J.biosoc.Sci, (2007) **39**, 59–78, © 2005 Cambridge University Press
doi:10.1017/S0021932005001161 First published online 7 Dec 2005

SPATIAL CORRELATES OF US HEIGHTS AND BODY MASS INDEXES, 2002

JOHN KOMLOS* AND BENJAMIN E. LAUDERDALE†

**Department of Economics, University of Munich, Germany and †Department of Politics, Princeton University, USA*

Summary. Aiming to further explore possible underlying causes of the recent remarkable stagnation and relative decline in American heights, this paper describes the result of analysis of the commercial US Sizing Survey (2002). Heights are correlated positively with income and education among both white males and females while Body Mass Index (BMI) is correlated negatively among females, as in other samples. In contrast to much of the literature, this paper considers geographic correlates of height such as local poverty rate, median income and population density at the zip code level of resolution. After adjusting for confounding factors that influence height such as income and education, population density is found to be strongly and negatively correlated with height among white men, but less so among white women. The effect on BMIs less convincing. Other ethnic groups are not analysed in detail because of the small number of observations available. Local economic conditions as measured by median income, unemployment and poverty rate do not have a strong correlation with height or BMI after adjusting for individual income and education.

Introduction

Why have the heights of Americans, the tallest in the world until the mid-20th century, stagnated recently while those of Western and Northern Europe have increased substantially? Western and Northern European heights generally surpassed American heights in the 1970s and the mean height difference is currently circa 2–6 cm (1–2 inches) (Fredriks *et al.*, 2000; Sunder, 2003). Not merely a question for anthropometricians, this is an issue of broad interest as the mean height of populations often reflects differences in health and longevity. It is well documented that early life nutrition and disease are the major environmental influences on terminal height (Waaler, 1984; Costa, 1993; Komlos & Cuff, 1998; Komlos & Baten, 1998). Height is an indicator of past environmental conditions (net); that is, it conveys information on the history of nutritional intake net of claims of disease incidence and other claims on nutrition such as work effort. Health delivery also plays a role as it affects the virulence and length of sickness. Although changes in population height are sensitive to current levels of nutrition during infancy, childhood and adolescence, mean adult attained height actually reflects the accumulated past nutritional experience of an individual over all of their growing years, particularly the fetal period, early childhood and adolescence. Similarly, the mean height of a population represents the cumulated nutritional status of that population during the periods of height growth of its members. Once growth ceases in early adulthood environmental conditions have no effect on height until shrinkage begins in old age. Hence, attained height can be influenced by genetics, food availability and access (quantity and quality), disease patterns, access to potable water, health services and parents' educational level and SES (occupation or income) that prevailed during the periods of height growth of the observed population (Floud, 1992). However, the disparity in height between Europe and North America has been challenging to explain in these terms, given the increasing economic prosperity experienced on both sides of the Atlantic.

While the income gradient in self-reported health is steeper in the US than in Canada (which has universal health insurance), the gap is smaller for the elderly, who are covered by universal health insurance in both countries (Decker & Remler, 2004). Similarly, Germans evaluate their own health status more positively than Americans (Komlos & Baur, 2004). Moreover, Swedish and Italian life expectancy exceeds that of the US by circa 2·8 years and 2·7 years respectively (Human Development Report 2005). Such findings, as well as the decline in height relative to European populations and rapidly increasing obesity in the US, all indicate that different political choices regarding health care distribution and/or individual choices regarding consumption (health production) might be the cause of stagnation in American height (Komlos *et al.*, 2004). In developed societies where caloric and protein intake is rarely limited by family income, height reflects less the economic output of a community and more its political and social choices that influence overall health during childhood development. This observed discrepancy between material welfare and biological well-being has motivated the formulation of a distinct concept of a biological standard of living (Komlos, 1989).

This paper examines spatial patterns in height and BMI in the US white population in the hope of shedding light on this conundrum. Because there are broad demographic differences between the US and Europe, it may be possible to explain the height discrepancy by linking height to demographic factors within the US. Thus, in addition to using the usual control variables 'own income' and 'education', the effect of such variables at the community level as population density, median income, unemployment rate, and poverty rate on height and BMI is examined. These local environmental factors might offer the opportunity for better describing gradients in American height as well as an avenue for analysing differences between the US and Europe in the future.

The Sizing Survey data

The US Sizing Survey (SizeUSA CD-ROM) was organized in 2002 by a company called '[TC]2' with funding from the US Department of Commerce, a number of clothing manufacturers, and the US military. The goal was to obtain data on the distribution of body size and body proportions in the US population for the purposes of creating better fitting off-the-shelf clothing (http://www.tc2.com or http://www.sizeUSA.com). As such, it contains a large number of variables on various body measurements relevant for the apparel industry. All body measurements with the exception of height and weight were measured with a three-dimensional structured white light technology full-body scanner. In contrast, heights and weights were measured rather crudely: to the nearest half-inch and to the nearest pound respectively. Nonetheless, such rounding of the height and weight measurements will not bias the results. The socioeconomic data are reported categorically for income, education, age and race/ethnicity.

A 'convenience' sample of 10,000 individuals (3689 male, 6311 female; 18 and older) was drawn at shopping centres and places of business around the country, in thirteen geographic clusters. Some data were also collected at universities, business corporate offices, trade show and convention centres, and apparel manufacturing locations. Hence, this is not a random sample. Nonetheless, it is valuable on account of the fact that one can seldom link height data to census sources through zip codes. Thus, it is worthwhile analysing this sample in spite of its shortcomings, as it has advantages as well. The white sample was obtained in cluster locations with wide variations in the number of individuals at each location: Buford, GA (106 individuals); Cary, NC (825); Chattanooga, TN (268); Columbia, MO (772); Dallas, TX (969); Glendale, CA (334); Lawrence, MA (658); Los Angeles, CA (62); Miami, FL (57); New York, NY (250); Portland, OR (156); San Francisco, CA (203); Winston Salem, NC (416).

This sampling procedure precludes making strong claims about the general US population as in the NHANES 1999–2002 survey (Table 1), which is designed to be representative of the US population. The [TC]2 company suggests that the observations have to be weighted in order to conform to NHANES values, but these weights have not been released. Body mass indexes are lower and heights are greater among the survey population than in the NHANES sample, which may indicate systematic bias, insofar as the 'convenience' survey is likely to sample from the more active part of the population with lower BMI values who travel to shopping malls (Table 1). The people sampled might also be more affluent than the typical member of the NHANES sample inasmuch as poorer people are not likely to be found in shopping malls, universities, businesses or convention centres. Moreover, higher BMI people may be less likely to volunteer to be scanned. This is a problem with any scanner survey, because the subjects need to be in spandex (or similar tight device). Since the clustered survey design may have the result of underestimating population variation, all regression analysis is under a generalized linear model that accounts for the risk of underestimating standard errors. The normal ANOVA is not well defined for these models and is not reported.

