

# Health, Height, Height Shrinkage and SES at Older Ages: Evidence from China

By Wei Huang, Xiaoyan Lei, Geert Ridder, John Strauss and Yaohui Zhao\*

*In this paper we build on the literature that examines associations between height and health outcomes of the elderly. We investigate the associations of height shrinkage at older ages with socio-economic status, finding that height shrinkage for both men and women is negatively associated with better schooling, current urban residence and household per capita expenditures. We then investigate the relationships between pre-shrinkage height, height shrinkage and a rich set of health outcomes of older respondents, finding that height shrinkage is positively associated with poor health outcomes across a variety of outcomes, being especially strong for cognition outcomes. (JEL D1, I12, J13)*

\* Huang: Department of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138 (email: [huangweipku@gmail.com](mailto:huangweipku@gmail.com)); Lei: China Center for Economic Research, National School of Development, Peking University, Beijing 100871, China (email: [xylei@ccer.edu.cn](mailto:xylei@ccer.edu.cn)); Ridder: Department of Economics, University of Southern California, 300 Kaprielian Hall, Los Angeles, CA 90089 (email: [ridder@usc.edu](mailto:ridder@usc.edu)); Strauss: Department of Economics, University of Southern California, 300 Kaprielian Hall, Los Angeles, CA 90089 (email: [jstrauss@usc.edu](mailto:jstrauss@usc.edu)); Zhao: China Center for Economic Research, National School of Development, Peking University, Beijing 100871, China (email: [yhzhao@ccer.edu.cn](mailto:yhzhao@ccer.edu.cn)). We thank two referees, Janet Currie, David Cutler, Richard Easterlin, Richard Freeman, Paul Frijters, Anastasia Gage, Amanda Kowalski, T. Paul Schultz, Yiqing Xu and Yi Zeng. We also thank participants at the 32<sup>nd</sup> Conference for Australian Health Economists, the CCER Academic Conference, the China Economic Annual, the CCER Labor Workshop, the Northeast Universities Development Economics Conference and the USC Development Economics Workshop. We are responsible for all remaining errors and omissions. This research is supported by grants from the National Institute of Aging, the U.S. National Science Foundation, the Natural Science Foundation of China, the Fogarty International Center, the World Bank, and the Peking University–Morgan Stanley Scholarship.

The rapid aging of population places health at older ages among the top public health priorities in recent years as the fraction of the population that is elderly has been rising. In countries such as China, rapid aging has occurred at much lower levels of national income and worse health conditions than was the case in industrial countries. The elderly in such countries were children when economic development and health conditions were far worse than today and their health as adults is likely to have been affected by such past conditions, more so than the elderly in industrial countries.

The effects of early-life health and environment on cognitive function, health, wellbeing and mortality have been documented by researchers across a range of disciplines, using data from many countries over the world (Elo and Preston, 1992; Barker, 1994; Nystrom Peck and Lundberg, 1995; Godfrey and Barker, 2000; Finch and Crimmins, 2004; Case et al., 2005; Case and Paxson, 2008b, 2009, 2010a, b; Alderman and Behrman, 2006; Zeng et al., 2007; Van den Berg, 2006; Smith, 2009; Huang and Elo, 2009, Almond and Currie, 2011).

There are several literatures that have used height to proxy for past health. In the large historical literature, adult height is taken to be an indicator of population health (Fogel, 1986, 2004; Steckel, 1995, 2009). The nutrition literature long ago established that child height is a very good summary measure of overall health of children (e.g. Martorell and Habicht, 1986). Adult height while reflective of the adolescent growth spurt, is also highly correlated with height during childhood.

A strong association has been found to exist among the elderly between measured height and cognitive ability including dementia, self-reported health, illness status and measures of depression (eg. Abbott et al., 1998; Schnaider Beerli et al., 2005; Case and Paxson, 2008a; Case and Paxson, 2010a; Deaton and Arora, 2009; Heineck, 2008; Maurer, 2010; Smith et al., 2012). Most of this literature is from industrial countries (Maurer, 2010 and Smith et al., 2012, are exceptions). The exact mechanisms are not completely known, as these studies are not structural, nor causal. Some of the pathway is likely to be through better health during the fetal period and childhood and prime-aged adulthood (e.g. Case and Paxson, 2008b or Smith, 2009), but other pathways exist as well, such as taller

people having more schooling and consequently making better health behavior decisions (e.g. Cutler and Lleras-Muney, 2010).

However, older people suffer height shrinkage during aging. Aging is associated with several physiological and biological changes, including body composition, such as an increase in body fat, a decrease in lean body mass and bone mass. Through such mechanisms as certain kinds of arthritis (such as ankylosing spondylitis), inflammation of spine joints, herniated disks, or kyphosis, these changes can lead to vertebral deformity, which can contribute to a reduction in height, or shrinkage (Kwok et al., 2002; Prothro and Rosenbloom, 1993; Roubenoff and Wilson, 1993). One health condition that can influence shrinkage directly and through many of these other proximate causes is osteoporosis. Early osteoporosis can be caused by mechanisms such as early menopause, diet, exercise, smoking, excess drinking and exposure to certain heavy metals such as lead. Some of these conditions may have their origins in childhood. Height shrinkage may thus be more severe in those with current health problems or problems from early childhood, or even be correlated with pre-shrinkage height itself. This suggests a potential relationship between shrinkage and different SES factors that are associated with current health and early childhood health problems. Furthermore, estimates of the impacts of height on other health measures, as we see in the current literature, may miss some important insights because shrinkage is not considered.

In this paper, we construct estimates of height shrinkage using techniques from the nutrition/human biology literature. We examine the SES correlates of shrinkage and use both estimated pre-shrinkage height and shrinkage as covariates in OLS regressions of health on height. The first step is to estimate the pre-shrinkage height of the seniors. This issue has a long history in the nutrition and human biology literature, even, especially, using cross-section data. Measured height is regressed on limb length(s) and the resultant prediction is used as a measure of pre-shrinkage height. This works because the limbs used in this literature do not generally shrink as people age. Lower leg length (Chumlea et al., 1985; Chumlea and Guo, 1993; Prothro and Rosenbloom, 1993; Myers et al., 1994; Zhang et al., 1998; Bermudez, 1999; Li et al., 2000; Cheng et al., 2000; Pini et al.,

2001; Knous and Arisawa, 2002), arm span from roughly the shoulder to the wrist (Kwok and Whitelaw, 1991; Kwok et al., 2002), total arm length (Mitchell and Lipschitz, 1982; Haboubi et al., 1990; Auyeung and Lee, 2001), upper arm or humeral length, tibia length (Haboubi et al., 1990) and fibula length (Auyeung and Lee, 2001) have all been employed to estimate pre-shrinkage stature.

Most of this literature simply uses lower leg or arm length to predict height using older populations. In some cases a younger population, that should not have shrunk yet, is used to estimate the function relating height to limb length(s), and the parameters used to predict height for an older population that has been shrinking, assuming that the parameters are the same for the two age groups (eg. Reeves et al., 1996). Rarely is an attempt made to actually estimate shrinkage and generally no attempt is made to relate shrinkage to socio-economic variables. Likewise very little literature exists that relates height shrinkage to other health variables, although see Hillier et al. (2012) for a recent, interesting exception relating shrinkage in older women to subsequent hip fractures and mortality. In contrast, there is some literature, though not large, that relates limb lengths to different health outcomes. The idea is that limb lengths are a proxy for childhood health influences. For instance Huang et al. (2008) use leg length and arm span to predict dementia in an older population. More interesting for our purposes, is a recent paper by Zhang et al. (2010) that uses another survey of Chinese elderly, the Chinese Longitudinal Healthy Longevity Survey, to relate arm length and knee height to cognitive impairment at older ages, finding strong negative associations between limb lengths and cognition among this older population. They do not, however examine the correlations between height shrinkage and cognition or the SES correlates of shrinkage.

Based on the national baseline data from the China Health and Retirement Longitudinal Study (CHARLS), this paper investigates the correlates of shrinkage with current SES and indicators of childhood health, and whether shrinkage is correlated with preshrinkage height and, separately, with limb lengths. We find strong negative associations between shrinkage and current measures of SES such as level of education, log of household per capita expenditure (*pce*), urban

residence, as well as strong correlations with county of current residence and province of birth. However the correlation between height shrinkage and non-location measures of childhood background are weak.

We then replace measured height by our pre-shrinkage height estimates, plus our estimates of height shrinkage, as covariates in regressions to investigate their associations with a rich set of later-life health variables: measures of cognitive functioning, hypertension, lung capacity, grip strength, balance, walking speed, self-reported general health, measures of physical functioning, activities of daily living (ADLs), instrumental activities of daily living (IADLs), the Center for Epidemiologic Studies- Depression (CES-D) scale and expectation of surviving to age 75 (for respondents under 65). We find that even controlling for SES and early life health conditions, that pre-shrinkage height and especially height shrinkage have significant associations with these later life health conditions. In general height shrinkage is negatively correlated with good health outcomes and pre-shrinkage height positively so. For many of our health variables, height shrinkage has considerably larger coefficients (in absolute value), which are more likely to be statistically significant, than for preshrinkage height.

Given our shrinkage results, height shrinkage is more a marker for later-life health problems than for early life health conditions, in contrast to pre-shrinkage height which is a marker of early life health. Hence this evidence means that it is not only early life events that are associated with late life health outcomes (childhood background variables are jointly significant in these health regressions), but health insults in later life as well. Had we only examined the correlations between measured height and these health variables, we would have missed this important insight. By providing evidence of whether and how height shrinkage is correlated with SES, this paper also validates the concern raised, but not tested, by Case and Paxson (2008a) that individuals with poor health tend to shrink more than healthy ones.

This paper is organized as follows. In Section I we discuss our model and econometric framework to estimate height shrinkage and preshrinkage height. Section II discusses the data used in this paper and summary statistics. Section III

shows how we estimate the pre-shrinkage height function from a sample of "young" respondents and height shrinkage for our "older" sample. Section IV discusses the evidence on the association between height shrinkage and SES and pre-shrinkage height. Section V provides further evidence on the association between height shrinkage, pre-shrinkage height and our health measures of this older population. Section VI concludes.

## I. Theoretic Framework

In previous studies, what has interested many researchers is the association between height and health status. A prototypical regression is:

$$y = \mathbf{x}'\boldsymbol{\theta} + p\alpha + u \quad (1)$$

in which  $y$  stands for health variables like self-reported health, ADL disability or cognitive ability;  $p$  is respondents' preshrinkage height;  $\mathbf{x}$  is short for a set of co-variates, such as demographic variables, possibly SES, or perhaps other childhood health variables;  $u$  is the error term, which is assumed in the literature to be mean independent of height and control variables. We also assume that it is mean independent of the predictors of pre-shrinkage height.<sup>1</sup> Most researchers are interested in the coefficient  $\alpha$ . However, in most situations, pre-shrinkage height is unobserved and the interviewers have available only measured height ( $h$ ), which might have been contaminated by height shrinkage ( $s$ ). The regression thus estimated is:

$$y = \mathbf{x}'\tilde{\boldsymbol{\theta}} + h\tilde{\alpha} + \tilde{u} \quad (2)$$

and for the older population measured height ( $h$ ), as an identity equals pre-shrinkage height ( $p$ ) minus height shrinkage ( $s$ ) (in the younger sample, in principle, measured height equals preshrinkage height):

---

<sup>1</sup>Of course even height, though predetermined, will be correlated with unobserved variables so that the regression coefficients in (1) are not causal effects.

$$h \equiv p - s \tag{3}$$

Height shrinkage may be independent of pre-shrinkage height, which is easier to handle, but this may not be the case. On the one hand, pre-shrinkage height is a marker of early life health status, and healthier people might shrink less with aging, have less osteoporosis for example. On the other hand, taller people may lose more height if they suffer kyphosis or some other related diseases. In these situations, the coefficients on height and  $\mathbf{x}$  in (2) are a biased estimate of the coefficients on pre-shrinkage height and  $x$  in (1) because the error in (2) will contain shrinkage that is correlated with pre-shrinkage height and  $\mathbf{x}$ .<sup>2</sup>

Maurer (2010) assumed that lower leg length was correlated with pre-shrinkage height and not correlated with height shrinkage, and used lower leg length as an instrument for measured height, using data from Latin America and the Caribbean. Then he argued that the 2SLS estimation would give consistent results. However, if pre-shrinkage height is correlated with height shrinkage conditional on control variables and error term, this suggests that lower leg length may be correlated with the error term as well, then this 2SLS estimate will also be inconsistent. We will test this directly.

