males from all social strata that were hired and salaried soldiers.

[9] We thank Prof. Dr. F. Rösing, University of Ulm, for his kind verbal information.

[10] Naturally different regression formulas have to be used for females and males. Furthermore the high number of models is not only the result of different reconstruction models for inhumated and cremated bones - as the latter show diminution (because they were exposed to the heat) - ; only for the reconstruction of inhumations 43 models exist. On cremation at least 22 researchers worked on different models.

[11] Stature reconstructions by the also quite commonly used model by Olivier were sorted out as these are not based on the same Martin (1928) -measurements of the bones like the above mentioned ones.

[12] If we know that the excavators and principal investigators used only femurs, we can recalculate each stature estimate into the corresponding estimate of another reconstruction method. However

---

sometimes the authors do not mention if they estimated only with femurs or, for example, with a combination of bones (which would make the estimates more exact), like femur and tibia, or only with another bone, like tibia. In order to minimize recalculation error, we prefer to use the calculation method that most principal investigators employed.

[13] But note, that, if the height of especially small individuals (males below 160 cm) is reconstructed Pearson might be more efficient.

[14] It is the distance from the highest point of the caput to the lowest point of the *condylus tibialis*. According to J. Wahl the stature estimates with the Breitinger/Bach method comes nearest to the Rösing estimates. Because we want to compare Rösing values, it is best to use Breitinger and Bach measures. – We thank PD Dr. J. Wahl, University of Tübingen and Landesdenkmalamt (Bureau of ancient heritage and monuments) Baden-Württemberg for his kind verbal information.

[15] In order to test the robustness of our first step results, we ran the regressions with aggregate and disaggregated data (disaggregated by region and gender; gender disaggregated regression not reported here).

[16] The latter two are usually taken together because the share of upper class is very low, and the even the share of middle class is not very high in most historical populations.

[17] In some cases, grave goods that indicated higher social status were certainly lost or robbed. This means that middle and upper class status is probably somewhat underreported.

[18] To be sure, this refers to aggregated height. Individual height is clearly influenced by genetic factors. It is not fully clear whether some very isolated populations such as the Pygmies have a different height potential. Also the Japanese are sometimes outliers in regressions (although their strongest height increase after World War II correlates well with the introduction of dairy products). Maya children that were brought to the U.S. and enjoyed good nutrition converged rapidly to North American growth paths – but not fully (Bogin [1991]).

Earlier views that North-Eastern French are genetically taller were recently rejected: once milk production and income are controlled, the height difference disappears (Baten [1999b]). The finding that the Dutch were particularly short in the nutritional crisis of the mid-19th century and that the

---

Indians (of Asian origin) and central Asian nomads were particularly tall also speaks against the explanatory power of genetic factors (Steckel [2001]).

[19] There is some compression of joints in the lower half of the body as well, but the length of the femur does not change enough to make a difference: Almost all of the age related loss in height derives from the collapse of the intervertebral disks and, in some individuals, the collapse of the vertebral bodies in some individuals. Changes in femoral length terminate with the fusion of the epiphyses. The only way you could get changes in length after that would be through remodelling of the articular surfaces or bending of the bone itself. Both of these changes would only be seen in rare pathological conditions such as very severe osteoarthritis, femoral fractures, and perhaps osteomalacia. Thanks to friendly communication with Barry Bogin, Richard Steckel and Phil Walker.

[20] If this mechanism would be at work, we would expect relatively taller old people among more disadvantaged groups and shorter old people among the less disadvantaged groups. In fact, we find that the height at old ages was lower for example among males than among (more discriminated) females (not shown in the table).

[21] The relatively high R-squares should be regarded with caution; it stems mostly form the intension of the gender dummies.

[22] Allen (2000) finds a positive real wage development in the early modern period for the British urban case. The number of observations is (quite) comparable summarized by large region (see Table 1: 13th (c.: 82 to 195 cases), 14th (c.: 553 to 120 cases), 17th (c.: 80 to 58 cases)).

[23] We did not use the superior estimates of de Vries (1984) because they only start after 1500.

[24] Grain yields were falling between 1220 and 1320, see Grove (2002), figure 2.

[25] We experimented with local temperature series for the three regions North/Eastern, Central/Western and Southern Europe, but the differences between the series were extremely small, so that we abandoned this avenue of temperature measurement.

[26] In contrast to the general positive view of Roman bath Scobie (1986) argued that they were quite unhygienic (e.g. water was rarely changed; it was a meeting point of ill and healthy people).

[27] McKeown (1955) has argued that medical technology did not play an important role in societies before the 20th century.

---

[29] For the case of the Roman provinces some authors hypothesized that urban population was better supplied with food than their rural counterparts.

[30] The statistical insignificance is probably not caused by the fact that we controlled for social status in the time-generating regressions of table 2. We performed the same estimations without social status and all results were almost identical.

[31] In our data Scandinavia and Eastern Europe heights even declined significantly. This supports Steckels (forthcoming) finding of a long-term height decline in Northern Europe. We took care not to use the data again that he analyzed, except for very few heights we had collected at an early stage of our project (6 %).

[32] Note also the potential urban bias of real wage estimation (rural nutrition and living standard in contrast was crucially influenced by non-traded, high quality proteins (such as milk), see Baten [1999], van Zanden [1995]).

[33] Personal communication with Phillip Walker, Univ. of California at Santa Barbara.

[34] Before the fourth century, the direction of migration was the opposite, from the extremely densely populated Italy to the imperial provinces.