To explore the relationship between local environment and individual height and weight, the zip codes in the data were linked to summary data from the US 2000 Census compiled by Zip Code Tabulation Area (ZCTA). Zip codes are postal codes that are generally well correlated to geographic areas, especially in highly populated areas. However, since they are assigned for addresses, they do not actually have land area associated with them directly. Consequently, the Census Bureau created geographic areas, ZCTAs, which correspond to these addresses. Because there is some

Table 1. Height as a function of age for men and women in the Size USA and NHANES 1999–2002 surveys, by race

	Men						Women					
Age group	NHANES			Size USA			NHANES			Size USA		
	Height	BMI	Weight	Height	BMI	Weight	Height	BMI	Weight	Height	BMI	Weight
White	inches		pounds	inches		pounds	inches		pounds	inches		pounds
18–25	69·5	25·9	178·1	70·3	25·2	176·7	64·1	26·2	152·8	64·7	23·8	141·5
26–35	69·6	27·2	187·4	70·4	27·2	192·0	64·2	27·9	163·7	64·8	25·9	154·2
36–45	69·7	28·0	193·7	70·1	28·2	196·6	64·3	28·3	166·5	64·7	26·8	159·3
46–55	69·5	28·7	196·7	69·6	28·8	197·9	64·1	28·8	168·0	64·2	27·6	162·0
56–65	68·8	28·7	192·7	69·4	28·7	196·4	63·6	29·8	170·7	63·9	27·6	159·9
	cm		kg	cm		kg	cm		kg	cm		kg
18–25	176·5		81·0	178·6		80·3	162·8		69·5	164·3		64·3
26–35	176·8		85·2	178·8		87·3	163·1		74·4	164·6		70·1
36–45	177·0		88·0	178·1		89·4	163·3		75·7	164·3		72·4
46–55	176·5		89·4	176·8		90·0	162·8		76·4	163·1		73·6
56–65	174·8		87·6	176·3		89·3	161·5		77·6	162·3		72·7
Black	inches		pounds	inches		pounds	inches		pounds	inches		pounds
18–25	69·9	25·3	175·5	69·7	26·3	181·5	64·1	28·8	168·5	64·5	26·7	157·8
26–35	69·9	27·7	192·9	69·2	29·4	198·9	64·7	30·8	183·8	64·6	29·7	176·1
36–45	70·1	28·1	196·0	69·5	28·3	193·8	64·4	31·6	186·9	64·7	30·9	183·9
46–55	69·5	27·3	187·5	69·7	27·5	189·9	64·5	31·8	187·4	64·7	30·5	181·0
56–65	69·2	28·5	194·4	70·4	28·3	198·1	63·9	31·9	185·2	64·0	31·0	179·7
	cm		kg	cm		kg	cm		kg	cm		kg
18–25	177·5		79·8	177·0		82·5	162·8		76·6	163·8		71·7
26–35	177·5		87·7	175·8		90·4	164·3		83·5	164·1		80·0
36–45	178·1		89·1	176·5		88·1	163·6		85·0	164·3		83·6
46–55	176·5		85·2	177·0		86·3	163·8		85·2	164·3		82·3
56–65	175·8		88·4	178·8		90·0	162·3		84·2	162·6		81·7
Hispanic	inches		pounds	inches		pounds	inches		pounds	inches		pounds
18–25	68·0	25·8	169·7	67·6	26·1	170·1	62·6	27·0	150·1	62·8	25·1	140·1
26–35	66·9	27·4	174·7	66·8	26·7	181·9	62·7	28·4	158·2	62·5	27·6	152·8
36–45	67·7	28·0	182·1	67·0	29·2	186·4	62·4	29·2	161·3	62·2	29·0	159·0
46–55	66·8	28·5	181·0	66·7	28·4	179·1	62·0	30·1	164·3	62·1	29·7	162·9
56–65	66·3	27·7	175·6	65·3	30·0	181·3	61·5	30·7	164·8	61·3	29·1	155·2
	cm		kg	cm		kg	cm		kg	cm		kg
18–25	172·7		77·1	171·7		77·3	159·0		68·2	159·5		63·7
26–35	169·9		79·4	169·7		82·7	159·3		71·9	158·8		69·5
36–45	172·0		82·8	170·2		84·7	158·5		73·3	158·0		72·3
46–55	169·7		82·3	169·4		81·4	157·5		74·7	157·7		74·0
56–65	168·4		79·8	165·9		82·4	156·2		74·9	155·7		70·5

Note: while there is agreement on BMI and weight for white men, the Size USA data appear to overestimate male height. For white women the Size USA survey data overestimate height, underestimate weight and consequently strongly underestimate BMI.

freedom in demarcating the borders of such regions in unpopulated areas, the ZCTA codes include additional codes so that water and unpopulated areas are not included in the ZCTAs that correspond to zip codes (http://www.census.gov/geo/ZCTA/zcta.html).

Since height is correlated with socioeconomic status in the United States, as it is everywhere else (Komlos & Kriwy, 2003), those variables are adjusted for as well as the age groupings in all reported analysis. Ethnicity is categorized as White, Black, Hispanic, and Other, but only the records of 'Whites' were considered, because there were too few Blacks in the sample (627 men and 989 women). The white and black populations are not combined because the high levels of persistent residential segregation make the assumption that blacks and whites respond similarly to the influences of local environment unjustified. The other two groups – 'Hispanic' and 'Other' – are excluded from the analysis because of their greatly increased chance of being foreign born. Everyone in the 66+ group was also excluded. The elimination of 18–25-year-old men from the sample was considered due to the possibility that they may not be fully grown; however, they were included as it was found that their presence did not change any results significantly. These limitations leave 1524 men and 2903 women in the sample.

Individual predictors of height and weight

The heights of both men and women exhibit a positive correlation with income and educational attainment (Table 2, Column A). For these results, there was no adjustment for survey site since this regression is for the purpose of illustrating categorical trends in the data (in any case only two of these site factors are significant: women from Glendale, CA, are shorter than average and women from Dallas, TX, have greater BMI). The low R^2s reported in these regressions are the standard order of magnitude in adult populations insofar as most of the variation in final height is genetic and therefore unexplained. Recent increases in height have been slight since the birth cohorts of 1957–66: about 0·5 inches (1·25 cm) for both males and females. The more recent birth cohort 1967–1976 shows a mere 0·3 inch increase for women and none for men. In contrast, heights for European men and women have continued to increase since the 1950s and have now overtaken and exceeded American heights. Australian male heights increased by 2 cm, and female heights by 1 cm between the birth cohorts circa 1955–1970 (Henneberg, 2000). While the age and education effects on the height of men and women are similar, the income coefficient is larger for men than for women, a result not previously observed (Komlos & Baur, 2004). This disparity may reflect the fact that household incomes are reported in this sample. Since the incomes of men are more influential in determining household incomes – due to their higher individual incomes and higher rates of work – the weaker correlation between height and income for women is expected. Alternatively, greater individual height may lead to greater income (Heineck, 2004), which may be a stronger effect for men than for women given their greater mean income.