In this paper, we use lower leg length and upper arm length to predict pre-shrinkage height using different data on a younger population, instead of taking them as instruments directly. We use estimates of this height function to predict pre-shrinkage height and height shrinkage for respondents from an older population, aged 60 and over. Of course in doing so we have to assume that the relationships between limb lengths and heights are the same for the two age groups. There is not much literature on this, what there is we discuss below, along with limited evidence from CHARLS that we can provide.

Firstly, we estimate the following equation for the younger group:

---

<sup>2</sup>In Section V, we will provide some evidence that height shrinkage is correlated both with SES, with pre-shrinkage height and with lower leg length, thus with the error term,  $\tilde{u}$ .

$$h_y = \mathbf{z}_y' \boldsymbol{\gamma} + \eta_y \quad (4)$$

where  $z$  is a vector representing lower leg length, upper arm length (as an adult) and a Han ethnic dummy. The variables in  $x$  and  $z$  overlap because of the Han dummy, but there are variables (the limb lengths) in  $z$  that are excluded from  $x$ . We assume

$$\mathbb{E}(\eta_y | \mathbf{z}_y, \mathbf{x}_y) = 0$$

We then apply the estimated  $\gamma$  coefficients to the older age-group to estimate their pre-shrinkage height:

$$\hat{p}_o = \mathbf{z}_o' \hat{\boldsymbol{\gamma}} \quad (5)$$

Height shrinkage is defined as the difference between pre-shrinkage height and measured height, as in (3).

After estimating pre-shrinkage height and height shrinkage, we estimate the association between height shrinkage and SES variables, i.e. education levels, per capita *pce*, age dummies, living in an urban area, marital status and childhood background variables: having an urban childhood upbringing, schooling of each parent, whether each parent had died before the respondent was 18, a self-reported general health measure of the respondent's health before age 16 and dummies for province of birth.<sup>3</sup> To test potential bias in an IV specification using limb lengths as instruments, we add limb lengths as  $x$  covariates and test their association with shrinkage. In a second specification we add pre-shrinkage height. That is:

$$s = \mathbf{x}' \boldsymbol{\beta} + p\delta + \varepsilon \quad (6)$$

---

<sup>3</sup>The childhood background variables might be thought of as possible instrumental variables for limb lengths in (4), however this would require the assumption that the only influence of childhood background on pre-shrinkage height, height shrinkage and other height outcomes, is through limb lengths, which is not consistent with the recent literature on early childhood-later life health associations. In results not shown, apart from women having an urban upbringing for upper arm length, only the province of birth dummies are significantly related to limb lengths, among the childhood background variables available to us.



with  $\mathbb{E}(\varepsilon|x, p, z) = 0$ . Finally, we estimate (1) and (2), along with (7) below, to examine the associations between height and health:<sup>4</sup>

$$y = \mathbf{x}'\boldsymbol{\theta} + p\alpha + s\kappa + u \quad (7)$$

Separate OLS estimation of (4) and (6), or (4) and (7), is the optimal 2-step GMM estimator. Our standard errors for (1), (6) and (7) are corrected for the fact that we use predicted variables as dependent and/or independent variables. We derive the asymptotic variances in Appendix 1.

## II. Data

The China Health and Retirement Longitudinal Study (CHARLS) was initiated to study the elderly population of China. It is designed to be complementary to the Health and Retirement Study (HRS) in the United States and like surveys around the world. CHARLS covers 150 counties randomly chosen across China. Twenty-eight provinces are represented in the data.<sup>5</sup> Counties were grouped into 8 geographic regions, and stratified by rural/urban status and by per capita county GDP.<sup>6</sup> Counties were then sampled, stratified, with probability proportional to population (pps).<sup>7</sup> Within counties we sampled three administrative villages or urban neighborhoods (resident committees) as our primary sampling units (psu), again using pps.<sup>8</sup>

The sampling goal within primary sampling units was 24 households with an age eligible member, defined as a person aged 45 or older. Sampling rates varied by psu. We first mapped all of the dwellings in the psu, using Google Earth maps,

---

<sup>4</sup>Note that if we estimated (1), estimating pre-shrinkage height on the older sample, we would be using 2SLS, as in Maurer (2010). We would face the same issues we raised above.

<sup>5</sup>Tibet was excluded from the study. Two other provinces, Hainan and Ningxia, both very small in population size, are not represented among the CHARLS counties..

<sup>6</sup>Data sources were the *Population Statistics by County/City of PRC*, 2009 (data from 2008) and the provincial statistical yearbooks (for GDP per capita).

<sup>7</sup>This was done by listing the stratified counties and selecting counties with a fixed interval and random starting point. This way we ensure that all parts of the GDP per capita distribution are covered.

<sup>8</sup>Data on population sizes were provided by the National Bureau of Statistics (NBS).

adjusted from the ground by our mapping teams.<sup>9</sup> From this we obtained a sampling frame of dwelling doors. We then randomly sampled 80 doors, and obtained information on the age of the oldest person and whether the dwelling was vacant (which some were). Using this information, we calculated age eligibility rates. From this information we determined psu-specific sampling rates to ensure in expectation 24 age-eligible households and re-sampled from the initial dwelling list. If a dwelling had multiple households living in it, we randomly sampled one with an age-eligible person. Households were defined as living together, sharing meals and at least some other expenses. After sampling our final list of households, we again checked for age eligibility and then randomly sampled one person age 45 or over, and their spouse (no matter the age), as our respondents.

The national baseline was fielded from late summer 2011 until March 2012 (see Zhao et al., 2012, for details). Among all households the age eligibility rate was 62% and the response rate among eligible households was 85%, 91% among rural households and 79% for urban households.<sup>10</sup> These rates compare very well with other HRS surveys in their initial waves. Sample size is 17,705 individuals with non-missing ages.

We use two samples for this paper. We estimate our preshrinkage height prediction equation using a "young" sample of respondents and spouses aged 45-49, who have presumably not started to shrink yet, or if so, have only shrunk a very small amount on average. We then use respondents and spouses aged 60 and over to predict preshrinkage heights, calculate shrinkage and estimate our models. Of the 17,705 observations, 3,451 are between 45 and 49 and 7,736 are 60 and over. Approximately 25% of "young" respondents did not get their biomarkers taken, usually because they were busy at work and unavailable. Among the "old" sample, 18% did not get any biomarkers taken, usually because they were too frail to be

---

<sup>9</sup>CHARLS mapping staff first went to the areas with GPS devices and took readings of the administrative boundaries, which were used to extract the Google Earth maps. A few primary sampling units had unreadable or no Google Earth maps, in which case we constructed the maps from the ground. In all cases we checked the maps from the ground and added to them when they were not up to date.

<sup>10</sup>Of those who did not respond, about half refused and half could not be found.

measured. Non-measurement rates were higher among those over 80 years. In addition, some observations were dropped because they had missing heights or other key variables missing or out of reasonable range. We are left with 1,101 men and 1,508 women in the "young" sample and 2,953 men and 2,935 women in the "old" sample, who have complete height and limb measurements (fewer with all of the other health variables complete).

Anthropometric measures included respondent's standing height, upper arm length and lower leg length, all measured in millimeters. The summary statistics of these variables are shown in Panels A and B of Table 1 for the "young" and "old" samples, respectively. Height was measured using a stadiometer directly from the heel to the top of head with the elders standing up-right. Upper arm length was measured with a Martin caliper with the respondent standing and holding the left or right arm at a right angle. We measured from the acromion process of the scapula to the olecranon process. Lower leg length was also measured using a Martin caliper from the right knee joint to the ground (pictures of how limb lengths are measured, used in the CHARLS training, are provided in Appendix 2).. Measured heights are smaller for the older group, by some 4.0 cm for men and 4.6 cm for women. Much of this difference could be due to shrinkage, although it could also be that older birth cohorts were less tall. Comparing upper arm lengths, they are very close between the 45-49 and 60 and over groups, suggesting that shrinkage may be the more important explanation. On the other hand, lower leg lengths are about 0.6 cm smaller for the over 60 group, for both men and women, suggesting some possible cohort effects. Indeed, our estimates of shrinkage for the older population are 3.3cm for men and 3.8cm for women, a bit less than the differences in measured heights between the two age groups, consistent with part of the height differences being birth cohort differences.

#### **Table 1 about here**

As mentioned above, this study examines the associations between pre-shrinkage height and height shrinkage on different measures of health of older people. We start with cognition questions, which are grouped into three, following

McArdle (2010) and Smith et al. (2011). The first component is the Telephone Interview of Cognitive Status (TICS). There are ten questions in this part, from awareness of the date (using either solar or lunar calendar), the day of the week and season of the year, to successively subtracting 7 from 100. An index is formed of the number of correct answers. This is a measure of the mental intactness of the respondent (Smith et al., 2011). A second set of questions asks a respondent to recall a series of 10 simple nouns and to recall again after approximately 10 minutes. Following McArdle (2010), we average the number of correct answers as our dependent variable. This is a measure of episodic memory, and is a component of fluid intelligence (Smith et al., 2011). Finally respondents are shown a picture of two overlapping pentagons and asked to draw it. We score the answer as 1 if the respondent successfully performs this task.

We have several biomarker variables available. We measure blood pressure three times. We create a dummy variable equal to one if a respondent has hypertension. For this case we take means of systolic and diastolic measurements and assign a hypertensive status equal to one if mean systolic is 140 or greater or if mean diastolic is 90 or greater. In addition respondents self-report if they have been diagnosed by a doctor with hypertension and we include those cases as being hypertensive. Respondents blow into a peak flow meter three times to measure lung capacity and we take the average. Respondents have their grip strength measured by a dynamometer. Two measurements are taken from each hand. We use average measurement from the self-reported dominant hand. Respondents are given a balance test, whether they can stand semi-tandem or full tandem. Because most can stand full tandem, we create a dummy equal to 1 if they can do so.<sup>11</sup> Finally we conduct a timed walk of 4 meters, asking the respondent to walk at a "normal" speed.