[35] We also used plague dummies for the sixth, 14th, and 17th centuries and found them insignificant. In addition, we tested whether stature was higher in the second halves of those centuries, because the most violent plague waves occurred around mid-century, and the tall stature of the sixth and 14th centuries might have been caused by the lower population density afterwards. In fact, heights were half a centimeter taller in the second halves of those centuries, which is economically, but no statistically significant (not shown here).

ACKNOWLEDGEMENTS

We would like to thank Richard Steckel, Timothy Guinnane, Joachim Wahl, Alexander Moradi, Gerhard Kling, as well as Stephen Broadberry, Sevket Pamuk and Esteban Nicolini and other participants of the fifth European Historical Economics Society Conference at the

---

Universidad Carlos III, Madrid, Spain in July 2003, and other members of the Tuebingen Economic History Research Group, for their interesting comments on earlier versions. Jan Luiten van Zanden gave important hints to new data.

REFERENCES

Allen, R. (2000). Economic structure and agricultural productivity in Europe. *European Review of Economic history* 4.1, 1-26.

Armelagos, G. (1990). Health and Disease in Prehistoric Populations in Transition. In: A. Swedlund and G. Armelagos (eds.), *Disease in Populations in Transition.* New York: Bergin and Garvey. 127-144.

Bairoch, P. (1976). Population urbaine et taille des villes en Europe de 1600 á 1070. *Revue d'histoire économique et sociale* 53, 304-335.

Barber, B., and Bowsher, D. (2000). *The Eastern Cemetery of Roman London - Excavations 1983 – 1990*. MoLAS monograph 4. London: Museum of London Archaeology S.

Baten, J. (1999). *Ernährung und Wirtschaftliche Entwicklung in Bayern (1730-1880).* Stuttgart: Franz Steiner Verlag.

Baten, J. (1999b). Kartographische Residuenanalyse am Beispiel der Regionalökonomischen Lebensstandardforschung über Baden, Württemberg und Frankreich. In: D. Ebeling (ed.). *Historisch-thematische Kartographie: Konzepte, Methoden, Anwendungen*. Bielefeld: Verlag für Regionalgeschichte. 98-109.

Baten, J. (2000). Economic Development and the Distribution of Nutritional Resources in Bavaria, 1797-1839. *Journal of Income Distribution* 9, 89-106.

Baten, J. (2002). Climate, Grain Production and Nutritional Status in 18th Century Southern Germany. *Journal of European Economic History* 30.1, 9-47.

Baten, J., and Murray, J. (2000). Heights of Men and Women in Nineteenth Century Bavaria: Economic, Nutritional, and Disease Influences. *Explorations in Economic History* 37, 351-369.

Bisel, S. (1988). Nutrition in First Century Herculaneum. *Anthropologie* XXVI/1, 61-66.

Bogin, B. (1988). *Patterns of Human growth*. Cambridge: Cambridge University Press.

Bogin, B. (1991). Measurement of Growth Variability and Environmental Quality in Guatemalan Children. *Annals of Human Biology* 18, 285-294.

Bogin, B. (1995). Human growth: A practitioner's guide. *Practicing Anthropology* 17, 25-28.

Bogin, B. (1995). *Human Variability and Plasticity*. Cambridge Studies in Biological Anthropology, Cambridge: Cambridge University Press (ed. by C.G.N. Mascie-Taylor and B. Bogin).

Brothwell, D. (2003). Wealth, Health and theoretical Archaeology. In: *Wealth, Health and Human Remains in Archaeology*. Symposium Nederlands Museum Anthrop. Praehist. Amsterdam 2003. Amsterdam. W.H. Metz: 39-55.

Clark, G. (2001). The Secret Story of the Industrial Revolution. Paper presented to the Fourth European Historical Economics Conference Oxford 2001.

Cohen, M. (1989). *Health and the Rise of Civilization*. New Haven: Yale University Press.

Cohen, M., and Armelagos, G. (1984). *Palaeopathology at the Origins of Agriculture*. Orlando/London: Academic Press.

Craig, L., et al (2002). The 'Great Inventions' and Human Nutrition: the Case of Mechanical Refrigeration. Proceedings (on CDROM) of the first International Conference on Economics and Human Biology. Tuebingen 2002.

McCloskey, D., and Ziliak, S. (1996). The Standard Error of Regressions. *Journal of Economic Literature* XXXIV, 97-114.

Dupâquier, J., et al. (1988). *Histoire de la population Française*. Paris: Presses Univ. de France.

McEvedy, C., and Jones, R. (1980^2). *Atlas of World Population History*. London/ Harmondsworth: Penguin.

Federico, G. (2002). The World Economy 0-2000 AD: A review article". *European Review of Economic history* 6, 111-120.

Federico, G., and Malanima, P. (2002). Labour Productivity in Italian Agriculture 1000 – 2000. Proceedings (on CDROM) of the XIII International Economic History Association Congress. Buenos Aires 2002.

Galloway, P. (1986). Long-Term Fluctuations in Climate and Population in the Pre-Industrial Era". *Population and Development Review* 12.1, 1-24.

Grewe, K. (1986). Zur Wasserversorgung und Abwasserentsorgung in der Stadt um 1200. In: H. Steuer (ed.). *Zur Lebensweise in der Stadt um 1200*. Zeitschrift für Archäologie des Mittelalters / Beiheft 4. Köln: Rheinland-Verlag. .

Grove, J. (2002). Climatic Change in Northern Europe over the Last Two Thousand Years and its Possible Influence on Human Activity. In: Wefer, G. et al. (eds.). *Climate Development and history of the North Atlantic Realm. Hanse Conference on Climate History.* Berlin/Heidelberg: Springer. 313-326.

Heide, A. (1997). *Das Wetter und Klima römischer Antike im Westen des Reiches*. Unpublished Ph.D. thesis. Mainz.