In contrast to the height data, for which there is slightly greater variance among men, for weight the variance is much greater for women; this is reflected in differences

Table 2. Height coefficients for age, income and education factors in the sample of white men and women under age 65 (A)[a]. Regression results for height as a function of log (population density) (B)[b]

Factor	Men (A)			Men (B)			Women (A)			Women (B)		
	Coef.			Coef.			Coef.			Coef.		
	inches	cm	*t*	inches	cm	*t*	inches	cm	*t*	inches	cm	*t*
Intercept	71·64	181·97	249·8	**71·53**	**181·69**	**149·2**	65·54	166·47	225·8	**65·55**	**166·50**	**353·7**
Log (population density)				**− 0·56**	**− 1·42**	**− 4·3**				− 0·14	− 0·37	− 1·5
Age/birth cohort												
18–25 (1976–1983)	Ref.	—	—	Ref.		—	Ref.	—	—	Ref.		—
26–35 (1967–1976)	− 0·13	− 0·33	− 0·5	0·05	0·13	0·3	− 0·32	− 0·81	− 1·8	− 0·31	− 0·79	− 1·8
36–45 (1957–1966)	− 0·51	− 1·30	− 1·4	− 0·34	− 0·86	− 1·1	**− 0·47**	**− 1·19**	**− 2·0**	**− 0·48**	**− 1·22**	**− 2·1**
46–55 (1947–1956)	**− 1·10**	**− 2·79**	**− 2·9**	**− 0·92**	**− 2·34**	**− 2·8**	**− 0·94**	**− 2·39**	**− 4·2**	**− 0·95**	**− 2·41**	**− 4·3**
56–65 (1937–1946)	**− 1·25**	**− 3·18**	**− 3·2**	**− 1·13**	**− 2·87**	**− 3·2**	**− 1·16**	**− 2·95**	**− 4·6**	**− 1·15**	**− 2·92**	**− 4·7**
Income												
<25K	**− 1·01**	**− 2·57**	**− 4·8**	**− 0·92**	**− 2·34**	**− 4·2**	**− 0·76**	**− 1·93**	**− 2·5**	**− 0·76**	**− 1·93**	**− 2·6**
25–50K	**− 0·95**	**− 2·41**	**− 5·1**	**− 0·91**	**− 2·31**	**− 5·2**	− 0·41	− 1·04	− 1·6	− 0·41	− 1·04	− 1·6
50–75K	**− 0·63**	**− 1·60**	**− 3·7**	**− 0·65**	**− 1·65**	**− 4·2**	− 0·16	− 0·41	− 0·8	− 0·17	− 0·43	− 0·9
75–100K	**− 0·57**	**− 1·45**	**− 3·3**	**− 0·62**	**− 1·57**	**− 3·0**	− 0·06	− 0·15	− 0·2	− 0·06	− 0·15	− 0·3
100K+	Ref.	—	—	Ref.		—	Ref.	—	—	Ref.		—
Education												
<High school	**− 2·07**	**− 5·26**	**− 4·8**	**− 1·83**	**− 4·65**	**− 5·1**	**− 1·87**	**− 4·75**	**− 5·1**	**− 1·83**	**− 4·65**	**− 5·4**
High school	**− 0·83**	**− 2·11**	**− 3·1**	**− 0·77**	**− 1·96**	**− 3·1**	**− 0·59**	**− 1·50**	**− 3·2**	**− 0·58**	**− 1·47**	**− 3·4**
Some college	− 0·4	− 1·02	− 1·8	**− 0·43**	**− 1·09**	**− 2·0**	− 0·22	− 0·56	− 1·9	**− 0·25**	**− 0·64**	**− 2·4**
College	Ref.	—	—	Ref.	—	—	Ref.	—	—	Ref.		—
Postgraduate	− 0·33	− 0·84	− 1·1	− 0·40	− 1·02	− 1·7	0·16	0·41	1·6	0·14	0·36	1·4
Pseudo R^2	0·045			0·058			0·044			0·046		
AIC	33·10			35·09			32·74			34·74		

[a]All factors are reported with respect to the baseline individual: an 18–25-year-old, college-educated individual with income over US$100,000. [b]The intercept is given at the sample mean log (population density) of 3·17 (1500 per square mile). Note the improvement in pseudo R^2 for male height, but not for female height, with respect to the initial factor analysis in (A). Bold indicates that the coefficient is significant at the 5% level. For number of observations see Table 3. Pseudo R^2 (1 − Res./null deviance).

between education, income and age groups (Table 3, Column A). This might be explained by such physiological factors as the percentage of body fat among women differing from that of men. The increase in BMI for both men and women with age is found to be large and monotonic up to the 46–55 age group: men's BMI increases by 3·9 from the 18–25 group to the 46–55 group while women's BMI increases by 4·6 over that range. Men's BMI shows no significant correlation with income; however women's BMI has two levels, with the three income groups earning below US$75,000/year having 1·4 greater BMI than the two above that threshold. With respect to education, men again show no significant correlations with BMI. Among women, generally greater education is correlated with lower BMI, as in other samples (Komlos & Baur, 2004).

Local environmental predictors: population density

Considering quantitative measures of local conditions, population density is an appealing metric because density differences can indicate differences in how people live in different communities. Higher density communities are likely to have more immediately available health services and people may spend less time in cars, but may also have lower rates of voluntary exercise, higher stress and greater environmental pollution. Since the densities in the sample range from 10 persons/mile^2 to 100,000 persons/mile^2, a log transformation on population density is used (Fig. 1).

A few examples are useful in assessing this distribution. The densest zip code in the sample – '10009' on the Lower East Side of Manhattan – had a density of just over 100,000 per square mile. The zip code '02138' in Cambridge, MA, the urban residential and commercial area including Harvard University, has a density of 12,500 per square mile. A very wealthy suburban area outside of Los Angeles, '90210', has a density of 2300 per square mile. The code '66049', in Lawrence, KS, has a density of 410 per square mile. These four locations correspond to log(population densities) of 11·5, 9·4, 7·7 and 6·0. The population density distribution of these data is similar to that of the US population as a whole, although the mean is almost twice as high, because the rural areas (<300 persons/mile^2) are under-sampled and moderate density areas (300–3000 persons/mile^2) are over-sampled (US Census, 2000; US Census Bureau, 2003). The mean log(population density) for the sample is 7·31 (1500 per square mile) versus 6·77 (875 per square mile) for the US population. The number of observations suffices to analyse densities between 30 and 100,000 persons per square mile. The overall population density of the United States is 80 persons per square mile versus 320 in the European Union; however, these figures include all land area (including uninhabitable land area) and are not directly comparable to the figures obtained for zip codes – limited to inhabited areas.

Height is more strongly correlated with population density for men than it is for women. For men, after adjustment for individual income and education, an increase in population density by a factor of 10 corresponds to a change in height of $-0{\cdot}56$ ($\pm 0{\cdot}25$, $p=0{\cdot}00001$) inches ($-1{\cdot}4$ cm) (Table 2, Column B). It is true that height was measured to the nearest half-inch, but the estimated mean is still unbiased, and even if the variance is inflated the t test of significance is still valid. The relationship

Table 3. BMI coefficients for age, income and education factors in the sample of white men and women under age 65 (A). Regression results for BMI as a function of log (population density) (B)