The remaining health measures are self-reported. General health is reported on a scale: very good, good, fair, poor, very poor. We construct a binary variable equal to one if health is reported as very poor or poor, zero otherwise. Respondents

---

<sup>11</sup>Respondents under 70 are asked to stand in full tandem for 60 seconds, those 70 and over for 30. We include age dummies as covariates, which will capture this difference.

are asked about whether they have difficulty in performing certain classes of activities: physical functioning, ADLs and IADLs.<sup>12</sup> We count the number of items in each group that the respondent claims having difficulty in performing or cannot perform. The expected survival question asks respondents to rank their expectation of surviving to a specific older age on a five point scale, from almost impossible to almost certain. We group the bottom two answers, almost impossible and not very likely. Because different age groups are asked survival chances to different ages, we standardize by only using those respondents under age 65, who are asked their survival chances to age 75. Similarly, respondents answered a Chinese version of CES-D 10 questionnaire in the survey, which contained 10 questions about the respondents' depression status. Based on that, we constructed a CES-D scale, with range from 0 to 30.

Mean values and standard deviations of all the health variables are provided in Panel B of Table 1 for the "old" sample. As is generally the case, health measures for older women are worse than for men. This is true both for self-reported measures such as poor general health, difficulties with physical functioning or ADLs, and the CES-D depression scale, and for biomarkers such as hypertension, the cognition measures, grip strength and lung capacity.

Panel B also reports summary statistics of demographic variables like education level, log of household per capita expenditure (*pce*),<sup>13</sup> marital status and type of areas (urban/rural) where respondents live at the time of the survey. The current generation of elderly population in China has only a small amount of schooling, particularly among women. Fifty-four percent of women 60 and over are illiterate, twenty percent among men. Only 8% of older men and 3% of women have completed senior high school or more. However, 56% of men have completed primary school, and 35% of women. When we compare these numbers to the

---

<sup>12</sup>There are nine questions on physical functioning, ranging from having difficulty running or jogging 1 km, to walking 1 km, to carrying a heavy bag of groceries, to picking up a small coin. There are 6 ADL questions (eg. getting into and out of bed or using the toilet) and 5 IADL questions (eg. doing household chores, shopping, or managing money).

<sup>13</sup>Per capita expenditures include the value of food consumed from own production. We prefer *pce* to income because *pce* is measured with less error and better represents long-run resources, since households smooth their consumption over time.

parents of these elderly, some progress had been made, since 70% of fathers and over 90% of mothers are reported to be illiterate (no schooling or less than primary school completion-see Panel B). The preponderance of our respondents are still married, more so among men, since their spouses tend to be younger. An overwhelming majority, roughly 80% of older men and women, live in rural areas.

Childhood background variables are also reported in Panel B. An even larger percent, over 90, have a rural background as a child. About 10-12% of fathers died before the respondent was age 18 and about 6-7% of mothers. CHARLS has a retrospective question about general health before the respondent turned 16 (an average over that period), with categories excellent, very good, good, fair or poor. This has been successfully used by HRS and other aging surveys, including the CHARLS pilot, and has been linked to later life health outcomes (e.g. Smith, 2009). In the CHARLS sample, 6% of men and 9% of women report that their childhood health was poor. Finally CHARLS also elicits province of birth. Evidence on public health infrastructure for pre-revolutionary China is scant, but some evidence exists that in Beijing, better water and sanitation facilities were built between 1910 and 1920 (Campbell, 1997) and that led to a rapid decline in infant mortality in there. This would have affected our cohorts. For other major cities there is some, but not much, evidence that public health infrastructure was being built during that time period (Campbell, 1997).

### III. Estimation of Pre-shrinkage Height

Following the methodology in the medical literature, we use lower leg length and upper arm length and estimate gender-specific equations using measured height as the dependent variable. Additionally, we add quadratics in both limb lengths and interactions to allow for nonlinearities. We also add a Han dummy variable to pick up potential ethnic differences.<sup>14</sup>

---

<sup>14</sup>Ethnic differences in the proportions of limb lengths to height have been found in the literature (see Steele, 1987, for example). Age is not included. Age itself should only have an influence on pre-shrinkage height through birth cohort effects. These are likely but the sample we estimate our coefficients for only spans 10 years. We do try one specification that includes a linear trend in year of birth, but it is never significant at standard levels.

The steps to estimate pre-shrinkage height is as follows: first, we use data from the "young" group, aged 45-49, and regress measured height on lower leg length, upper arm length, their squares and interaction and the Han dummy. These coefficients are then applied in the "older" sample, those aged 60 and above, and the predicted value is the estimated pre-shrinkage height for this group. Some medical studies have used this approach, separating "young" and "old" groups, include Steele (1987) and Reeves et al. (1996).<sup>15</sup> A strong assumption is required that any secular changes in height across birth cohorts (which are important in China) do not change the relationship between height and limb length (see Leung et al., 1996 and Kwok et al., 2002).<sup>16</sup> While we cannot test this restriction directly, we provide some indirect evidence for the five year age group, 45-49. We add individual age dummies plus interactions between the age dummies and each of the limb lengths. We then test the joint significance of the age dummies plus limb interactions to examine whether the coefficients vary by age. While this is suggestive, it does not rule out possible change of coefficients across a wider age gap. The regressions are shown in Table 2.

## Table 2 about here

Columns (1)-(3) in Table 2 show the coefficients of the pre-shrinkage height function in the male sample and columns (4)-(6) for the female sample. We first show a linear specification in limb lengths and the Han dummy, then add quadratics and an interaction, and finally add the age dummies and limb interactions. In the linear models, an increase of 1cm in lower leg length is associated with an increase of 1cm in measured height for men and 0.93 for women,<sup>17</sup> as one would expect if shrinkage has not yet begun. The coefficients on upper arm length are 0.69 and

---

<sup>15</sup>However, most of the medical literature estimates the coefficients using the same age-group sample as is used to predict preshrinkage height.

<sup>16</sup>Kwok et al. (2002) use data on an older sample in China to estimate their prediction equation, but they first remove observations who have symptoms of vertebral deformity based on x-ray images. The remaining older respondents should not have experienced shrinkage. Using these data, they find the same ratio of total arm span to height for younger and older men, but a slightly higher ratio for older women.

<sup>17</sup>Neither coefficient is significantly different from one at standard levels.

0.78 for men and women respectively, each significant at .001. The Han dummy is positive for both men and women, but significant (at 5%) only for women. The quadratics and interaction are always jointly significant at under .001, while the age dummies and limb interactions are not jointly significant at standard levels; hence we use columns (2) and (5) as our preferred estimates. The marginal effects on height of both lower leg and upper arm lengths are positive over the entire distribution, and convex. The  $R^2$ 's are over .51 for both men and women.<sup>18</sup>

After we obtain our pre-shrinkage height estimates for the 60 and older group, height shrinkage is defined as the estimated pre-shrinkage height less the current measured height. The summary of our estimates are shown in Panel B in Table 1. Mean height shrinkage is 3.3 cm for men and 3.8 cm for women, which is consistent with findings in the human biology literature that women have more problems with vertebral deformity (see Kwok et al., 2002).

### Figure 1 about here

Figure 1 shows the age pattern of measured height, pre-shrinkage height and height shrinkage by gender. The top two figures show non-parametric graphs of measured height and pre-shrinkage height as a function of age. And the bottom two graphs show the pattern of height shrinkage and age for males and females respectively.<sup>19</sup> From the top two graphs, estimated pre-shrinkage height does not decline much with age, a little more for men than for women. However measured height does decline with age, indicating that height shrinkage increases, as shown in the bottom two figures. Our pre-shrinkage height estimates do not correct for mortality selection. If we assume that respondents who survived to older ages are those who were taller and less frail, then adding those who died back would result in pre-shrinkage heights declining with age. This is what we would expect if older

---

<sup>18</sup>Many of the medical papers obtain higher  $R^2$ s for their height prediction equations, but they generally have extremely small samples and highly controlled, usually laboratory, venues in which the measurements are conducted. These settings should minimize measurement error, compared to a large-scale population survey such as CHARLS..

<sup>19</sup>The non-parametric curves are calculated using a Jianqing Fan (1992) locally weighted regression smoother, which allows the data to determine the shape of the function.



birth cohorts faced worse health conditions at birth, and in early life.

As a check on our preshrinkage height estimates, we compare our CHARLS preshrinkage heights for the sample aged 60-69 in 2011, by year of birth, to measured heights in another data source, the China Health and Nutrition Survey (CHNS). We use the same birth year cohorts in both data sets, but in the CHNS data, we can measure heights of these cohorts 20 years earlier, in 1991, when they would be aged 40-49, and so should not have begun to shrink much yet. We thus expect their measured heights in 1991 to be close to our estimated preshrinkage heights in the CHARLS data for the same birth year cohort.<sup>20</sup> The CHNS data in 1991 only covers 8 provinces, not 28 provinces as in CHARLS. To make the comparisons cleaner, we restrict the CHARLS data to the same 8 provinces as covered in CHNS. The results are shown in Table 3. Comparing mean heights by birth year cohort between being measured in 1991 in CHNS and in 2011 in CHARLS, heights in 1991 are higher, by generally between 1.5-3.2 cm for men, depending on the age, and between 1.3-5.3cm for women. These differences are consistent with shrinkage. Comparing mean heights in 1991 with estimated preshrinkage heights from CHARLS, the results show a close correspondence. For women, the differences between the CHARLS estimated preshrinkage heights and the CHNS measured heights is very small, generally under 0.7cm in absolute value and often less than 0.5cm. For men, aged 60-64 in 2011 (40-44 in 1991), the differences are very small as well; they increase some for those aged 63-69 in 2011, which may indicate that there is some shrinkage that has begun in this age group.<sup>21</sup>

### **Table 3 about here**

---

<sup>20</sup>We thank David Cutler for this idea. Note that the CHNS data do not include limb lengths, so we cannot use the CHARLS preshrinkage height function estimates to predict individual preshrinkage heights with CHNS observations. Also only one height observation per person is available in CHNS, so it is not possible to take differences in height measurements to measure shrinkage directly.

<sup>21</sup>Plotting the ratio of lower leg or upper arm length to measured height in the CHARLS data does show a slight increase for those in their late 40s, which is consistent with this conjecture.

#### IV. Height Shrinkage, Pre-Shrinkage Height and SES

Very few studies have been able to measure height shrinkage and we know precious little about the correlations between shrinkage and later life SES, early life health conditions and family background. Further, as noted, any correlations between height shrinkage and upper arm and lower leg length are important since they determine whether an IV estimator using lower leg and upper arm lengths as IVs for measured height in health equations is consistent. Table 4 shows the gender-specific results of the association between SES, early life conditions, upper arm and lower leg length, and height shrinkage. All regressions control for basic demographic variables, including dummy variables for age, Han ethnicity, marital status, urban residence and current residential county. We also include covariates measuring early life conditions, including dummies for province of birth, urban upbringing before age 16, for schooling of the father and mother, for whether the father and mother died by respondent's age 18, and for whether the respondent reported being in poor health on average before age 16. In columns 2 and 5 we add preshrinkage height. In columns 3 and 6 we replace preshrinkage height with lower leg and upper arm lengths in order to test for bias with an IV estimator of (2). All estimates correct standard errors for the fact that we predict shrinkage and pre-shrinkage heights (see Appendix 1 for detailed derivations).

**Table 4 about here**

From these estimates, we find that the SES variables are very important predictors of height shrinkage; the Wald tests are all highly significant. Dummy coefficients of level of education are negative, monotonically declining with higher education and jointly significant. For men, having primary schooling, compared to being illiterate, is associated with shrinkage of 0.9cm less. This is a large difference when compared to mean shrinkage for men of 3.3cm. For women, the difference is smaller, 0.6cm, but still large compared to the mean female shrinkage of 3.8cm. For men or women with senior high school or more, the additional shrinkage compared to people with primary schooling is 1cm for men, or slightly more for

women. So these regression coefficients are large. One potential explanation can be that people with higher education level are more likely to have had better health behaviors when younger. They are also likely to have had better health during childhood, perhaps in ways not measured by our childhood general health variable.