Herrmann, B. (1985). Parasitologisch-epidemiologische Auswertungen mittelalterlicher Kloaken. *Zeitschrift für Archäologie des Mittelalters* 13, 131-152.

Hermann, B. (1987). Anthropologische Zugänge zu Bevölkerung und Bevölkerungsentwicklung im Mittelalter. In: Hermann, B., and Spandel, R. (eds.).

Determinanten der Bevölkerungsentwicklung im Mittelalter. Weinheim: Acta Humaniora, VCH Verlagsgesellschaft. 55-72.

Huntley, B., et al. (2002). Holocene Climatic History of Northern Europe as Recorded by Vegetation Changes: Possible Fluctuations upon Human Activity. In: Wefer, G. et al. (eds.). *Climate Development and history of the North Atlantic Realm. Hanse Conference on Climate History.* Berlin/Heidelberg: Springer. 223-232.

Karpf, G., and Volk, (1973). Medizingeschichtliche, anthropologische-paläopathologische und paläodemographische Befunde des römischen Brandgräberfeldes von Schallstadt. *Archäologisches Korrespondenzblatt* 3, 51-58.

Klasen, S. (1996). Nutrition, Health, and Mortality in Sub-Saharan Africa: Is there a Gender Bias? *Journal of Development Studies* 32, 913-933.

Klasen, S. (forthcoming). Does Gender Inequality in Education Reduce Economic Development: Evidence from Cross-Country Regressions. *World Bank Economic Review.*

Kelly, M. (no year given), *Climate and Pre-Industrial Growth.* working-paper on: http://www.ucd.ie/economic/staff/mkelly/papers/climate.pdf.

Koepke, N. (2002). Regional Differences and Temporal Development of the Quality of Nutrition in the Roman Provinces of Germania and Raetia from the First Century to the Fourth Century AD”. Proceedings (on CDROM) of the XIII International Economic History Association Congress. Buenos Aires 2002.

Komlos, John (1985). Stature and Nutrition in the Habsburg Monarchy: The Standard of Living and Economic Development, *American Historical Review* 90, 1149-1161.

Komlos, J. (1989). *Nutrition and Economic Development in the Eighteenth Century Habsburg Monarchy: An anthropometric History.* Princeton, NJ : Princeton University Press.

Komlos, J., and Baten, J. (1998). Height and the Standard of Living. *Journal of Economic History* 57.3, 866-870.

Komlos, J., Hau, M., and Bourguinat, N. (2003). The Anthropometric History of Early-Modern France. *European Review of Economic History*, forthcoming.

Mc Keown, T. (1955). Medical Evidence Related to English Population Changes in the Eighteenth Century. *Population Studies* IX, 119-141.

Maat, G. (2003). Male Stature. A Parameter of Health and Wealth in the Low Countries, 50-1997 AD. In: *Wealth, Health and Human Remains in Archaeology*. Symposium Nederlands Museum Anthrop. Praehist. Amsterdam 2003. Amsterdam. W.H. Metz: 57-88.

Maddison, A. (2001). *The World Economy: A Millennial Perspective*. Paris: Development Centre of the OECD.

Malanima, P. (forthcoming). *L'economia italiana della crescita medivevale alla crescita contemporana*.

Mays, S. (1995). Killing the Unwanted Child. *British Archaeology 2*. http://www.britarch.ac.uk/ba/ba2/ba2feat.html#mays

Mokyr, J. (1992). *Lever of Riches: Technological Creativity and Economic Progress*. Oxford: Oxford University Press.

Mokyr, J., and O'Grada, C. (2002), What Do People Die of During Famines? The Great Irish Famine in comparative perspective. *European Review of Economic History* 6.3, 339-363.

Ortner, D. (1998). Male-female Immune Reactivity and its Implications for Interpreting Evidence in Human Skeletal Paleopathology. In: Grauer, A., and Stuart-Macadam (eds.), *Sex and Gender in Paleopathological Perspective.* Cambridge: Cambridge University Press. 79-92.

Osmani, S. and Sen, A. (2003). The hidden penalties of gender inequality: foetal origins of ill-health. *Economics and Human Biology* 1, 105-121.

Pfister, C. (1988^{3}). *Klimageschichte der Schweiz 1525 – 1860. Das Klima der Schweiz von 1525 – 1860 und seine Bedeutung in der Geschichte von Bevölkerung und Landwirtschaft*. Bern/Stuttgart. Haupt Verlag.

Pomerantz, K. (2002). *The Great Divergence: Europe, China, and the Making of the Modern World Economy*. Princeton: Princeton University Press.

Pradhan, M., Sahn, D., and Younger, S. (2002). Decomposing World Health Inequality". Paper presented to the First International Conference of Economics and Human Biology. Tübingen 2002.

Quiroga Valle, G. (1998). Height Evolution in Spain, 1893-1954: An Analysis by Regions and Professions. In: Komlos, J., and Baten, J. (eds.). *The Biological Standard of Living in Comparative Perspective*. Stuttgart: Franz Steiner Verlag: 359-383.

O'Rourke, K., and Williamson, J. (2002). When Did Globalisation Begin?. *European Review of Economic History* 6.1, 23-50.

Rösing, F., and Caselitz, (1998). Der Mensch selbst als Indikator für wirtschaftliche, soziale und politische Rahmenbedingungen von Burg und Stadt im Mittelalter. In: Biller, T., et al. (ed.). *Schloss Tirol. Saalbauten und Burgen des 12. Jahrhunderts in Mitteleuropa*. Munich: Deutscher Kunstverlag. 241-248.

Scobie, A. (1986). Slums, Sanitation, and Mortality in the Roman World. *Klio* 68.2, 399-433.