	Men (A)			Men (B)			Women (A)			Women (B)		
Factor	*n*	Coef.	*t*	*n*	Coef.	*t*	*n*	Coef.	*t*	*n*	Coef.	*t*
Intercept	**1524**	**24·2**	**52·9**	**1524**	**24·61**	52·5	**2903**	**21·8**	**72·8**	**2903**	**21·81**	**52·4**
Log (population density)				1524	− 0·32	− 1·7				2903	− 0·33	− 1·2
Age/birth cohort												
18–25 (1976–1983)	469	Ref.	—	469	Ref.	—	670	Ref.	—	670	Ref.	—
26–35 (1967–1976)	**280**	**2·1**	**5·0**	**280**	**2·25**	**5·3**	**611**	**2·7**	**5·8**	**611**	**2·72**	**5·8**
36–45 (1957–1966)	**335**	**3·2**	**5·5**	**335**	**3·29**	**5·5**	**636**	**3·7**	**12·5**	**636**	**3·71**	**13·9**
46–55 (1947–1956)	**270**	**3·9**	**8·0**	**270**	**3·98**	**8·3**	**620**	**4·6**	**14·4**	**620**	**4·56**	**14·9**
56–65 (1937–1946)	**173**	**3·8**	**10·6**	**173**	**3·90**	**10·6**	**366**	**4·4**	**12·8**	**366**	**4·37**	**13·5**
Income												
<25K	458	− 0·0	− 0·0	458	0·04	0·1	**614**	**1·3**	**3·9**	**614**	**1·38**	**3·9**
25–50K	321	0·7	1·8	321	0·75	1·8	**764**	**1·4**	**3·7**	**764**	**1·43**	**3·9**
50–75K	289	0·5	1·5	289	0·50	1·4	**645**	**1·4**	**3·7**	**645**	**1·36**	**3·5**
75–100K	201	− 0·2	− 0·7	201	− 0·25	− 0·7	431	− 0·1	− 0·3	431	− 0·13	− 0·4
100K+	254	Ref.	—	254	Ref.	—	449	Ref.	—	449	Ref.	—
Education												
<High school	53	0·5	0·7	53	0·64	0·9	**56**	**3·0**	**6·2**	**56**	**3·12**	**5·9**
High school	298	0·2	0·4	298	0·19	0·5	**442**	**0·6**	**2·4**	**442**	**0·61**	**2·4**
Some college	485	0·5	1·2	485	0·46	1·1	**986**	**1·3**	**5·1**	**986**	**1·22**	**5·4**
College	412	Ref.	—	412	Ref.	—	968	Ref.	—	968	Ref.	—
Postgraduate	275	− 0·3	− 0·9	275	− 0·35	− 1·0	451	− 0·2	− 1·2	451	− 0·26	− 1·2
Pseudo R^2	0·090			0·092			0·085			0·086		
AIC	34·03			36·02			34·36			36·36		

The intercept is given at the sample mean log(population density) of 3·17. The changes in pseudo R^2 versus regression (A) are very small; population density provides little additional explanatory value over the income, education and age factors for BMI. All factors are reported with respect to the baseline individual: an 18–25-year-old college-educated individual with income over US$100,000.
For number of observations see Table 3.
Pseudo R^2 (1 − res./null deviance).

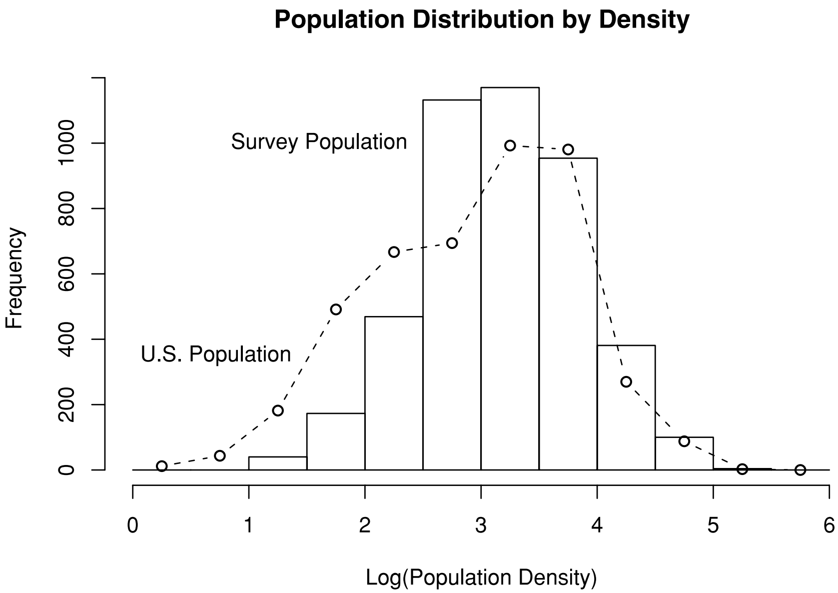

Fig. 1. Distribution of log (population density) in the Sizing Survey sub-sample. Superimposed on this histogram is the distribution for the entire US population, according to the US Census (normalized to the Sizing Survey sample size). The Sizing Survey over-counts moderate densities (300–3000 per square mile) at the expense of low densities, which are poorly represented in the sample. Only results between log (population density) 1·5 and 5 can be considered as robust. The mean log (population density) for the sample is 3·17 (1500 per square mile) versus 2·94 (875 per square mile) for the US population. Source: US Census (2000) (Summary File 1).

between height and population density is linear over the range where there are sufficient data to estimate a mean height as a function of density (Fig. 2). The difference in mean height over this range (between circa 100 and 30,000 persons/mile^2) is about 1·75 inches (4·5 cm) – a very large difference indeed. With direct adjustment for cluster site the magnitude of the correlation is reduced from $-0{\cdot}56$ to $-0{\cdot}36$ ($\pm 0{\cdot}34$) with $p=0{\cdot}04$ and remains linear over the full range of data after adjustment, indicating that within individual clusters there is still a height gradient.

Information is not provided on the length of time the surveyed population had been resident within the respective Zip Code Tabulation Area (ZCTA). In this present age of geographic mobility, although it is true that the surveyed population probably did not grow up in the same ZCTA, it is possible that to some extent the movement was across roughly similar population density areas.

Using a similar linear model for white women, an increase in population density by a factor of 10 is correlated to a reduction in height by 0·14 ($\pm 0{\cdot}19$, $p=0{\cdot}13$) inches ($-0{\cdot}4$ cm) (Table 2, Column B). However, the relationship between height and population density for women appears to be two-level (above and below a density of 10,000 per square mile) rather than linear (Fig. 3). The heights in the low- and medium-density regime exceed those in at high densities by 0·6 inches ($p<0{\cdot}001$)

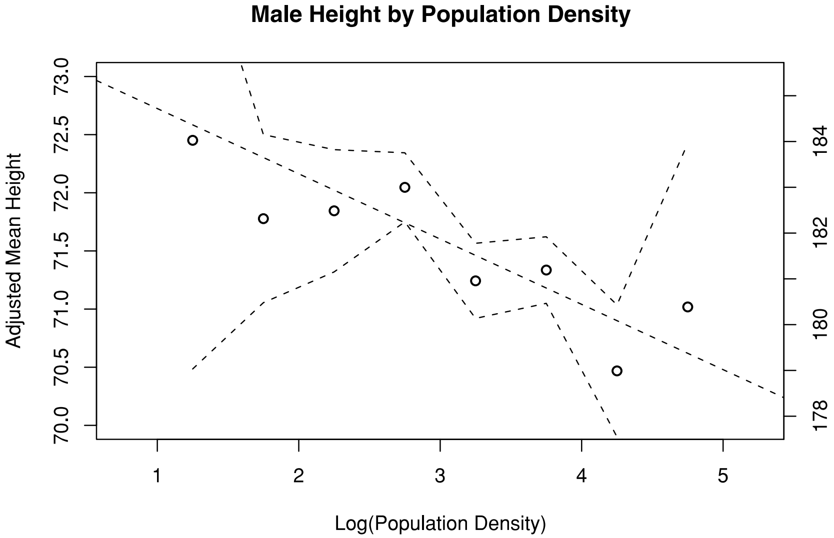

Fig. 2. Male height, adjusted for age, education and income categories, as a function of log (population density). Each point is the mean adjusted height for densities between adjacent integer values of log (population density). The values are set to the reference values of the regression in Table 4. The 95% confidence interval for each point is depicted with dashed lines.