Household log *pce* is negatively associated with height shrinkage, especially for men. A one standard deviation on log *pce* is approximately 0.9, so a two standard deviation increase in log *pce* would be associated with a reduction in shrinkage of 0.4cm, a much lower difference than we see for schooling. It is likely that higher income people may be able to purchase better medical care and nutritious food for themselves, although there is likely to exist reverse causality as well, which may explain why the coefficients are more negative for men.

Being currently married is associated with less shrinkage for men, but not significant. Marriage is often found to be correlated positively with better health and more happiness, and is associated with better labor market outcomes for men, so this is not surprising, though, again, we must remember that these estimates are not necessarily causal. Not surprisingly, there are very strong positive associations between shrinkage and age. Currently living in an urban area is significantly associated with less shrinkage for both men and women, on the order of 1cm less for women, and almost that much for men. These effects are similar to the schooling differences from illiterate to primary school. The county dummies are jointly significant at under the .001 level. This is consistent with results such as Strauss et al. (2010), who find very strong community effects on health outcomes for the elderly in China, using the CHARLS Pilot data. Early childhood background and health are not jointly significant in these regressions. However, having had poor childhood health is associated with more shrinkage for women, significant at 10%, although the magnitude of the difference is only half of the difference between those currently living in urban versus rural areas (and of the difference between those who are illiterate and those with primary schooling). Dummies for birth provinces are jointly significant, for both men and women.

Table 4 also demonstrates joint significance of lower leg and upper arm lengths for men at the 10% level, though they are not significant for women. A one

standard deviation increase in lower leg length for men is associated with shrinkage of 0.3cm less. Preshrinkage height is positively associated with shrinkage for women, although not quite significant at 10% and the magnitude of the coefficient is small. The evidence for men is consistent with IV estimates of measured height using limb lengths as instruments, shown below, being inconsistent.

## V. Results: Impact of Estimated Height on Health Outcomes

Since there is a growing literature, cited above, that investigates how height is associated with other adult health outcomes, it is of interest to explore this with our estimates of preshrinkage heights and height shrinkage. We do not claim causality from these regressions, because of the usual problem of omitted variables, but also because in some instances reverse causality is possible.<sup>22</sup> The procedure is to regress our health measures first on measured height and control variables to get our baseline estimates. Then we replace measured height by predicted pre-shrinkage height and then add height shrinkage. Standard errors are again corrected for predicted shrinkage and preshrinkage heights. We also use an IV estimator for measured height, using limb lengths, their quadratics and interaction as instruments, to compare to our OLS estimates. Since some health outcomes are missing for some observations, the number of observations differs by outcome.

Table 5.1-5.3 show the results from the regressions of our health measures on height.<sup>23</sup> The same demographic and SES controls, and controls for early life conditions that we use in Table 4 are added in all the regressions. We start with the cognition outcomes in Table 5.1. Measured height is positively and significantly associated with all of the cognition measures for both men and women.

---

<sup>22</sup>One potential example is with the depression score and shrinkage. Depression is associated in women with early menopause. Menopause in turn is associated with osteoporosis, which can lead to shrinkage. Now our depression score is current and may not indicate past episodes, but we also know that if a person has had one bout of depression, that increases the likelihood of more, later.

<sup>23</sup>Some studies, such as Case and Paxson (2008a) and Smith et al. (2012) omit current SES variables in order to estimate the associations of health with height, which they argue works through childhood health. However, since we are interested in associations of later life health measures with height shrinkage, which is associated with later life events as well as ones in childhood, we include current SES measures.

Case and Paxson (2008a) find such relationships among the older population in the United States using the Health and Retirement Study (HRS) when they do not include current SES, notably schooling. However when they add schooling height many times becomes insignificant, unlike our results for cognition (also see Smith et al., 2012 for evidence on China of height associations with health measures of the elderly, also when current SES variables are not included). A likely mechanism for this relationship is the positive association between childhood height and childhood and later cognition. There exists a large literature on early child height impacts on later child cognition; Case and Paxson (2008b) is a recent example (see Glewwe and Miguel, 2008 and Strauss and Thomas, 2008 for reviews). Since childhood heights are strongly related to adult heights and cognition skills persist from childhood through adulthood, it is not surprising to see this relationship among older persons. However, the magnitude of the our estimated changes in cognition associated with a one or even two standard deviation increase in height, are smaller than estimated change associated with increasing the level of education from illiteracy to completing primary school. This is consistent with the results in Case and Paxson (2008a).

**Table 5.1 about here**

When we replace measured heights by preshrinkage heights and height shrinkage, preshrinkage height has the same positive, significant association for women with the TICS and draw a picture variables, though not for the word recall. For men, preshrinkage height is positive and significantly associated with the TICS and word recall. Height shrinkage is, however, strongly, negatively correlated with all of these measures and for both men and women, suggesting that a part, perhaps a large part, of the association between measured height and cognition occurs through height shrinkage. Indeed, for men, the magnitude of the shrinkage coefficients is substantially larger, compared to the coefficients on measured height. As is true for measured height, the associated differences in cognition from height shrinkage is much smaller than the associated difference between levels of

schooling.<sup>24</sup> Furthermore, the shrinkage coefficients are often larger in magnitude than the coefficients on preshrinkage height, which are only significant in half of the six cases. Hence if we had used only measured height as our measure of height, we might have wrongly concluded that it is height from childhood, as proxied by preshrinkage height, that was correlated with these later life health variables, instead of height shrinkage.

Shrinkage in turn, as is clear from results in Table 4, is more strongly associated with current SES covariates than with childhood background variables. In contrast, both current and childhood SES and childhood health variables are each jointly significantly associated with the cognition outcomes, as are the current county of residence dummies and, for women, the birth provinces. So these results imply that later life cognition and health are associated with health events throughout the life cycle and in later life, not just from early childhood, and that the association with height is more with shrinkage than preshrinkage height.<sup>25</sup> We also see in Tables 5.1-5.3 that using the Maurer IV approach for measured height results in coefficients that are fairly close to the coefficients on preshrinkage height when we include both preshrinkage height and shrinkage. However, the comparison of preshrinkage height and shrinkage coefficients are not forthcoming from the IV results.

## Table 5.2 about here

In Table 5.2 we show results for the biomarkers. Preshrinkage height is significantly, positively related to lung capacity and grip strength, and height shrinkage is significantly, negatively related to both outcomes. In contrast to the cognition results, the size of the shrinkage coefficients are larger for these outcomes.

---

<sup>24</sup>If there were no shrinkage, the number of words correctly recalled is estimated to be 0.06 higher for women, the mean number of words correctly recalled being 2.85, so a small effect. In contrast, the conditional difference in word recalled correctly between women with no schooling and those with primary schooling is 0.65 (results not shown).

<sup>25</sup>Note that if we have to use only measured height as our covariate, but that the correct specification is using both preshrinkage height and shrinkage, we implicitly assume that the coefficients on preshrinkage height and shrinkage are equal in magnitude and of opposite signs. While there are some cases in which this seems to hold, there are others, like draw a line, where it does not.

The difference in lung capacity or grip strength for men if there was no shrinkage is predicted to be only a little less than the predicted difference between illiterate and literate men.<sup>26</sup> Men who have shrunk more take more time to do the timed walk, but for women shrinkage and preshrinkage heights are not related to walk time. Hypertension and ability to balance are also unrelated to both preshrinkage height and shrinkage. Current and childhood SES are related to many of these outcomes, as are current county of residence and province of birth.

Table 5.3 has results for self-reported health outcomes. As can be seen, shrinkage is generally related to worse outcomes, and more strongly and significantly so than is preshrinkage height positively correlated with these outcomes. Examining the OLS or IV results for measured height would, again, not provide this result. For men, this is so for the CES-D depression index, the self-assessed likelihood of not surviving to age 75 (for those 65 and younger), the number of measures of physical functioning that the respondent reports having difficulty with, and having poor or very poor general health. For women, shrinkage is significantly associated with having more difficulties with measures of physical functioning, ADLs, IADLs and the likelihood of not surviving to age 75.. What is surprising about these results is that preshrinkage height has positive, significant at the 10% level, associations with bad outcomes for ADLs and IADLs for women, although not for men. This is very unlike the results for cognition and the biomarkers. Mortality selection could be partly causing this, but we cannot be more than speculative on this point.

### **Table 5.3 about here**

---

<sup>26</sup>For example, lung capacity would be 5.3 higher with no shrinkage, compared to a difference of 7.3 between illiterates and those completing primary school. Both of these differences are small compared to mean levels of lung capacity or to the standard deviations. The differences in predicted grip strength are a little higher relative to mean or standard deviations of grip strength.

## VI. Conclusions

According to Barker (1994), childhood health in uterus has a lasting impact on health, including at old ages. Height has been used widely as an indicator in part of childhood health. However, because height shrinks with aging, it suffers a measurement error problem when studying its impact on health outcomes at older ages.

Based on unique data of Chinese aged 45 and older, we address this problem by making use of upper arm and lower leg lengths to construct estimates of the relationship between these limb lengths and measured height, on a population aged 45-49, and then use these estimates to estimate preshrinkage height and height shrinkage on a population 60 years and older. We then investigate the association between height shrinkage, SES variables and variables measuring different dimensions of childhood health. We follow this exercise by examining the associations between measured height on the one hand, or pre-shrinkage height and shrinkage on the other, and a rich set of health variables, including measures of cognition, biomarkers, as well as various self-reported health measures.

The results in this paper show that shrinkage and socio-economic variables such as schooling, household per capita expenditure and current urban residence are negatively correlated for both men and women. Differences are largest across levels of schooling and urban/rural residence. Childhood background factors are much more weakly associated with shrinkage, excepting province of birth, which is highly significant. These results are consistent with the concerns raised in Case and Paxson (2008a).

Height shrinkage, and to a lesser extent, pre-shrinkage height, are also correlated with many later life health outcomes, particularly cognition and biomarker measures. The shrinkage coefficients tend to be larger than for preshrinkage height, suggesting that current health issues are important in understanding health of the elderly, not just events in early childhood. In general the more the shrinkage the worse are these other health outcomes.



## Appendix 1: Asymptotic Variances

Table 4 uses a constructed dependent variable, height shrinkage, while Tables 5.1-5.3 use predicted preshrinkage height and, in some specifications, height shrinkage, as right hand side variables. Furthermore, the predicted preshrinkage height coefficients are derived from a different sample. This suggests that a 2 sample GMM procedure might be appropriate (eg. Ridder and Moffitt, 2007), however we are not using the standard setup because we do not use all of the variables in the second stage to predict preshrinkage height. We derive the asymptotic variances here.

### Shrinkage as dependent variable

The regression with shrinkage as the dependent variable is

$$s = \mathbf{x}'\boldsymbol{\beta} + p\delta + \varepsilon$$

with  $\mathbb{E}(\varepsilon|\mathbf{x}, p, \mathbf{z}) = 0$ . Substitution of (4) gives

$$\mathbf{z}'\boldsymbol{\gamma} - h = \mathbf{x}'\boldsymbol{\beta} + \mathbf{z}'\boldsymbol{\gamma}\delta + (\delta - 1)\eta + \varepsilon = \mathbf{x}'\boldsymbol{\beta} + \mathbf{z}'\boldsymbol{\gamma}\delta + \zeta$$

with  $\zeta = (\delta - 1)\eta + \varepsilon$ . Rearranging gives

$$-h = \mathbf{x}'\boldsymbol{\beta} + \mathbf{z}'\boldsymbol{\gamma}(\delta - 1) + \zeta = \mathbf{x}'\boldsymbol{\beta} + \mathbf{z}'\boldsymbol{\gamma}\alpha + \zeta$$

Note that  $\alpha = -1$  if pre-shrinkage height is not in the relation.