Schweich, M., and Knüsel, C. (2002). Bio-cultural Effects on Stature and Body Proportions in European Archaeological Populations." Proceedings (on CDROM) of the first International Conference on Economics and Human Biology. Tuebingen 2002.

Steckel, R. (1995). Stature and the Standard of Living. *Journal of Economic Literature* 33.4, 1903-1940.

Steckel, R. (2003). Research project: A history of health in Europe from the late Paleolithic era to the present, *Economics & Human Biology* I 1, 139-142.

Steckel, R. (forthcoming). Net Nutrition over the Past Millennium: Methodology and some Results for Northern Europe. *Social Science History* 2004.

Steckel, R., and Prince, J. (2001). Tallest in the World: Native Americans of the Great Plains in the Nineteenth Century. *American Economic Review* 91, 287-294.

Steckel, R., and Rose, J. (2002). *The Backbone of History: Health and Nutrition in the Western Hemisphere*. Cambridge: Cambridge University Press.

Ulrich-Bochsler, S. (1996). Anthropologische Befunde zu Frau und Kind in Mittelalter und Neuzeit unter besonderer Berücksichtigung der Stellung der Früh- und Neugeborenen: soziobiologische und soziokulturelle Aspekte im Lichte von Archäologie, Geschichte, Volkskunde und Medizingeschichte. Basel, Univ., Diss. Mikrofiche.

Volk, et al. (1988). Frauen am römischen Limes. Anthropologische Rekonstruktion der Bevölkerung von Kastell und Vicus Schirenhof. In: M. Hilgarth, Mönig-Schuth, M. (Hrsg.). *Festschrift H.-G. Hillemanns*. Konstanz: Druckerei am Fischmarkt. 549-587.

de Vries, J.(1984). *European Urbanization 1500 – 1800*. London: Methuen.

Wrigley, E., and Schofield, R. (1981). *The Population History of England 1541 – 1871: A reconstruction*. Cambridge/Mass.: Harvard University Press.

Wurm, H. (1985). Über die Durchschnittlichen Körperhöhen der Sozialen Mittel- und Untersichten im Mitteleuropäischen Germanischen Siedlungsraum vom Frühmittelalter bis zur Neuzeit. *Anthropologischer Anzeiger* 43, 11-30.

Wurm, H. (1985b). Die Körperhöhe deutscher Harnischträger: Ein Beitrag zur Frage der durchschnittlichen Körperhöhen der Soldaten im Deutschen Reich im 16. und 17. Jahrhundert unter Berücksichtigung regionaler und sozialer Unterschiede und anthropologischen Vergleichmaterials. *Zeitschrift für Morphologie und Anthropologie* 75.2, 155-188.

van Zanden, J.L. (1995). Tracing the Beginning of the Kuznets Curve: Western Europe during the Early Modern Period. *Economic History Review* XLVIII.4, 643-664.

van Zanden, J.L. (1999). Wages and Standard of Living in Europe, 1500 – 1800. *European Review of Economic History* 3.2, 175-197.

Tübinger Wirtschaftswissenschaftliche Diskussionsbeiträge

Die Liste der hier aufgeführten Diskussionsbeiträge beginnt mit der Nummer 177 im Jahr 2000. Die Texte können direkt aus dem Internet bezogen werden. Sollte ein Interesse an früher erschienenen Diskussionsbeiträgen bestehen, kann die vollständige Liste im Internet eingesehen werden. Die Volltexte der dort bis Nummer 144 aufgeführten Diskussionsbeiträge können nur direkt über die Autoren angefordert werden.

177. **Stadler, Manfred und Stephan O. Hornig:** Wettbewerb bei unvollständiger Information: Informationsaustausch oder stillschweigende Kollusion? Januar 2000.
178. **Jung, C. Robert und Roman Liesenfeld:** Estimating Time Series Models for Count Data Using Efficient Importance Sampling, Januar 2000.
179. **Stadler, Manfred und Rüdiger Wapler:** Arbeitsmarkttheorie, Februar 2000.
180. **Wapler, Rüdiger:** Unions, Monopolistic Competition and Unemployment, Februar 2000.
181. **Hornig, Stephan O.:** When Do Firms Exchange Information?, März 2000.
182. **Preuße, Heinz Gert:** Entwicklungen in der US-amerikanischen Außenhandelspolitik seit der Gründung der Nordamerikanischen Freihandelszone (NAFTA), März 2000.
183. **Preuße, Heinz Gert:** Sechs Jahre Nordamerikanisches Freihandelsabkommen (NAFTA) - Eine Bestandsaufnahme, März 2000.
184. **Starbatty, Joachim:** Struktur- und Industriepolitik in einer Welt konstitutioneller Unwissenheit, März 2000.
185. **Woeckener, Bernd:** Spatial Competition of Multi-Product Retail Stores with Store-Specific Variety Effects, April 2000.
186. **Bayer, Stefan:** Altruism and Egoism: Measurable by Utility Discount Rates?, April 2000.
187. **Bayer, Stefan:** Generation Adjusted Discounting in Long-term Decision-making, Mai 2000.
188. **Cansier, Dieter:** Freifahrerverhalten und Selbstverpflichtungen im Umweltschutz, Mai 2000.
189. **Kellerhals, B. Philipp und Rainer Schöbel:** The Dynamic Behavior of Closed-End Funds and its Implication for Pricing, Forecasting and Trading, Juli 2000.
190. **Bühler, Wolfgang , Korn Olaf und Rainer Schöbel:** Pricing and Hedging of Oil Futures –A Unifying Approach-, Juli 2000.
191. **Woeckener, Bernd:** Spatial Competition with an Outside Good: a Note, August 2000.
192. **Woeckener, Bernd:** Standards Wars, August 2000.