(1·5 cm). Adjusting for cluster site eliminates all dependence of height on population density for women (p=0·95).

Similar results are found for the correlation of BMI with local population density for both men and women; however they are not statistically significant. A factor of 10 increase in population density is found to be correlated to a decrease in BMI of 0·32 for men (p=0·09) and 0·33 women (p=0·23) (Table 3, Column B). However, the plots of mean BMI with respect to density indicate a non-linear relationship for men (Fig. 4) while they plausibly support an approximately linear model for women holding income and education constant (Fig. 5). The magnitude of the correlation increases (and reaches the 95% confidence threshold for men) after adjustment for cluster: − 0·24 ± 0·23 (p=0·04) for men and − 0·26 ± 0·19 (p=0·008) for women. The magnitude of this increase is not itself significant, but such an increase would indicate that weights are correlated with population density more strongly within each cluster than they are across all clusters. This perhaps indicates that sorting, rather than inherent effects of particular densities, is the cause for such differences: people who have higher BMIs are choosing to live in lower density areas within the region near their employment. Since adjustment for education and income may not be perfect, we may be observing that within a region, the higher density areas are on average wealthier. Since weight is negatively correlated with income for women, an imperfect adjustment would create just this type of in-cluster gradient.

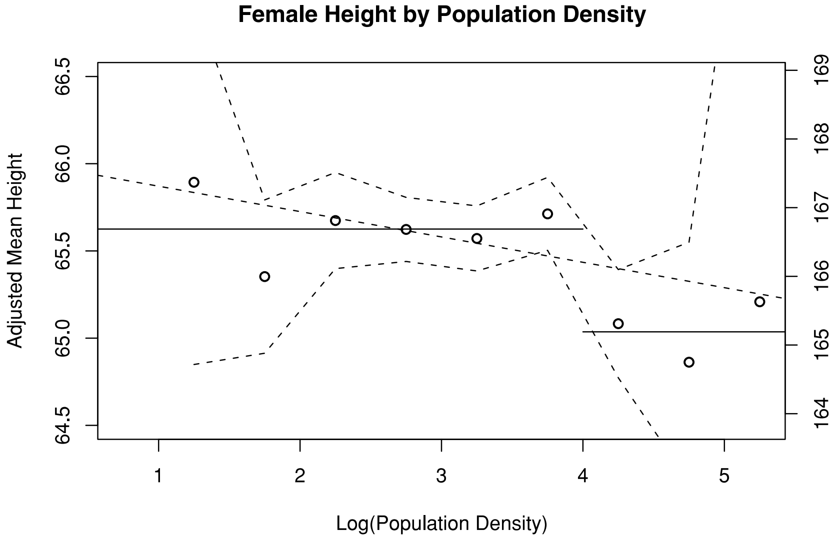

Fig. 3. Female height, adjusted for age, education and income categories, as a function of log (population density). Each point is the mean adjusted height for densities between adjacent integer values of log (population density). The values are set to the reference values of the regression in Table 4. The 95% confidence interval for each point is depicted with dashed lines. The regression line described in Table 4 is plotted as a dashed line because there is evidence from the plot that the dominant effect is not linear, but two-levelled. Plotted as two solid lines at 65·62 inches (166·7 cm) and 65·03 inches (165·2 cm) are the mean values for women living at population densities above and below 10,000 per square mile.

Foreign born

The sample does allow the direct identification of the ‘foreign born’; however, since it does have ‘Hispanic’ and ‘Other’ categories, the foreign born are probably largely eliminated through the study’s exclusion criterion. Given that only 14% of current US foreign-born individuals originated in Europe while 53% originated in Latin America and 25% in Asia and that foreign-born individuals comprise 12% of the US population (US Census Current Population Survey, 2003), most immigrants should be identified as ‘Hispanic’ or ‘Other’. Most importantly, the individuals likely to be foreign born are eliminated through the ‘Hispanic’ and ‘Other’ categories at rates expected if such categories were eliminating most of the foreign born (Fig. 6). The number of individuals eliminated by this method as a function of fraction foreign born was found to be both uniformly greater than, and to scale with the fraction of individuals who are foreign born in those districts. However, while this indicates that most of the foreign born are probably already being eliminated from the sample – consistent with the fact that most of the foreign-born population in the US is non-white – it is not certain that foreign-born individuals have been entirely excluded.

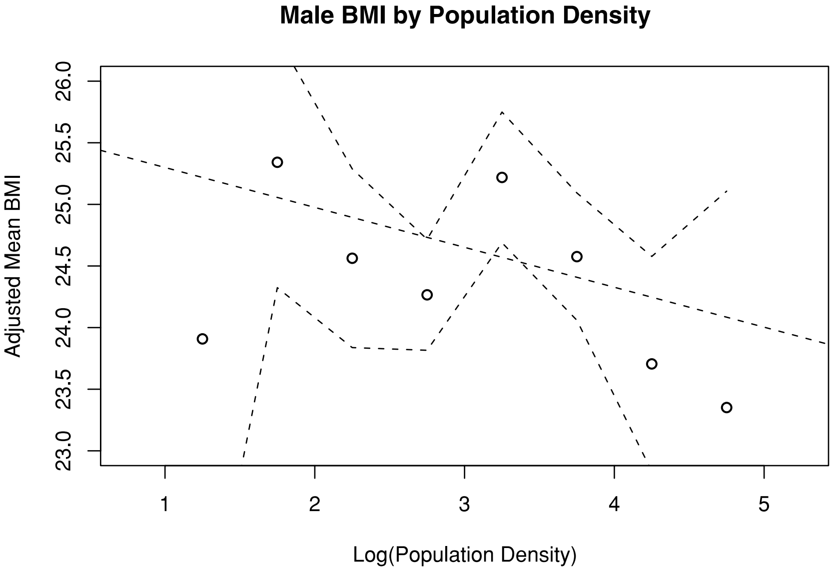

Fig. 4. Male BMI, adjusted for age, education and income categories, as a function of log (population density). Each point is the mean adjusted height for densities between adjacent integer values of log (population density). The values are set to the reference values of the regression in Table 5. The linear model (dashed line) described in Table 5 does not fit the data. The 95% confidence interval for each point is depicted with dashed lines.