This equation is estimated on the older sample with  $\boldsymbol{\gamma}$  being estimated on the younger sample. The unconditional moment restrictions are

$$\mathbb{E}_O \left[ \begin{pmatrix} \mathbf{x} \\ \mathbf{z} \end{pmatrix} (-h - \mathbf{x}'\boldsymbol{\beta} - \mathbf{z}'\boldsymbol{\gamma}\alpha) \right] = 0$$

$$\mathbb{E}_Y[\mathbf{z}(p - \mathbf{z}'\boldsymbol{\gamma})] = 0$$

Define

$$m_O(\beta, \alpha, \gamma) = \begin{pmatrix} \mathbf{x} \\ \mathbf{z} \end{pmatrix} (-h - \mathbf{x}'\beta - \mathbf{z}'\gamma\alpha)$$

$$m_Y(\beta, \alpha, \gamma) = \mathbf{z}(p - \mathbf{z}'\gamma)$$

We have

$$\mathbb{E}_O[\mathbf{m}_O(\beta_0, \alpha_0, \gamma_0)\mathbf{m}_Y(\beta_0, \alpha_0, \gamma_0)'] = 0$$

so that the variance matrix of the moment conditions is

$$\mathbf{W} = \begin{pmatrix} \sigma_\zeta^2 \mathbb{E}_O(\mathbf{x}\mathbf{x}') & \sigma_\zeta^2 \mathbb{E}_O(\mathbf{x}\mathbf{z}') & 0 \\ \sigma_\zeta^2 \mathbb{E}_O(\mathbf{z}\mathbf{x}') & \sigma_\zeta^2 \mathbb{E}_O(\mathbf{z}\mathbf{z}') & 0 \\ 0 & 0 & \sigma_\eta^2 \mathbb{E}_Y(\mathbf{z}\mathbf{z}') \end{pmatrix}$$

with estimator

$$\hat{\mathbf{W}} = \begin{pmatrix} \hat{\sigma}_\zeta^2 \frac{\mathbf{X}_O' \mathbf{X}_O}{N_1} & \hat{\sigma}_\zeta^2 \frac{\mathbf{X}_O' \mathbf{Z}_O}{N_1} & 0 \\ \hat{\sigma}_\zeta^2 \frac{\mathbf{Z}_O' \mathbf{X}_O}{N_1} & \hat{\sigma}_\zeta^2 \frac{\mathbf{Z}_O' \mathbf{Z}_O}{N_1} & 0 \\ 0 & 0 & \hat{\sigma}_\eta^2 \frac{\mathbf{Z}_Y' \mathbf{Z}_Y}{N_2} \end{pmatrix}$$

where  $\mathbf{X}_O, \mathbf{Z}_O$  are the matrices with covariates for the old sample and  $\mathbf{Z}_Y$  is the matrix with covariates of the young sample.

The inverse of  $\hat{\mathbf{W}}$  is block diagonal and that implies that the optimal GMM estimator with weighting matrix  $\hat{\mathbf{W}}^{-1}$  is the solution to

$$\mathbf{X}_O'(-h_O - \mathbf{X}_O\beta - \mathbf{Z}_O\gamma\alpha) = 0$$

$$\mathbf{Z}_O'(-h_O - \mathbf{X}_O\beta - \mathbf{Z}_O\gamma\alpha) = 0$$

$$\mathbf{Z}_Y'(p_Y - \mathbf{Z}_Y\gamma) = 0$$

Therefore the optimal GMM estimator regresses  $p$  on  $\mathbf{z}$  in the young sample and uses these estimates in the old sample. We can therefore rewrite the moment

function for the old sample as (we use the same notation for the moment function)

$$m_O(\beta, \alpha, \gamma) = \begin{pmatrix} \mathbf{x} \\ \mathbf{z}'\gamma \end{pmatrix} (-h - \mathbf{x}'\beta - \mathbf{z}'\gamma\alpha)$$

and use unweighted GMM.

To obtain the asymptotic variance we first compute

$$\begin{aligned} \frac{\partial \mathbf{m}_O}{\partial \beta'}(\beta, \alpha, \gamma) &= - \begin{pmatrix} \mathbf{x} \\ \mathbf{z}'\gamma \end{pmatrix} \mathbf{x}' \\ \frac{\partial \mathbf{m}_O}{\partial \alpha}(\beta, \alpha, \gamma) &= - \begin{pmatrix} \mathbf{x} \\ \mathbf{z}'\gamma \end{pmatrix} \mathbf{z}'\gamma \\ \frac{\partial \mathbf{m}_O}{\partial \gamma'}(\beta, \alpha, \gamma) &= - \begin{pmatrix} \mathbf{x} \\ \mathbf{z}'\gamma \end{pmatrix} \alpha \mathbf{z}' + \begin{pmatrix} \mathbf{0} \\ \mathbf{z}' \end{pmatrix} (-h - \mathbf{x}'\beta - \mathbf{z}'\gamma\alpha) \\ \frac{\partial \mathbf{m}_Y}{\partial \beta'}(\beta, \alpha, \gamma) &= \mathbf{0} \\ \frac{\partial \mathbf{m}_Y}{\partial \alpha}(\beta, \alpha, \gamma) &= \mathbf{0} \\ \frac{\partial \mathbf{m}_Y}{\partial \gamma'}(\beta, \alpha, \gamma) &= -\mathbf{z}\mathbf{z}' \end{aligned}$$

The variance matrix of

$$\begin{pmatrix} \hat{\beta} \\ \hat{\alpha} \\ \hat{\gamma} \end{pmatrix}$$

is

$$(\mathbf{A}'\mathbf{W}^{-1}\mathbf{A})^{-1}$$

with

$$\mathbf{A} = \begin{pmatrix} -\mathbb{E}_O(\mathbf{x}\mathbf{x}') & -\mathbb{E}_O(\mathbf{x}\mathbf{z}')\gamma & -\mathbb{E}_O(\mathbf{x}\mathbf{z}')\alpha \\ -\gamma'\mathbb{E}_O(\mathbf{z}\mathbf{x}') & -\gamma'\mathbb{E}_O(\mathbf{z}\mathbf{z}')\gamma & -\gamma'\mathbb{E}_O(\mathbf{z}\mathbf{z}')\alpha \\ 0 & 0 & -\mathbb{E}_Y(\mathbf{z}\mathbf{z}') \end{pmatrix}$$

and (using the same notation for the variance of the new moment conditions)

$$\mathbf{W} = \begin{pmatrix} \sigma_{\zeta}^2 \mathbb{E}_O(\mathbf{x}\mathbf{x}') & \sigma_{\zeta}^2 \mathbb{E}_O(\mathbf{x}\mathbf{z}')\gamma & \mathbf{0} \\ \sigma_{\zeta}^2 \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{x}') & \sigma_{\zeta}^2 \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{z}')\gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_{\eta}^2 \mathbb{E}_Y(\mathbf{z}\mathbf{z}') \end{pmatrix}$$

Now

$$\mathbf{A}'\mathbf{W}^{-1}\mathbf{A} = \begin{pmatrix} \sigma_{\zeta}^{-2} \mathbb{E}_O(\mathbf{x}\mathbf{x}') & \sigma_{\zeta}^{-2} \mathbb{E}_O(\mathbf{x}\mathbf{z}')\gamma & \sigma_{\zeta}^{-2} \mathbb{E}_O(\mathbf{x}\mathbf{z}')\alpha \\ \sigma_{\zeta}^{-2} \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{x}') & \sigma_{\zeta}^{-2} \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{z}')\gamma & \sigma_{\zeta}^{-2} \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{z}')\alpha \\ \sigma_{\zeta}^{-2} \mathbb{E}_O(\mathbf{z}\mathbf{x}')\alpha & \sigma_{\zeta}^{-2} \mathbb{E}_O(\mathbf{z}\mathbf{z}')\gamma'\alpha & \sigma_{\eta}^{-2} \mathbb{E}_Y(\mathbf{z}\mathbf{z}') + \sigma_{\zeta}^{-2} \mathbf{C} \end{pmatrix}$$

with

$$\mathbf{C} = \alpha^2 \begin{pmatrix} \mathbb{E}_O(\mathbf{z}\mathbf{x}') & \mathbb{E}_O(\mathbf{z}\mathbf{z}')\gamma' \end{pmatrix} \begin{pmatrix} \mathbb{E}_O(\mathbf{x}\mathbf{x}') & \mathbb{E}_O(\mathbf{x}\mathbf{z}')\gamma \\ \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{x}') & \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{z}')\gamma \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}_O(\mathbf{x}\mathbf{z}') \\ \gamma' \mathbb{E}_O(\mathbf{z}\mathbf{z}') \end{pmatrix}$$

If pre-shrinkage height is excluded, i.e.  $\alpha = -1$  the second row and column in  $\mathbf{A}'\mathbf{W}^{-1}\mathbf{A}$  are removed, we substitute  $\alpha = -1$ , and

$$\mathbf{C} = \mathbb{E}_O(\mathbf{z}\mathbf{x}') (\mathbb{E}_O(\mathbf{x}\mathbf{x}'))^{-1} \mathbb{E}_O(\mathbf{x}\mathbf{z}')$$

The resulting variance matrix is for

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix}$$

We estimate  $\mathbb{E}_O(\mathbf{x}\mathbf{x}')$  by

$$\frac{\mathbf{X}'_O \mathbf{X}_O}{N_1}$$

and same for other moments for the older population. For the younger population

we estimate  $\mathbb{E}_Y(\mathbf{z}\mathbf{z}')$  by

$$\frac{\mathbf{Z}_Y' \mathbf{Z}_Y}{N_2}$$

The variances  $\sigma_\eta^2$  and  $\sigma_\zeta^2$  are estimated in the usual way. Note that we do not have to make an assumption on the correlation of  $\eta$  and  $\varepsilon$  for the older population (which would not be identified).

### Pre-shrinkage height and shrinkage as independent variables

The basic regression is now:

$$y = \mathbf{x}'\boldsymbol{\theta} + p\alpha + s\kappa + u$$

with  $\mathbb{E}(u|\mathbf{x}, p, s) = 0$ . Substitution of the prediction equation gives

$$y = \mathbf{x}'\boldsymbol{\theta} + \mathbf{z}'\boldsymbol{\gamma}\alpha + (\mathbf{z}'\boldsymbol{\gamma} - h)\kappa + (\alpha + \kappa)\eta + u = \mathbf{x}'\boldsymbol{\theta} + \mathbf{z}'\boldsymbol{\gamma}(\alpha + \kappa) - h\kappa + (\alpha + \kappa)\eta + u = \mathbf{x}'\boldsymbol{\theta} + \mathbf{z}'\boldsymbol{\gamma}\mu + h\nu + \zeta$$

with  $\mu = \alpha + \kappa$  and  $\nu = -\kappa$  and  $\zeta = \mu\eta + u$ .