193. **Opper, Sonja und Joachim Starbatty:** Reflections on the Extension of Human Rights from the Economic Perspective, September 2000.
194. **Hornig, Stephan und Manfred Stadler:** No Information Sharing in Oligopoly: The Case of Price Competition with Cost Uncertainty, Oktober 2000.
195. **Duijm, Bernhard:** A First Evaluation of the Institutional Framework for European Monetary Policy, Oktober 2000.
196. **Edlund, Lena und Evelyn Korn:** An Economic Theory of Prostitution, Oktober 2000.
197. **Bayer, Stefan und Claudia Kemfert:** Reaching National Kyoto-Targets in Germany by Mainting a Sustainable Development, Oktober 2000.
198. **Preusse, Heinz Gert:** MERCOSUR – Another Failed Move Towards Regional Integration? November 2000.
199. **Böckem, Sabine und Ulf Schiller**: Contracting with Poor Agents, November 2000.
200. **Schiller, Ulf:** Decentralized Information Acquisition and the Internal Provision of Capital, November 2000.
201. **Leitner, Frank:** Die Entstehung von Runs auf Banken unter verschiedenen Umweltbedingungen, Dezember 2000.
202. **Gampfer, Ralf:** Die optimale Versteigerungsreihenfolge in sequentiellen Zweitpreisauktionen bei Synergieeffekten, Dezember 2000.
203. **Eisele, Florian, Werner Neus und Andreas Walter:** Zinsswaps – Funktionsweise, Bewertung und Diskussion, Januar 2001.
204. **Jung, Robert und Andrew R. Tremayne:** Testing Serial Dependence in Time Series Models of Counts Against Some INARMA Alternatives, Januar 2001.
205. **Heilig, Stephan und Rainer Schöbel:** Controlling Chaos in a Model with Heterogeneous Beliefs, Januar 2001.
206. **Wapler, Rüdiger:** Unions, Growth and Unemployment, Februar 2001.
207. **Woeckener, Bernd:** Compatibility decisions, horizontal product differentiation, and standards wars, Mai 2001.
208. **Kellerhals, B. Philipp und Rainer Schöbel:** Risk Attitudes of Bond Investors, Mai 2001.
209. **Kellerhals, B. Philipp:** Pricing Electricity Forwards under Stochastic Volatility, Mai 2001.
210. **Wapler, Rüdiger:** Unions, Efficiency Wages and Unemployment, August 2001.
211. **Starbatty, Joachim:** Globalisierung und die EU als „sicherer Hafen" – einige ordnungspolitische Anmerkungen, Juli 2001.
212. **Kiesewetter, Dirk und Rainer Niemann:** Beiträge und Rentenzahlungen in einer entscheidungsneutralen Einkommensteuer, August 2001.
213. **Schnabl, Gunther und Dirk Baur:** Purchasing Power Parity: Granger Causality Tests for the Yen-Dollar Exchange Rate, August 2001.
214. **Baten, Jörg:** Neue Quellen für die unternehmenshistorische Analyse, August 2001.
215. **Baten, Jörg:** Expansion und Überleben von Unternehmen in der „Ersten Phase der Globalisierung", August 2001.
216. **Baten, Jörg:** Große und kleine Unternehmen in der Krise von 1900-1902, August 2001.
217. **Baten Jörg:** Produktivitätsvorteil in kleinen und mittelgroßen Industrieunternehmen, Sicherheit in Großunternehmen? Die Gesamtfaktorproduktivität um 1900, August 2001.
218. **Schnabl, Gunther:** Weak Economy and Strong Currency – the Origins of the Strong Yen in the 1990's, August 2001.
219. **Ronning, Gerd:** Estimation of Discrete Choice Models with Minimal Variation of Alternative-Specific Variables, September 2001.
220. **Stadler, Manfred und Rüdiger Wapler:** Endogenous Skilled-Biased Technological Change and Matching Unemployment, September 2001.