Hence, whether the inability to identify foreign-born individuals directly in the sample could have led to the large observed negative relationship between height and population density among white men is explored further. For each factor of 10 in a particular density range the fraction foreign born in the zip code area was found (US Census Bureau, 2003). Among these 0·14 were of European origin (US Census, 2000). Multiplying these two fractions yields an estimate of the foreign-born white men in a zip code area in a particular density range. That number of individuals was then eliminated from the sample; the total number of observations thus eliminated is 24 out of 1525 – about 1·6% of the data. In order to bias the results against our hypothesis the shortest individuals within each density range in proportion to the estimated number of foreign-born individuals in a zip code area were removed. The regression slope is thereby slightly reduced (in absolute) magnitude to − 0·4 inches (− 1·1 cm) but remains strongly significant. This still implies a gap of about 1·5 inches (3·75 cm) over the density range. The real influence of the foreign born would be smaller than this, as the maximum effect is estimated. A similar exercise was done for white women under the two-level model, eliminating 60 individuals (38 below the cut-off density of 8100 and 22 above that level). The height difference is reduced slightly from 0·59 to 0·46 inches (1·2 cm) ($p=0·001$). Thus, even under these adverse assumptions, the negative association between physical stature and population density of residence remains substantial and strongly significant.

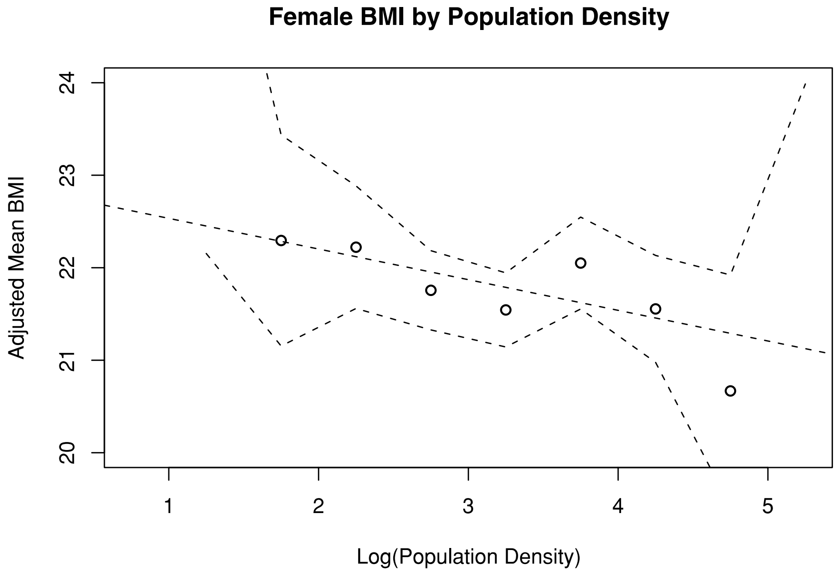

Fig. 5. Female BMI, adjusted for age, education and income categories, as a function of log (population density). Each point is the mean adjusted height for densities between adjacent integer values of log (population density). The values are set to the reference values of the regression in Table 5. The linear model (dashed line) described in Table 5 is mostly followed by the data, but the higher density groups seem not to follow this model. The 95% confidence interval for each point is depicted with dashed lines.

However, if the fraction of foreign-born individuals in a zip code is included in the regressions, the sizes of all population density correlations are reduced to insignificance. The authors believe that such an adjustment is inappropriate, for two reasons. First, as the model uses individual factors such as personal income and personal environment expressed through population density to predict height, including the number of foreign-born individuals in a person's zip code is conceptually problematic. Second, the fraction foreign born in a zip code is very highly correlated with population density (r=0·63), so the regression becomes unstable when both variables are included.

Local environmental predictors: economic conditions

Marginal positive correlations between male height and local economic conditions as measured by poverty rate, unemployment rate and median income were examined. Insofar as these results were not significant after adjustment for population density, little evidence was found that local economic conditions considered separately from personal income have any predictive power regarding height or BMI (Table 4). There may be a significant negative association between male BMI and local poverty rate, even after adjusting for individual income and education, but this is only observable when the non-significant unemployment rate and median income are included in the regression, so this result should not be given much credence.

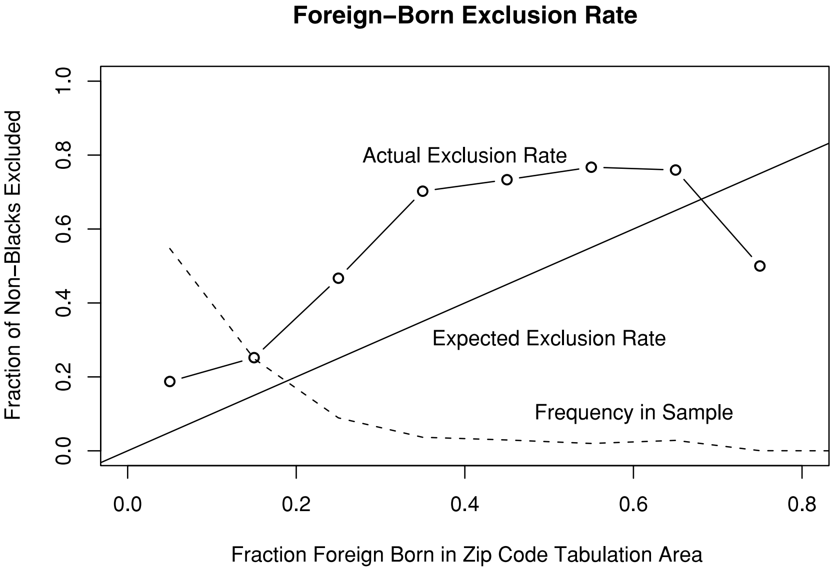

Fig. 6. Rate of exclusion from the non-black sample as a function of fraction foreign born in zip code tabulation area. A straight line is plotted at the rate of rejection expected if all individuals in the 'Hispanic' and 'Other' categories were foreign born. The rate of rejection is higher than this baseline (the final point corresponds to just four individuals, of which two are eliminated) and scales properly as a function of fraction foreign born in zip code. This suggests that most of the foreign-born individuals are successfully rejected. The dashed line indicates the distribution of individuals in the final sample as a function of fraction foreign born.

Conclusion

The survey under consideration is not a random sample of the US population. It under-sampled those living at low population densities (under 300 people per square mile) and over-sampled those living in moderate densities (between 300 and 3000 people per square mile). The bias is linked to its strategy to collect samples mainly at shopping malls. This had the consequence of obtaining an anthropometrically biased sample, especially among women. While young men (18–25) in the sample are 0·8 inches taller than the national average, young women in the sample are 11 pounds lighter and 0·6 inches taller than the national average. Evidently, the individuals shopping in the malls survey are more active and hence less likely to be obese than the general population.