If we read for  $\mathbf{x}'$  the vector  $(\mathbf{x}' \ h)$  and for  $\boldsymbol{\beta}'$  the vector  $(\boldsymbol{\theta}' \ \nu)$  and for  $\delta$  the parameter  $\mu$ , then we see that the variance matrix in the previous section applies with these changes (if we use the same estimator). This gives us the variance matrix of

$$\begin{pmatrix} \hat{\theta} \\ \hat{\nu} \\ \hat{\mu} \\ \hat{\gamma} \end{pmatrix}$$

From this we easily obtain the variance matrix of the original parameters.

## References:

- Abbott, Robert, Lon White , G. Webster Ross, Helen Petrovich, Kamal Masaki, David Snowdon and J. David Curb. 1998. "Height as a Marker of Childhood Development and Late-life Cognitive Function: the Honolulu-Asia aging study." *Pediatrics*, 102(3): 602–609.
- Alderman, Harold and Jere Behrman. 2006. Reducing the incidence of low birth weight in low income countries has substantial economic benefits", *World Bank Research Observer*, 21(1):25-48.
- Almond, Douglas and Janet Currie. 2011. "Killing me softly: The fetal origins hypothesis", *Journal of Economic Perspectives*, 25(3):153-172.
- Auyeung, Tung Wai and Jenny S. W. Lee. 2001. "Estimation of height in older Chinese adults by measuring limb length", *Journal of the American Geriatrics Society*, 49(5):684-685.
- Barker, David. 1994. *Mothers, babies and health in later life*. London: BMJ Publishing Group.
- Bermudez, Odilia I., Emily K. Becker and Katherine L. Tucker. 1999. "Development of Sex-specific Equations for Estimating Stature of Frail Elderly Hispanics living in the Northeastern United States." *American Journal of Clinical Nutrition*, 69: 992-8.
- Campbell, Colin. 1997. "Public Health Efforts in China Before 1949 and Their Effects on Mortality: The Case of Beijing", *Social Science History*, 21(2):179-218.
- Case, Anne, Angela Fertig and Christine Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics*, 24:365 - 389
- Case, Anne and Christine Paxson. 2008a. "Height, Health and Cognitive Function at Older Ages." *American Economic Review, Papers and Proceedings*, 98: 463-7.
- Case, Anne and Christine Paxson. 2008b. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy*, 116 (3):299-332 .

Case, Anne and Christine Paxson. 2009. "Early Life Health and Cognitive Function in Old Age", *American Economic Review Papers and Proceedings*, 99(2):104-109.

Case, Anne, Christine Paxson and Mahnaz Islam. 2009. "Making Sense of the Labor Market Height Premium: Evidence from the British Household Panel Survey." *Economic Letters*, 102: 174-6.

Case, Anne and Christine Paxson. 2010a. "Causes and Consequences of Early Life Health." NBER Working Paper 15637.

Case, Anne and Christine Paxson. 2010b. "The Long Reach of Childhood Health and Circumstance: Evidence from the White Hall II study." NBER Working Paper 15640.

Cheng, H.S., L.C. See and Y. H. Shieh. 2001. "Estimating Stature from Knee height for Adults in Taiwan." *Chang Gung Medical Journal*, 24(9): 547-56.

Chumlea, William C., Alex F. Roche and Maria L.Steinbaugh. 1985. "Estimating Stature from Knee height for persons 60 to 90 years of age." *Journal of American Geriatrics Society* 33(2), 116-20.

Chumlea, William.C. and Shumei Guo. 1992. "Equations for Predicting Stature in white and black Elderly Individuals." *Journal of Gerontology*, 47: M197-203.

Chumlea, W.C., S.S. Guo, K. Wholihan, D. Cockram, R.J. Kuczmarski and C.L. Johnson. 1998. "Stature Prediction Equations for Elderly non-Hispanic White, non-Hispanic Black, and Mexican-American Persons Developed from NHANES III data." *Journal of American Diet Association*, 98:137-42.

Cutler, David and Adriana Lleras-Muney. 2010. "Understanding differences in health behaviors by education", *Journal of Health Economics*, 29(1):1-28

Deaton, Angus and Raksha Arora. 2009. "Life at the top: the benefits of height." *Economics and Human Biology*, 7(2): 133-6.

Elo, Irma and Samuel Preston. 1992. "Effects of early life conditions on adult mortality", *Population Index*, 58(2):186-211.

Fan Jianqing. (1992) "Design-Adaptive nonparametric regression." *Journal of the American Statistical Association*, 87(420): 998-1004.

Finch, Caleb and Eileen Crimmins. 2004. "Inflammatory exposure and historical changes in human life spans", *Science*, 305:1736-1739.

Fogel, Robert. 1986. "Physical growth as a measure of the economic well-being of populations: The eighteenth and nineteenth centuries", in F. Falkner and J.M. Tanner (eds.), *Human growth: A comprehensive treatise, Volume 3*, New York: Plenum Press.

Fogel, Robert. 2004. *Escape from hunger and premature death:1700-2100*. Cambridge: Cambridge University Press.

Folstein, Marshal, Susan Folstein, and Paul McHugh, Mini-mental state: A practical method for grading the cognitive state of patients for the clinician, *Journal of Psychiatric Research*, 12(3) 189-198, 1975.

Glewwe, Paul and Edward Miguel. 2008. "The impact of child health and nutrition on education in less developed countries", in T.Paul Schultz and John Strauss (eds.), *Handbook of Development Economics, Volume 4*, Amsterdam: North Holland Press.

Godfrey, Keith and David Barker. 2000. "Fetal nutrition and adult disease", *American Journal of Clinical Nutrition*, 71(suppl):1344S-1352S.

Haboubi, N.Y., P.R. Hudson and M.S. Pathy. 1990. "Measurement of height in the elderly", *Journal of the American Geriatrics Society*, 8(9):1008-1010.

Heineck Guido. 2009. "Too Tall to be Smart? The Relationship between Height and Cognitive Abilities." *Economics Letters*,105: 78-80.

Hillier, Teresa, Li-Yung Lui, Deborah Kado, ES LeBlanc, Kimberly Vesco, Douglas Bauer, Jane Cauley, Kristine Ensrud, Dennis Black, Marc Hochberg and Steven Cummings. 2012. "Height loss in older women: Risk of hip fracture and mortality independent of vertebral fractures", *Journal of Bone and Mineral Research*, 27(1):153-159.



- Huang, Cheng and Irma T. Elo. 2009. "Mortality of the Oldest Old Chinese: The Role of Early-life Nutritional Status, Socio-economic Conditions, and Sibling Sex-composition." *Population Studies*, 63(1): 7-20.
- Huang, T., M. Carlson, A. Fitzpatrick, L. Kuller, L. Fried and P. Zandi. 2008. "Knee height and arm span: A reflection of early-life environment and risk of dementia", *Neurology*, 70:1818-1826.
- Leung, S., J. Lau, Y. Xu, L. Tse, K. Huen, G. Wong, W. Law, V. Yeung, W. Yeung and N. Leung. 1996. "Secular changes in standing height, sitting height and sexual maturation of Chinese-the Hong Kong growth study 1993", *Annals of Human Biology*, 23(4):297-306.
- Li Edmund, Emily Tang, Cora Wong, Susan Lui, Veronica Chan and David Dai. 2000. "Predicting stature from Knee height in Chinese Elderly Subjects." *Asia Pacific Journal of Clinical Nutrition*, 9: 252-55.
- Knous, Barbara and Masako Arisawa 2002 "Estimation of Height in Elderly Japanese using Region-specific Knee height equations." *American Journal of Human Biology*, 14(3): 300-7.
- Kwok, T and M.N. Whitelaw. 1991. "The use of Armspan in Nutritional Assessment of the Elderly." *Journal of American Geriatrics Society*, 39(5): 492-6.
- Kwok, T., E. Lau and J. Woo. 2002 "The Prediction of height by Armspan in Older Chinese people." *Annals of Human Biology*, 29(6): 649-56.
- Kuh, Diana and Michael Wadsworth. 1989. "Parental height-childhood environment and subsequent adult height in a national birth cohort." *International Journal of Epidemiology*, 18: 663-668.
- McArdle, John. 2010. "Contemporary Challenges of Longitudinal Measurement Using HRS Data", in G. Walford, E. Tucker & M. Viswanathan (eds.), *The SAGE Handbook of Measurement*. London: SAGE Press.
- Martorell, Reynaldo and Jean-Pierre Habicht. 1986. "Growth in early childhood in developing countries", in F. Falkner and J.M. Tanner (eds.), *Human growth: A comprehensive treatise, Volume 3*, New York: Plenum Press.

- Mitchell C.O. and David Lipschitz., 1982. "Arm Length Measurement as an Alternative to Height in Nutritional Assessment of the Elderly." *Journal of Parenteral and Enteral Nutrition*, 6(3): 226.
- Maurer, Jurgen 2010. "Height, education and later life cognition in Latin America and the Caribbean", *Economic and Human Biology*, 8(2):168-176.
- Myers, S.A., S. Takiguchi and M. Yu. 1994. "Stature Estimated from Knee height in Elderly Japanese Americans." *Journal of American Geriatrics Society*, 42(2):157-60
- Nystrom Peck, Maria and Olle Lundberg. 1995. "Short stature as an effect of economic and social conditions in childhood." *Social Science and Medicine*, 41: 733–738.
- Perissinotto, Egle, Claudia Pisent, Giuseppe Sergi, Francesco Grigoletto, Giuliano Enzi and ILSA Working Group. 2002. "Anthropometric Measurements in the Elderly: Age and Gender differences." *British Journal of Nutrition*, 87(2): 177–186.
- Pini, R., E. Tonon, M.C. Cavallini, F. Bencini, M. diBari, G. Masotti and N. Marchionni 2001 "Accuracy of Equations for Predicting Stature from Knee height, and Assessment of Statural Loss in an Older Italian Population." *Journal of Gerontology A Biological Science*, 56(1):B3-7.
- Prothro, J.W. and C.A. Rosenbloom. 1993. "Physical Measurements in an Elderly Black Population: Knee height as the Dominant indicator of Stature." *Journal of Gerontology*, 48: M15–M18.
- Reeves, S., C. Varakamin and C. Henry. 1996. "The relationship between arm-span measurement and height with special reference to gender and ethnicity", *European Journal of Clinical Nutrition*, 50:398-400.
- Ridder, Geert and Robert Moffitt. 2007. "The Econometrics of Data Combination", in J. Heckman and E. Leamer (eds.), *Handbook of Econometrics, Volume 6 part 2*, Amsterdam: North Holland Press.
- Roubenoff, R. and P. Wilson, 1993. "The advantage of knee height over height as an

index of stature in expression of body composition in adults.” *American Journal of Clinical Nutrition* 57(5), 609–613.

Schnaider Beerli, Michal, Michael Davidson, Jeremy Silverman, Shlomo Noy, James Schmeidler, and Uri Goldbourt. 2005. “Relationship between body height and dementia.” *American Journal of Geriatric Psychiatry*, 13: 116–123.

Smith, James P. 2009 “The Impact of Childhood Health on Adult Labor Market Outcomes.” *Review of Economics and Statistics*, 91 (3): 478-489.

Smith, James P., John J. McArdle and Robert Willis. 2010. "Cognition and Economic Outcomes in the Health and Retirement Study, *Economic Journal*, 120(548):F363-F380.

Smith James P., Yan Shen, John Strauss, Zhe Yang and Yaohui Zhao. 2012. “The effects of childhood health on adult health and SES in China”, *Economic Development and Cultural Change*, forthcoming.