221. **Preusse, Heinz G.:** How Do Latin Americans Think About the Economic Reforms of the 1990s?, September 2001.
222. **Hanke, Ingo:** Multiple Equilibria Currency Crises with Uncertainty about Fundamental Data, November 2000.
223. **Starbatty, Joachim:** Zivilcourage als Voraussetzung der Freiheit – Beispiele aus der Wirtschaftspolitik - , Oktober 2001.
224. **Kiesewetter, Dirk:** Zur steuerlichen Vorteilhaftigkeit der Riester-Rente, Dezember 2001.
225. **Neubecker, Leslie:** Aktienkursorientierte Management-Entlohnung: Ein Wettbewerbshemmnis im Boom?, Dezember 2001.
226. **Gampfer, Ralf:** Internetauktionen als Beschaffungsinstrument: Eigenständige oder Integrierte Lösung?, Dezember 2001.
227. **Buchmüller, Patrik:** Die Berücksichtigung des operationellen Risikos in der Neuen Basler Eigenkapitalvereinbarung, Dezember 2001.
228. **Starbatty, Joachim:** Röpkes Beitrag zur Sozialen Marktwirtschaft, Januar 2002.
229. **Nufer, Gerd:** Bestimmung und Analyse der Erfolgsfaktoren von Marketing-Events anhand des Beispiels DFB-adidas-Cup, März 2002.
230. **Schnabl, Gunther:** Asymmetry in US-Japanese Foreign Exchange Policy: Shifting the Adjustment Burden to Japan, März 2002.
231. **Gampfer, Ralf:** Fallende Preise in Sequentiellen Auktionen: Das Beispiel des Gebrauchtwagenhandels, März 2002.
232. **Baur, Dirk:** The Persistence and Asymmetry of Time-Varying Correlations, März 2002.
233. **Bachmann, Mark:** Ermittlung und Relevanz effektiver Steuersätze. Teil 1: Anwendungsbereich und Modellerweiterungen, März 2002.
234. **Knirsch, Deborah:** Ermittlung und Relevanz effektiver Steuersätze. Teil 2: Der Einfluss der Komplexitätsreduktion von Steuerbemessungsgrundlagen, März 2002.
235. **Neubecker, Leslie:** Aktienkursorientierte Managemententlohnung bei korrelierter Entwicklung der Marktnachfrage, März 2002.
236. **Kukuk, Martin und Manfred Stadler:** Rivalry and Innovation Races, März 2002.
237. **Stadler, Manfred:** Leistungsorientierte Besoldung von Hochschullehrern auf der Grundlage objektiv meßbarer Kriterien?, März 2002.
238. **Eisele, Florian, Habermann, Markus und Ralf Oesterle:** Die Beteiligungskriterien für eine Venture Capital Finanzierung – Eine empirische Analyse der phasenbezogenen Bedeutung, März 2002.
239. **Niemann, Rainer und Dirk Kiesewetter:** Zur steuerlichen Vorteilhaftigkeit von Kapitallebensversicherungen, März 2002.
240. **Hornig, Stephan:** Information Exchange with Cost Uncertainty: An Alternative Approach with New Results, Juni 2002.
241. **Niemann, Rainer, Bachmann, Mark und Deborah Knirsch:** Was leisten die Effektivsteuersätze des European Tax Analyzer?, Juni 2002.
242. **Kiesewetter, Dirk:** Tax Neutrality and Business Taxation in Russia: A Propsal for a Consumption-Based Reform of the Russian Income and Profit Tax, Juni 2002.
243. **McKinnon, Ronald und Gunther Schnabl:** Synchronized Business Cycles in East Asia and Fluctuations in the Yen/Dollar Exchange Rate, Juli 2002.
244. **Neus, Werner:** Fusionsanreize, strategische Managerentlohnung und die Frage des geeigneten Unternehmensziels, Juli 2002.
245. **Blüml, Björn und Werner Neus:** Grenzüberschreitende Schuldverträge und Souveränitätsrisiken, Juli 2002.

---

246. **Starbatty, Joachim:** Die Abschaffung der DM ist noch keine Bereitschaft zur politischen Union, Juli 2002.
247. **Schnabl, Gunther:** Fear of Floating in Japan? A Bank of Japan Monetary Policy Reaction Function, September 2002.
248. **Brassat, Marcel und Dirk Kiesewetter:** Steuervorteile durch Versorgungszusagen in Arbeitsverträgen, September 2002.
249. **Knirsch, Deborah:** Neutrality-Based Effective Tax Rates, September 2002.
250. **Neubecker, Leslie:** The Strategic Effect of Debt in Dynamic Price Competition with Fluctuating Demand, November 2002.
251. **Baur, Dirk und Robert Jung:** Return an Volatility Linkages Between the US and the German Stock Market, Dezember 2002.
252. **McKinnon, Ronald und Gunther Schnabl:** The East Asian Dollar Standard, Fear of Floating, and Original Sin, Januar 2003.
253. **Schulze, Niels und Dirk Baur:** Coexceedances in Financial Markets – A Quantile Regression Analysis of Contagion, Februar 2003.
254. **Bayer, Stefan:** Possibilities and Limitations of Economically Valuating Ecological Damages, Februar 2003.
255. **Stadler, Manfred:** Innovation and Growth: The Role of Labor-Force Qualification, März 2003.
256. **Licht, Georg und Manfred Stadler:** Auswirkungen öffentlicher Forschungsförderung auf die private F&E-Tätigkeit: Eine mikroökonometrische Evaluation, März 2003.
257. **Neubecker, Leslie und Manfred Stadler:** Endogenous Merger Formation in Asymmetric Markets: A Reformulation, März 2003.
258. **Neubecker, Leslie und Manfred Stadler:** In Hunt for Size: Merger Formation in the Oil Industry, März 2003.
259. **Niemann, Rainer:** Wie schädlich ist die Mindestbesteuerung? Steuerparadoxa in der Verlustverrechung, April 2003.
260.
261. **Neubecker, Leslie:** Does Cooperation in R&D Foster Tacit Collusion?, Juni 2003.
262. **Buchmüller, Patrik und Christian Macht:** Wahlrechte von Banken und Aufsicht bei der Umsetzung von Basel II, Juni 2003.
263. **McKinnon, Ronald und Gunther Schnabl:** China: A Stabilizing or Deflationary Influence in East Asia? The Problem of Conflicted Virtue, Juni 2003.
264. **Thaut, Michael:** Die individuelle Vorteilhaftigkeit der privaten Rentenversicherung – Steuervorteile, Lebenserwartung und Stornorisiken, Juli 2003.
265. **Köpke, Nikola und Jörg Baten:** The Biological Standard of Living in Europe During the Last Two Millennia, September 2003.