Nonetheless, the strong negative correlation between white male height and population density of residence is striking. The difference in men's height between low- and high-density residences is quite substantial: circa 1·5–1·75 inches (3·75–4 cm), after controlling for own income and education. This relationship is also negative among white women, though smaller: about 0·5 inches (1·25 cm), and significant only between those living at densities greater and less than about 10,000 persons per square mile. It is unlikely that these results are influenced by the study's limited ability to

Table 4. Regression results for white male and female height and BMI as a function of local economic conditions

	Men			Women		
Height	inches	cm	*t* value	inches	cm	*t* value
(Intercept)	74·18	188·43	11·93	67·65	171·82	26·30
log (population density)	− **0·54**	− **1·37**	**3·61**	− 0·12	− 0·30	− 1·20
log (median household income)	− 0·21	− 0·53	− 0·16	− 0·33	− 0·84	− 0·59
Unemployment rate	− 0·05	− 0·12	− 1·87	0·01	0·02	0·31
Poverty rate	0·02	0·06	0·80	− 0·02	− 0·05	− 0·97
BMI	index		*t* value	Index		*t* value
(Intercept)	34·50		5·23	36·88		4·74
log (population density)	− 0·20		− 0·92	− 0·17		− 0·55
log (median household income)	− 1·82		− 1·30	− 2·93		− 1·70
Unemployment rate	0·05		0·86	− 0·07		− 1·39
Poverty rate	− **0·10**		**2·46**	− 0·03		− 0·70

There are no significant correlations between these variables after adjustment for individual income, education and age (included in regression, but not depicted here) except for a slight negative association between male BMI and the poverty rate. Median household income is in dollars; unemployment rate and poverty rate are in per cent. Intercept term is at all shown variables equal to zero and should not be directly interpreted.

screen out foreign-born individuals, who are more heavily concentrated in urban areas. This is the case insofar as the negative association held up even after reasonable adjustments were made for the foreign born in a particular zip code area. The fact that this is a convenience (rather than a random) sample should not affect the main results because many explanatory variables in the regressions were controlled for. There is no reasonable reason why taller people would have been relatively over-sampled at lower population densities. Nonetheless there are many omitted variables such as the size and socioeconomic status of the parents, birthplace of the respondents, and number of siblings in the family.

That height declines with increasing community size in the US has been observed previously, even if the observed gradient was only 0·5 inches (1·25 cm) and the sample of soldiers did not exclude the foreign born (Karpinos, 1958, p. 308). Moreover, it has been often observed that New England – among the most densely populated section of the US – has had the shortest white population ever since the early 19th century (Karpinos, 1958; US Department of Health and Human Services, 1981, p. 33). However, the finding of urban disadvantage is contrary to many studies, though these have generally not held income and education constant (US Department of Health and Human Services, 1981, p. 31). Even if the relationship between health and urban life is ambiguous (Galea *et al.*, 2005), in many European contexts the relationship between height and population density tends to be positive since about the turn of the 20th century. After about 1900, with the rise of refrigerated transport, and the improvements in sanitation, the advantages of urban communities with higher income groups and superior access to medical care tended to outweigh the higher disease loads and higher food prices (Eveleth & Tanner, 1990, p. 202). The urban height advantage in contemporary Germany, for instance, is substantial only among males and is strongest among the lower income and education groups, above all in the former GDR (Komlos & Kriwy, 2002, 2003). On average German men in cities are 1·7 cm taller than rural (village) residents. Furthermore, recruits from Oslo and its surroundings are tallest in Norway, (Sunder, 2003), and according to the French Decennial Health Survey 2003, male heights in Paris are 1 cm above the average (personal communication from Alain Paraponaris), but in the Netherlands, urbanization does not have a major impact on heights (Fredriks, 2004, p. 30).

The negative correlations observed are robust between suburban or small town densities and those found in cities. The sample sizes are too small at truly rural densities to gauge the trend there. Nonetheless, this study's findings might indicate the influence of urban disamenities on height, such as pollution (poorer quality of air) and the quicker spread of contagious diseases. There are also negative neighbourhood externalities in 'inner city' slums. Moreover, urban lifestyles associated with a fast-food consuming culture, more stressful living – 'life in the fast lane', a fast-paced life that might lead to the disregard of children's needs, and thereby affect children's growth (Phelps, 2003, p. 104; Okunade & Karakus, 2003). In other words, there might be net health effects of urban living compared with small towns and suburbs even after controlling for individual social status. While urban areas have better access to health care facilities, the residents might not be as keen on using those opportunities. It might also be that the birth rate among urban poor is greater than among those living elsewhere.

Ideally it would have been good to have information on the income of the parents of the individuals in the sample, instead of their current income. With upward social mobility the adults in the sample might well be shorter than their current income would indicate. Moreover, it was not possible to control for heterogeneity in the ethnic background of the white population or for the geographic distribution of health care quality. Because terminal height can only be influenced by environmental conditions prior to reaching adulthood, spatial movement of individuals between childhood or adolescence and adulthood would diminish differences caused by disparate childhood conditions (if the movement is to locations that are substantially different from the location of childhood). However, the same movement of individuals could lead to the creation of geographic height gradients if a mechanism exists by which individuals sort into certain types of areas. While the US census does not find any difference between absolute movement rates for men and women that might explain the strong observed population density correlation for men (US Census Bureau, 2003; Annual Social and Economic Supplement, tabulated at http://www.infoplease.com/ipa/A0922200.html), it may be that the choice of movement destination rather than the number of movements is at issue. There is greater socioeconomic mobility among women in the United States; daughters' incomes are less strongly correlated with their fathers' incomes than sons' incomes are (Peters, 1992). Consequently, two hypotheses emerge for the disparity between men and women. First, it may be that men carry forward more of their childhood differences due to less mobility-induced mixing, which presumes that living at low densities is, in fact, beneficial. Second, it may be that men are choosing the destination of their movement based on factors that are correlated with their own height – a sorting hypothesis. In other words, men either select themselves in such a way that taller men choose to live in low-density environments, or there are spillover effects that have a negative impact on the height of men in high-density areas to the extent that current residence type correlates positively with residence type during childhood. Hence, much more work needs to be done along these lines before there can be more certitude pertaining to the finding that heights correlate negatively with population density. Nonetheless, the results are intriguing and worth pursuing.

Body mass index also varies somewhat by residence type, with people living in low densities weighing more than those who live in high-density areas, a disparity that seems to be stronger within individual communities than it is across the country as a whole. Some evidence was found that population density is more strongly correlated with BMI within a community than across communities, with larger individuals living in lower densities within their area regardless of that area's absolute density. The fact that there is a simultaneous negative correlation with BMI complicates the process of linking the results pertaining to height directly to claims about biological welfare.

The hypothesis that higher densities lead to lower height would indicate that American suburbs and moderately sized towns – the low end of the sampled density range – are providing the best mix of benefits for biological welfare: easy access to medical care and few of the negative environmental conditions found in urban areas. To confirm such a conclusion, it would be necessary to show a link between these densities and health, perhaps by using a sample of children. Supporting the sorting hypothesis is the fact that height has been shown to predict personal economic success (Persico *et al.*, 2004), and more strongly for men than for women (Heineck, 2004). If those who are successful are preferentially moving to suburbs with low density, this might create the height gradient observed, though adjusting for individual income would compensate for such a difference if the income data in the sample were accurate. Given the small number of income categories and the fact that income is self-reported, this study's income adjustment might well be incomplete. Moreover, a serious potential problem is that it was not possible to adjust for cost-of-living indexes at the local level, implying that nominal income rather than real income was actually being used in the analysis. Finally, a psychological explanation might be applicable as well, if among those with equal income, individuals who are physically larger make choices about financial allocation regarding the size of their residence that push them towards lower density areas.