Steckel, Richard 1995. “Stature and the standard of living”, *Journal of Economic Literature*, 33(4):1903-1940.

Steckel, Richard 2009. “Heights and human welfare: recent developments and new directions”, *Explorations in Economic History*, 46(1):1-23.

Steele, Marilyn and Joyce Mattox 1987. "Correlation of arm-span and height in young women of two races", *Annals of Human Biology*, 14(5):445-447.

Strauss, John and Duncan Thomas. 2008. "Health over the life course", in T.Paul Schultz and John Strauss (eds.), *Handbook of Development Economics, Volume 4*, Amsterdam: North Holland Press.

Strauss, John, Xiaoyan Lei, Albert Park, Yan Shen, James P. Smith, Zhe Yang, and Yaohui Zhao. 2010. “Health Outcomes and Socio-economic Status Among the Elderly in China: Evidence from the CHARLS Pilot”, *Journal of Population Ageing*, 3(3):111-142.

Thomas, Duncan, John Strauss and Maria-Helena Henriques. 1991. “How Does Mother’s Education Affect Child Height.” *Journal of Human Resources*, 26(2):

183-211.

Van den Berg, Gerard, Maarten Lindeboom and France Portrait. 2006. "Economic conditions early in life and individual mortality." *American Economic Review*, 96: 290–302.

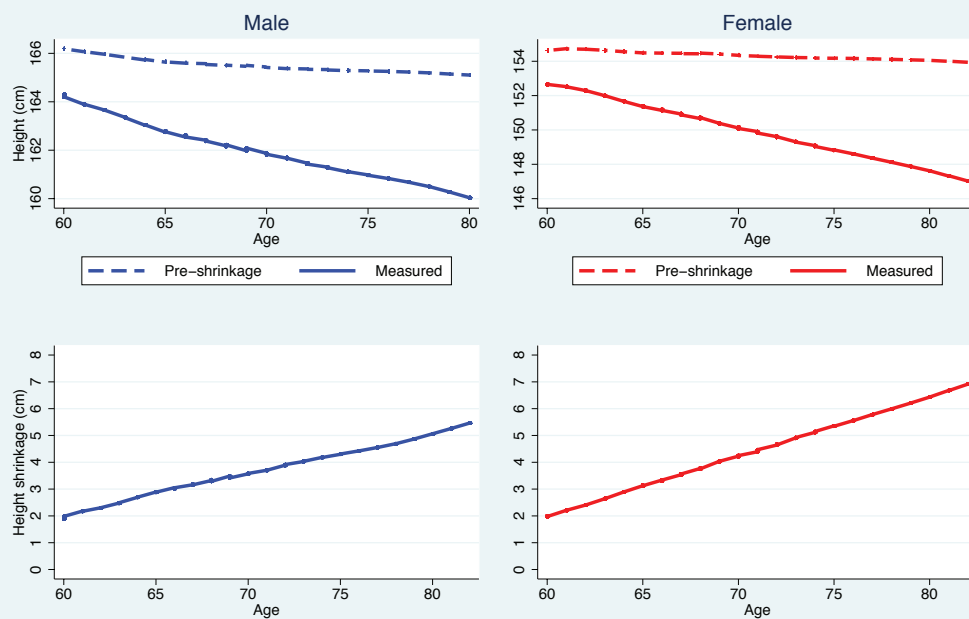
Zeng, Yi, Danan Gu and Kenneth C. Land. 2007. "The Association of Childhood Socioeconomic Conditions with Healthy Longevity at the Oldest-Old Ages in China." *Demography*, 44(3): 497-518

Zhang H, B.H. Hsu-Hage and M.L. Wahlqvist. 1998 "The use of knee height to estimate maximum stature in elderly Chinese." *Journal of Nutrition, Health & Aging*, 2(2):84-7.

Zhang, Zhenmei, Danan Gu and Mark Hayward. 2010. "Childhood nutritional deprivation and cognitive impairment among older Chinese people", *Social Science and Medicine*, 71:941-949.

Zhao, Yaohui, John Strauss, Gonghuan Yang, John Giles, Yisong Hu, and Albert Park. 2012. "The CHARLS User Guide", China Center for Economic Research, Peking University.

Fig 1. Measured height, pre-shrinkage height and height shrinkage



Data source: CHARLS 2011.

Table 1: Summary Statistics

Variable	Male sample			Female sample		
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
<b>Panel A: Younger sample (45 &lt;= Age &lt;= 49)</b>						
Height (cm)	1101	166.35	6.16	1508	155.20	5.86
Upper arm length (cm)	1101	35.20	2.37	1508	32.60	2.20
Lower leg length (cm)	1101	50.00	3.09	1508	46.52	2.98
Age	1101	47.31	1.31	1508	47.25	1.35
Han race	1101	0.94	0.24	1508	0.92	0.27
<b>Panel B: Older sample (Age &gt;= 60)</b>						
<b>Biological Measures and Demographics</b>						
Height (cm)	2953	162.33	6.79	2935	150.65	6.50
Upper arm (cm)	2953	35.22	2.39	2935	32.49	2.20
Lower leg (cm)	2953	49.45	2.92	2935	45.89	2.97
Pre-shrinkage height (cm)	2953	165.62	4.32	2935	154.44	4.05
Height shrinkage (cm)	2953	3.30	5.05	2935	3.80	5.02
Age	2953	68.06	6.41	2935	68.11	6.93
Han race	2953	0.94	0.23	2935	0.93	0.25
<b>Cognitive abilities</b>						
TICS (0-5)	2941	7.17	2.72	2915	5.26	3.15
Words recall (0-10)	2720	3.10	1.67	2619	2.85	1.74
Draw a figure successfully (0-1)	2941	0.64	0.48	2915	0.40	0.49
<b>Biomarkers</b>						
Hypertension (0-1)	2939	0.48	0.50	2921	0.57	0.50
Lung capacity	2799	263.96	113.48	2673	195.17	79.39
Grip strength	2903	32.70	9.04	2847	21.67	7.34
Balance (0-1)	2850	0.75	0.43	2767	0.60	0.49
Walk time	2780	4.42	2.10	2714	5.01	2.68
<b>General health</b>						
Poor general health (0-1)	2953	0.29	0.46	2934	0.36	0.48
Physical function (0-7)	2741	1.13	1.40	2500	1.64	1.59
ADLs (0-6)	2931	0.39	0.99	2907	0.56	1.16
IADLs (0-5)	2937	0.49	1.09	2913	0.75	1.27
CESD (0 - 30)	2773	8.04	5.98	2669	10.26	6.77
Life expectation pessimistic (0-1)	2599	0.33	0.47	2451	0.38	0.49
<b>Adulthood Socio-Economic Status</b>						
Education levels						
Illiterate	2953	0.20	0.40	2935	0.54	0.50
Primary	2953	0.56	0.50	2935	0.35	0.48
Junior	2953	0.16	0.37	2935	0.07	0.26
Senior and above	2953	0.08	0.27	2935	0.03	0.18
Log (Expenditure <i>per capita</i> )	2953	8.35	0.91	2935	8.29	0.93
Married	2953	0.86	0.34	2935	0.71	0.46
Urban	2953	0.18	0.38	2935	0.21	0.41
<b>Childhood Socio-Economic Status</b>						
Urban before 16	2953	0.08	0.28	2935	0.08	0.27
Childhood Health fair and better	2953	0.93	0.25	2935	0.90	0.30
Childhood Health poor	2953	0.06	0.24	2935	0.09	0.29
Childhood health missing	2953	0.01	0.08	2935	0.01	0.09
Father illiterate	2953	0.66	0.48	2935	0.70	0.46
Father literate	2953	0.29	0.46	2935	0.24	0.43
Father education missing	2953	0.05	0.22	2935	0.06	0.24
Mother illiterate	2953	0.91	0.28	2935	0.93	0.26
Mother literate	2953	0.05	0.22	2935	0.04	0.19
Mother education missing	2953	0.04	0.18	2935	0.03	0.18
Father alive before 18	2953	0.87	0.34	2935	0.89	0.31
Father dead before 18	2953	0.12	0.33	2935	0.10	0.30
Father death missing	2953	0.01	0.10	2935	0.01	0.08
Mother alive before 18	2953	0.91	0.29	2935	0.93	0.25
Mother dead before 18	2953	0.07	0.26	2935	0.06	0.23
Mother death missing	2953	0.01	0.12	2935	0.01	0.10

Notes: Data source is CHARLS 2011.

Table 2: Preshrinkage Height Prediction

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Male sample		Measured Height (cm)			
			Female sample			
Upper arm length	0.694*** (0.115)	-0.0819 (1.712)	0.563 (0.381)	0.776*** (0.0946)	-1.993* (1.169)	0.498* (0.276)
Lower leg length	0.996*** (0.105)	-6.689*** (1.245)	0.952*** (0.325)	0.931*** (0.0735)	-4.116*** (1.107)	0.938*** (0.182)
Upper arm length square		0.00268 (0.0240)			0.0561** (0.0226)	
Lower leg length square		0.0761*** (0.0122)			0.0632*** (0.0149)	
Interaction of arm length and leg length		0.00982 (0.0330)			-0.0196 (0.0221)	
Han race	0.397 (0.564)	0.716 (0.589)	0.367 (0.587)	1.056** (0.426)	1.260*** (0.408)	0.896** (0.430)
Age dummies						
45 (Reference)						
46			1.678 (15.93)			-11.56 (10.30)
47			-7.384 (17.57)			-15.77 (11.73)
48			-15.24 (16.26)			-9.856 (9.858)
49			-12.63 (16.60)			-13.19 (10.15)
Age dummies interacting upper arm length						
45 & Arm (Reference)						
46 & Arm			-0.146 (0.417)			0.217 (0.311)
47 & Arm			0.0730 (0.462)			0.402 (0.337)
48 & Arm			0.0520 (0.433)			-0.0234 (0.323)
49 & Arm			0.708 (0.446)			0.732** (0.353)
Age dummies interacting lower leg length						
45 & Leg (Reference)						
46 & Leg			0.0703 (0.363)			0.0828 (0.221)
47 & Leg			0.102 (0.420)			0.0468 (0.252)
48 & Leg			0.258 (0.374)			0.220 (0.225)
49 & Leg			-0.243 (0.397)			-0.246 (0.239)
Constant	91.78*** (4.536)	291.4*** (47.90)	98.59*** (13.81)	85.63*** (3.263)	243.2*** (19.51)	94.92*** (8.104)
Observations	1101	1101	1101	1508	1508	1508
R-square	0.443	0.518	0.458	0.452	0.513	0.465
F test for All limbs	140.3	136.5	12.14	228.6	180.3	30.56
P Value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
F tests - Limbs quadratic terms		16.90			23.25	
P Value		[0.000]			[0.000]	
F test - Age dummies plus interactions with limbs			1.182			1.242
P Value			[0.291]			[0.248]

Notes: Data source is CHARLS 2011. Sample used are those aged between 45 and 49. Coefficients in Columns (2) and (5) are used to predict pre-shrinkage height in older sample (60+).