II. Description of our variables

height	by century and European regions (north-east, central-east, Mediterranean) measured in cm, of both male and female individuals (adjusted to male level, regionally adjusted to the average of the European region)
climate warm	series of glacier movements and tree-ring-series from Switzerland, Greenland and Sweden
gender inequality	measured by height differences of male and female individuals as % of male heights
urban share	urbanization rate for the period of: 0-1000 are rough estimates based on the adjusted share of skeletons of our sample that were found in urban places: For Central-Western Europe we did set the 1st century to 1%, the „Roman centuries“ to 5%, the following early middle ages to 1%. For Northern/Eastern Europe Antiquity and Early Middle Ages were assigned 1%. The Mediterranean was assigned 10% for centuries 1-4 and 5% after the decline of the Roman empire. 1000-1300 is based on Federico/ Malanima (2000) and Malanima (forthcoming), with adjustments for Central-West and North-Eastern Europe; 1300-1800 for ‘Mediterranean’ is based on Federico/ Malanima (2000) and Malanima (forthcoming); weighted by population average (McEvedy 1300-1800). As we have mostly data on Sweden and Denmark, the Scandinavia weight is based on these countries.
population density	created for each country by using the information from McEvedy and Jones (1980[2]); except for other

	data for the following regions: for England from 1500 to 1800 Wrigley/Schofield (1981); as the data provided by E. Wrigley/ R. Schofield (1981) is only for England we completed these with the Irish and Scottish data from McEvedy/Jones. We calculated the weighted mean for the three ‘large regions’ (of ‘Mediterranean’, North-Eastern-Europe, Central-Western Europe) by averaging according to population size. Criterion for inclusion of a country was the number of cases. The ‘Mediterranean’ region is constructed with data on Iberia and Italy. We took as parts of ‘North-Eastern-Europe’: Poland, Czechoslovak Republic, Austria, Hungary, Denmark and Sweden. ‘Central-Western Europe’ in our case is: Switzerland, Belgium, Germany, Netherlands and Great Britain & Ireland.
Roman bath / other technology	1 for institution of *balnea* and other aspects of Roman technology in ancient Italy and the Roman provinces
social inequality	measured by height differences of individuals of ‘high/middle’ economic status and ‘low/unknown’ status

III. Estimates with and without group measurements

Our sample consists of 2774 individual heights and 200 aggregated height figures that refer to the remaining 6703 cases. Such a combination of grouped and individual data is a complicated econometric issue, because the better documented cases deserve a higher weight in the regressions, but as their distribution is unknown, the calculation of usual t-statistics requires additional assumptions. In addition, the weight should not be arbitrary, it should be greater than 1 for the reasons mentioned, but it should also be less than the number of individuals over which the average was calculated.

In a first step, we basically used the square root of the number of cases as importance weight in WLS regressions, so that the aggregated cases get a slightly smaller influence on the results than the individual measurements.

The basic econometric issue is this: As we have N people with a mean height of H and put in H as the obs and sqrt(H) as the weight, then at the very least, this observation is stochastic while the others are not (the standard errors can be both, too large and too small). So our standard errors are all not appropriately estimated because of this fact. There is also an efficiency issue, because we do not really have N observations, we just have the mean (this issue means that the standard errors might be too large). There are strong variance and covariance assumptions needed to make regression with grouped data equal to regression with individual data.

We do not directly investigate those assumptions. Since we have a large data set for most centuries, we instead simply drop the grouped (aggregated) observations and compare the estimates of those individual-only heights and the full data set (Figure A.1). The results are very robust, but the 14th to 16th centuries are poorly documented without the grouped observations. In all three centuries, there are less than 20 observations (if we would show those estimates based on N=13, N=6, and N=1, the estimates would be 1-2 cm below the estimates that included the grouped observations). For all other centuries, the height estimates based on all cases with Square Root Weighted Least Squares and with individual data only are highly correlated. If the estimates are performed by the three European regions, the correspondence is similar (except that there are more poorly documented centuries in which the deviations are naturally larger).

Figure A.1

HEIGHT DEVELOPMENT, 1st TO 18th CENTURIES: ALL CASES AND AGGREGATED HEIGHTS EXCLUDED

(IN CM, MALE AND FEMALE)

Source: see table 1. The level of heights was adjusted to male heights of an average European (using the regional coefficients and weighting them with sample weights).

Additional measurement issues

We excluded among the grouped data all height estimates that were not exclusively based on one gender.

IV. Stature reconstruction methods

How to estimate human stature from long-bones and cremated bones

Basics

Archaeological excavators typically do not find complete skeletons in cemeteries of early settlements, but fortunately methods have been developed to estimate human stature, even if only one long bone survived, or a (certain) part of the long bone.[1][1] The length of a long bone represents an approximately constant proportion of height. Anthropologists did not yet reach full agreement about the optimal estimation procedure. Naturally all the (available) formulas only give an estimate of height. The femur (thigh) gives the best approximation to reconstruct height of inhumated skeletons; luckily it is the most frequently survived bone.[2][2]

Regression models

---

[1][1] If only a part of the long bone is available, its length of the long bone has to be reconstructed first before height can be reconstructed; because two regressions are needed in this case, the measurement error of the result is larger.

But of course the calculation of heights has a smaller estimation error if the preservation of the bones is good: completely preserved long-bones provide more exact results than cremated bones.

If we want to compare height of inhumated and cremated individuals we have to take into account that the margin of error is methodologically determined larger for cremated bones.

[2][2] In contrast to the older hypotheses that apart from the femur the humerus results in a smaller measurement error than radius and tibia (e.g. Wurm (1986) 157), current researchers came to the conclusion that normally the long bones of the lower extremities (means apart from the femur: the tibia) give the best approximation (Herrmann et al. (1990) 93 f.).

A large number of various regression models[3][3] exists to reconstruct adult[4][4] height[5][5]: in order to find the best model anthropologists evaluated various reconstruction models for each long bone for both genders[6][6] (fig.1), which provide different results.[7][7]

No general consensus exists yet: anthropological analyses of the various excavation sites are based on these different reconstruction methods. Therefore it makes no sense the reported stature estimates directly. What could we do to enlarge our data base (instead of working only with those observations that were calculated with only one method by the principal investigators)? It was necessary to construct algorithms in order to standardize all height estimations to the most appropriate method.