In sum, while the results should be interpreted with caution as rural densities are under-sampled in this survey, it might well be worthwhile to entertain the hypothesis further that the fact that the US has been lagging behind Western and Northern Europeans in physical stature is mostly a large-city phenomenon, at least among men. Consider that young adult, rural, low-status West German men are about 177·4 cm tall, the same as their US counterparts (with a high-school education) living at the average population density of 1500 persons/mile^2. Similarly, while upper-status rural West German men enjoy a 3·6 cm height advantage, US upper-status men (with a college education, earning US$100K+) living at average densities enjoy a 4·3 cm height advantage over lower-status men. Hence, there is not a consequential difference in height between German and American men (either low or upper status) living at lower than average population densities. However, German men living at high densities are 1·7 cm taller than those living at low densities, while American are nearly 2·0 cm shorter. Hence, the gap between the urban German and US males widens to some 3·7 cm (1·5 inches). This pattern does not quite hold up among females, however, insofar as German women are taller than American women at both low and high population densities (by circa 2·0 cm and 3·3 cm respectively). This might be associated with the higher levels of poverty in the US than in other OECD countries. Defining poverty as having a household income 50% below the median in the population implies that the US has the highest rate of poverty of children in OECD countries. While children in Nordic countries have a poverty rate of about 4%, the US has a rate in excess of 20%, twice as high as those obtained in Holland and Germany (UNICEF, 2005, p. 5). Admittedly, it was not possible to link these patterns satisfactorily to the patterns outlined above; nonetheless, the results reported here do point to the need for further research considering the relationship between living environment and lifestyle as a means to explain why Americans have fallen behind Western and Northern Europeans in height since the 1960s, but surged well ahead in weight since the 1980s.

Acknowledgments

The authors appreciate comments from David Allison, Barry Bogin, Marek Brabec, Bruce Bradtmiller, Noël Cameron, Francesco Cinnirella, Dora Costa, Joan Costa-Font, Timothy Cuff, Soumyananda Dinda, Irvin Emanuel, Andrew Foster, Ralf Futselaar, Michael Haines, Bernard Harris, Maciej Henneberg, Michale Hermanussen, Laurent Heyberger, Richard Jantz, Ichiro Kawachi, Charles Kenny, Stephan Klasen, Martin Linsi, Reynaldo Martorell, Udaya S. Mishra, Albert A. Okunade, Alain Paraponaris, Andrew Postlewaite, Virginia Rauh, Friedrich W Rösing, Tom Samaras, T. Paul Schultz, Marco Sunder, Walter Willett and Barbara L. Wolfe on earlier versions of the paper. Any remaining errors are the authors' own.

References

Costa, D. L. (1993) Height, weight, wartime stress, and older age mortality: evidence from the union army records. *Explorations in Economic History* **30**, 424–449.

Decker, S. L. & Remler, D. (2004) *How Much Might Universal Health Insurance Reduce Socioeconomic Disparities in Health? A Comparison of the US and Canada*. NBER Working Paper No. W10715, August.

Eveleth, P. B. & Tanner, J. M. (1990) *Worldwide Variation in Human Growth*. 2nd edn. Cambridge University Press, Cambridge.

Floud, R. (1992) Anthropometric measures of nutritional status in industrialized societies: Europe and North America since 1750. In Osmani, S. R. (ed.) *Nutrition and Poverty*. Clarendon Press, Oxford, pp. 219–242.

Fredriks, A. M. (2004) *Growth Diagrams, 1997. Fourth Dutch Nation-wide Survey*. Author's publication, Amsterdam.

Fredriks, A. M., van Buuren, S., Burgmeijer, R. J., Meulmeester, J. F., Beuker, R. J., Brugman, E., Roede, M. J., Verloove-Vanhorick, S. P. & Wit, J. M. (2000) Continuing positive secular growth change in the Netherlands 1955–1997. *Pediatric Research* **47**, 316–323.

Galea, S., Freudenberg, N. & Vlahov, D. (2005) Cities and population health. *Social Science and Medicine* **60**, 1017–1033.

Heineck, G. (2005) Up in the skies? The relationship between height and earnings in Germany. *Review of Labour Economics and Industrial Relations* **19**(3), 469–489.

Henneberg, M. (2000) Possible causes of secular trends in body size – lessons from the Southern Hemisphere. In Schultz, M. *et al.* (eds) *Homo-unsere Herkunft und Zukunft. Proceedings. 4. Kongress der Gesellschaft für Anthropologie, Potsdam 25. bis 28. September 2000.* Cuvillier Verlag, Göttingen, pp. 234–237.

Karpinos, B. D. (1958) Height and weight of selective service registrants processed for military service during World War II. *Human Biology* **40**, 292–321.

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

Komlos, J. & Baten, J. (1998) *The Biological Standard of Living in Comparative Perspectives.* Proceedings of a Conference Held in Munich, January 18–23, 1997. Franz Steiner Verlag, Stuttgart.

Komlos, J. & Baur, M. (2004) From the tallest to (one of) the fattest: the enigmatic fate of the size of the American population in the twentieth century. *Economics and Human Biology* **2**, 57–74.

Komlos, J. & Cuff, T. (eds) (1998) *Classics in Anthropometric History*. Scripta Mercaturae Verlag, St Katherinen.

Komlos, J. & Kriwy, P. (2002) Social status and adult heights in the two Germanies. *Annals of Human Biology* **29**, 641–648.

Komlos, J. & Kriwy, P. (2003) The biological standard of living in the two Germanies. *German Economic Review* **4**, 493–507.

Komlos, J., Smith, P. & Bogin, B. (2004) Obesity and the rate of time preference: is there a connection? *Journal of Biosocial Science* **36**, 209–219.

Okunade, A. A. & Karakus, M. (2003) Mortality from breast carcinoma among US women: The role and implications of socio-economics, heterogeneous insurance, screening mammography and geography. *Health Care Management Science* **6**, 237–248.

Persico, N., Postlewaite, A. & Silverman, D. (2004) The effect of adolescent experience on labor market outcomes: the case of height. *Journal of Political Economy* **112**, 1019–1053.

Peters, H. E. (1992) Patterns of intergenerational mobility in income and earnings. *Review of Economics and Statistics* **74**, 456–466.

Phelps, C. E. (2003) *Health Economics*. 3rd edn. Addison Wesley, Boston.

SizeUSA CD-ROM http://www.sizeusa.tc2.com. Jim Lovejoy, jlovejo@tc2.com; or Karen Davis, kdavis@tc2.com, $[TC]^2$ – Turning Research Into Reality, 211 Gregson Drive, Cary, NC 27511, USA; tel: 919–380–2184 or (919) 380–2177; fax - (919) 380–2181.

Sunder, M. (2003) The making of giants in a welfare state: the Norwegian experience in the 20th century. *Economics and Human Biology* **1**, 267–276.

Waaler, H. (1984) Height, weight, and mortality: The Norwegian experience. *Acta Medica Scandicana* (Supplement) **679**, 1–56.

UNICEF (2005) *Child Poverty in Rich Countries 2005*. Innocenti Report Card No. 6. UNICEF, Innocenti Research Centre, Florence.

United Nations Development Programme (2002) *Human Development Report 2002*. Oxford University Press, New York.

US Census Bureau (2003) *Current Population Survey, Annual Social and Economic Supplement: Foreign-Born Population by World Region of Birth, US Citizenship Status, and Year of Entry: 2003.* http://www.census.gov/population/socdemo/foreign/ppl-174/tab02–06.pdf.

US Census (2000) *Summary File 1*. http://www.census.gov/Press-Release/www/2002/sumfile1.html.

US Department of Health and Human Services (1981) *Height and Weight of Adults Ages 18–74 Years by Socioeconomic and Geographic Variables United States.* National Center for Health Statistics, Series 11, No. 224. PHS 81–1674.