Which is the best reconstruction model?

Most anthropologists and archaeologists used the following five models (s. tab.1): one method was exclusively used for cremation analysis (Rösing) and four methods to

---

[3][3] Naturally different regression formulas have to be used for females and males.

Furthermore the high number of models is not only the result of different reconstruction models for inhumated and cremated bones - as the latter show diminution (because they were exposed to the heat) - ; only for the reconstruction of inhumations 43 models exist. On cremation at least 22 researchers worked on different models: recapitulation of the research history see Wahl (1982) and Heussner (1987): esp. Dijkstra, Müller, Wahl, Rösing.

[4][4] It is also possible to reconstruct height of non-adult individuals, best by the regressions by Telkkä et al. (1962) for 3 different age groups (age under 1a, 1-9years and 10-15 years). But these are rarely used, as the gender of the child must be known.

[5][5] Discussed e.g. by Hermann et a. (1990), Rösing (1988), Wurm (1985), Wurm (1986) and Wurm and Leimeister (1986).

[6][6] E.g. Bach (1965); Breitinger (1938); Dupertius/Hadden (1951); Manouvrier (1893); Rösing (1977); Trotter und Gleser (1952); Trotter (1970); Olivier et al. (1978); Pearson (1899); Wolanski (1953).

[7][7] If we compare the models that reconstruct height by exactly the same length of bone (e.g. femur f1) the result can vary several centimetres (irrespective the standard error).

reconstruct heights of inhumated individuals; these are: Breitinger (for males) and Bach (for females), Trotter/Gleser as well as Pearson (respectively Wolanski), and less often Manouvrier.[8][8]

Table. 1

THE FIVE MOST COMMON RECONSTRUCTION MODELS

method [9][9]	**N**	**% (rounded)**
Breitinger and Bach	1421	47.9
Trotter/Gleser[w]	538	18.1
Pearson (incl. Wolanski)	336	11.3
Manouvrier	172	5.9
mixed	102	3.4
Rösing (cremated bones only)	153	5.1

We have to calculate conversion factors to get consistency of the data set. Which is the best regression model to use, into which the other equation should be converted? We decided to use the reconstruction formulas after Breitinger/Bach as the 'basic' model, because it is the most commonly used model in the data we collected by now.[10][10] Furthermore it gives the most exact results (lowest mean of differences in stature) for the relevant spectrum of centimetres[11][11]: For the height range 164 – 178.9 cm for

---

[8][8] Stature reconstructions by the also quite commonly used model by Olivier were sorted out as these are not based on the same Martin (1928) -measurements of the bones like the above mentioned ones.

[9][9] For another 231 individuals (7.8%) femurs were reported without reconstruction, and 21 heights (0.7%) were reported with mixed methods.

[10][10] If we know that the excavators and principal investigators used only femurs, we can recalculate each stature estimate into the corresponding estimate of another reconstruction method. However sometimes the authors do not mention if they estimated only with femurs or, for example, with a combination of bones (which would make the estimates more exact), like femur and tibia, or only with another bone, like tibia. In order to minimize recalculation error, we prefer to use the calculation method that most principal investigators employed.

[11][11] Compare Formicola (1993). But note, that, if the height of especially small individuals (males below 160 cm) is reconstructed Pearson might be more efficient.

males Formicola (1993) arrives at the lowest estimation error for Breitinger, whereas the Trotter/Gleser and Pearson methods[12][12] are less efficient.[13][13]

A third argument for Breitinger is that Trotter/Gleser arrive at higher, and Pearson at lower height estimates, whereas Breitinger/Bach estimates fall between them. Thus Breitinger/Bach estimates are normally close to a "compromise" estimate (that would "average" the three most widely used methods (s. example, tab.2)).

Table 2

MEAN MALE HEIGHT OF OUR SAMPLE, CONVERTED MEASUREMENTS IN COMPARISON

Breitinger	Pearson	Trotter/Gleser[w]
169.7	166.1	170.7

In our case 92.2% of the female observations and 89.3% of the male observations fall into the range mentioned above.

And last but not least the Breitinger and Bach measurements are the best to compare with measurements after Rösing (for cremated bones), which can not be converted, because the reconstruction formulas are based on different parts of bones[14][14]: in the case of cremated individuals these are the diameters of long bone heads (e.g. femur F18, see fig. 2), but in the case of inhumated skeletons these are the length of the long bones (e.g. femur F1 'largest length'[15][15], see fig.2).[16][16]

---

[12][12] He does not included Manouvrier's method.

[13][13] That is also the case for the Søjvold (1990).

[14][14] Measurements used for all methods after R. Martin (1928).

[15][15] It is the distance from the highest point of the caput to the lowest point of the condylus tibialis.

[16][16] According to J. Wahl the stature estimates with the Breitinger/Bach method comes nearest to the Rösing estimates. Because we want to compare Rösing values, it is best to use Breitinger and Bach

-

Figure 1

ADULT SKELETON, FRONTAL VIEW.
Long bones used for stature reconstruction models in capitals

-

---

measures. – We thank Dr. J. Wahl, University of Tübingen and Landesdenkmalamt (Bureau of ancient heritage and monuments) Baden-Württemberg for his kind verbal information.

Fig. 2.

MOST OFTEN USED LONG BONE MEASUREMENTS
FOR HEIGHT ESTIMATES: FEMUR

for inhumations:

measurement F1: ‚largest length’: distance of the highest point of the head of femur/ *caput femoris* to the lowest point of the medial condyle/ *Condylus tibialis (= C. medialis)*

for cremations:

measurement F18: ‘vertical diameter of the femur head’: linear distance of the highest to the lowest point of the *caput femoris*