Table 3: Height Comparison with CHNS 1991

Age in 2011	CHNS 1991			CHARLS 2011			Differences	
	MH	# Obs.	Age in 1991	MH	PH	# Obs.	MH 2011 - MH 1991	PH 2011 - MH 1991
<b>Panel A: Male sample</b>								
60	165.64 [5.667]	106	40	163.05 [6.048]	165.50 [5.119]	91	-2.59	-0.14
61	166.71 [5.819]	84	41	164.85 [5.947]	166.58 [4.627]	85	-1.86	-0.13
62	165.66 [6.391]	82	42	162.66 [5.992]	165.70 [3.398]	114	-3.00	0.04
63	165.43 [5.609]	75	43	163.94 [6.568]	166.35 [5.261]	82	-1.50	0.92
64	164.81 [7.259]	79	44	163.39 [7.350]	165.57 [4.422]	97	-1.41	0.76
65	163.76 [6.708]	98	45	163.69 [7.135]	165.90 [4.766]	76	-0.06	2.14
66	164.16 [6.396]	62	46	162.52 [6.664]	164.95 [4.103]	61	-1.64	0.79
67	163.36 [5.885]	57	47	161.47 [7.655]	164.50 [5.381]	63	-1.90	1.13
68	163.94 [6.415]	59	48	161.50 [6.462]	164.76 [3.829]	70	-2.45	0.82
69	163.77 [6.453]	63	49	160.59 [6.530]	165.58 [4.361]	63	-3.17	1.81
<b>Panel B: Female sample</b>								
60	154.60 [5.886]	104	40	151.90 [5.801]	154.57 [4.125]	123	-2.71	-0.03
61	153.98 [5.899]	101	41	151.68 [6.051]	154.47 [3.639]	105	-2.29	0.50
62	154.19 [4.964]	97	42	152.93 [5.596]	154.71 [3.753]	94	-1.27	0.52
63	154.03 [5.859]	89	43	151.97 [6.942]	154.93 [4.140]	74	-2.06	0.90
64	155.49 [6.032]	80	44	152.04 [6.080]	154.17 [3.636]	98	-3.45	-1.31
65	152.16 [6.189]	86	45	150.09 [6.471]	153.72 [4.056]	76	-2.08	1.56
66	153.61 [5.869]	61	46	150.50 [7.179]	154.17 [4.315]	69	-3.11	0.55
67	153.28 [4.554]	58	47	148.84 [5.879]	154.44 [4.391]	60	-4.44	1.16
68	153.49 [5.383]	69	48	148.23 [5.558]	153.67 [4.207]	51	-5.26	0.18
69	153.96 [6.537]	67	49	150.39 [5.903]	153.94 [4.024]	59	-3.57	-0.03

Notes: MH - Measured Height; PH - Pre-shrinkage Height. Standard deviations are in brackets. Unit of heights is centimeters. CHNS 1991 was fielded in eight provinces: Liaoning, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou. The CHARLS data are from the same provinces.



Table 4: Preshrinkage height and SES

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Male Sample		Height shrinkage (cm)		Female Sample	
Pre-shrinkage height (cm)		0.007 (0.0275)			0.044 (0.0283)	
Lower leg length			-0.104** (0.0466)			0.066 (0.0467)
Upper arm length			0.053 (0.0576)			-0.033 (0.0622)
Education level						
Illiterate (Reference)						
Primary	-0.928*** (0.2375)	-0.925*** (0.2376)	-0.939*** (0.2373)	-0.594*** (0.2044)	-0.603*** (0.2043)	-0.599*** (0.2044)
Junior	-1.597*** (0.3144)	-1.597*** (0.3145)	-1.589*** (0.3141)	-0.976** (0.3866)	-1.012*** (0.3865)	-0.996** (0.3866)
Senior and above	-1.922*** (0.4042)	-1.920*** (0.4042)	-1.948*** (0.4038)	-1.728*** (0.5571)	-1.743*** (0.5568)	-1.749*** (0.5569)
Log expenditure <i>per capita</i>	-0.217** (0.1045)	-0.218** (0.1046)	-0.204* (0.1045)	-0.086 (0.1018)	-0.091 (0.1017)	-0.090 (0.1017)
Married	-0.334 (0.2547)	-0.335 (0.2547)	-0.321 (0.2544)	-0.024 (0.2041)	-0.032 (0.2040)	-0.022 (0.2040)
Urban	-0.858** (0.3487)	-0.862** (0.3488)	-0.835** (0.3483)	-1.099*** (0.3159)	-1.118*** (0.3158)	-1.119*** (0.3159)
Han race	-0.141 (0.6594)	-0.142 (0.6583)	-0.148 (0.6590)	0.641 (0.5853)	0.589 (0.5802)	0.636 (0.5851)
Urban Area before 16 year-old	-0.347 (0.4020)	-0.347 (0.4021)	-0.344 (0.4016)	-0.384 (0.3880)	-0.411 (0.3879)	-0.404 (0.3880)
Childhood health poor	0.031 (0.3555)	0.034 (0.3556)	0.006 (0.3552)	0.513* (0.2939)	0.526* (0.2937)	0.534* (0.2938)
Father's literate	0.074 (0.2023)	0.072 (0.2023)	0.095 (0.2020)	0.074 (0.2145)	0.064 (0.2144)	0.064 (0.2145)
Mother's literate	-0.062 (0.4065)	-0.061 (0.4066)	-0.070 (0.4061)	0.160 (0.4676)	0.174 (0.4673)	0.199 (0.4678)
Father's dead before 18	0.163 (0.2651)	0.162 (0.2652)	0.187 (0.2649)	0.423 (0.2896)	0.445 (0.2896)	0.434 (0.2896)
Mother's death before 18	0.424 (0.3288)	0.421 (0.3289)	0.450 (0.3282)	0.211 (0.3742)	0.198 (0.3740)	0.214 (0.3741)
Age categories						
60 - 64 (Reference)						
65 - 69	0.760*** (0.2156)	0.763*** (0.2156)	0.737*** (0.2154)	1.085*** (0.2159)	1.103*** (0.2159)	1.093*** (0.2160)
70 - 74	0.886*** (0.2418)	0.887*** (0.2418)	0.877*** (0.2415)	1.653*** (0.2488)	1.656*** (0.2487)	1.662*** (0.2488)
75 - 79	1.606*** (0.2956)	1.612*** (0.2958)	1.581*** (0.2954)	3.180*** (0.3031)	3.210*** (0.3031)	3.204*** (0.3032)
80 - 84	2.335*** (0.4253)	2.338*** (0.4253)	2.326*** (0.4248)	3.336*** (0.4124)	3.362*** (0.4122)	3.366*** (0.4124)
85 +	3.087*** (0.7180)	3.091*** (0.7181)	2.984*** (0.7176)	5.646*** (0.5770)	5.658*** (0.5767)	5.677*** (0.5770)
Constant	2.927 (3.1213)	1.749 (5.5651)	6.295* (3.8560)	2.391 (2.12437)	-4.240 (4.8218)	0.484 (2.9550)
Observations	2,953	2,953	2,953	2,935	2,935	2,935
R-squared	0.274	0.274	0.276	0.285	0.286	0.286
Birth province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Current county dummies	Yes	Yes	Yes	Yes	Yes	Yes
<b>Wald tests</b>						
Adult SES variables	58.384	58.491	57.288	36.154	37.561	37.342
P value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Age category dummies	61.127	61.339	59.321	192.519	194.947	195.057
P value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Childhood SES variables	7.118	7.057	7.441	10.456	10.894	11.050
P value	[0.789]	[0.794]	[0.762]	[0.490]	[0.452]	[0.439]
Birth province dummies	30.290	30.286	30.377	39.145	39.955	40.479
P value	[0.065]	[0.065]	[0.064]	[0.006]	[0.005]	[0.004]
Current county dummies	448.144	446.369	455.192	376.873	368.428	370.915
P value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Upper arm and lower leg			5.058			2.076
P value			[0.080]			[0.354]

Notes: Missing dummies are added if available. Standard errors adjusted for predicted variables are in parentheses. In Wald tests, adult SES variables include urban, married, education levels and income per capita, childhood SES variables include living in urban area before 16 year-old, childhood health status, parents' education, parents' death before 18 years old.

[illegible]

Table 5.1 (continued): Height Shrinkage, Pre-shrinkage Height and Cognitive Abilities

Dependent variable	(9)	(10)	(11)	(12)
	OLS	2SLS	OLS	OLS
<b>Panel A: Male Sample</b>				
Measured height	0.0049*** (0.00130)	0.0020 (0.00191)		
Pre-shrinkage height			0.0018 (0.00195)	0.0018 (0.00195)
Height shrinkage				-0.0074*** (0.00176)
Observations	2941	2941	2941	2941
R-square	0.199	0.197	0.195	0.200
<b>Wald tests</b>				
Adult SES variables			364.191	338.556
P value			0.000	0.000
Age category dummies			49.160	41.184
P value			0.000	0.000
Childhood SES variables			25.231	24.794
P value			0.008	0.010
Birth province dummies			19.613	17.694
P value			0.482	0.608
Current county dummies			301.732	303.959
P value			0.000	0.000
<b>T-test: MH + PH = 0</b>				2.148
P value				0.032
<b>Panel B: Female sample</b>				
Measured height	0.0045*** (0.00148)	0.0056*** (0.00211)		
Pre-shrinkage height			0.0052*** (0.00204)	0.0054*** (0.00203)
Height shrinkage				-0.0038*** (0.00174)
Observations	2915	2915	2915	2915
R-square	0.224	0.224	0.223	0.225
<b>Wald tests</b>				
Adult SES variables			386.200	375.602
P value			0.000	0.000
Age category dummies			40.433	31.892
P value			0.000	0.000
Childhood SES variables			30.168	30.613
P value			0.001	0.001
Birth province dummies			27.982	26.282
P value			0.110	0.157
Current county dummies			401.491	397.046
P value			0.000	0.000
<b>T-test: MH + PH = 0</b>				0.594
P value				0.552
County dummies	Yes	Yes	Yes	Yes

Notes: Data source is CHARLS 2011. Columns (1), (5) and (9) show OLS results with robust standard errors are in parentheses. Columns (2), (6) and (10) show 2SLS results with instrumental variables being upper arm length, lower leg length, quadratics and the interaction between upper arm and lower leg length. Robust standard errors are in parentheses. In other columns, standard errors adjusted for predicted variables are in parentheses. All regressions include adult SES variables, age category dummies, childhood SES variables, birth province dummies and current county dummies. Missing dummies are added, if available. In Wald tests, adult SES variables include urban, married, education levels and log expenditure per capita; childhood SES variables include living in urban area before 16 year-old, childhood health status, parents' education, parents' death before 18 years old.

[illegible]

[illegible]

County dummies

Table 5.2 (continued): Height Shrinkage, Pre-shrinkage Height and Biomarkers

Dependent variable	(17)	(18)	(19)	(20)
	OLS	2SLS	OLS	OLS
<b>Panel A: Male Sample</b>				
Measured height	-0.0142** (0.00621)	-0.0148* (0.00785)		
Pre-shrinkage height			-0.0117 (0.00852)	-0.0118 (0.00852)
Height shrinkage				0.0164** (0.00788)
Observations	2780	2780	2780	2780
R-square	0.093	0.093	0.092	0.093
<b>Wald tests</b>				
Adult SES variables			35.320	31.917
P value			0.000	0.000
Age category dummies			113.240	104.349
P value			0.000	0.000
Childhood SES variables			4.409	4.213
P value			0.956	0.963
Birth province dummies			27.691	27.537
P value			0.117	0.121
Current county dummies			936.804	941.687
P value			0.000	0.000
<b>T-test: MH + PH = 0</b>				0.395
P value				0.693
<b>Panel B: Female sample</b>				
Measured height	-0.0154 (0.00947)	-0.0220* (0.01234)		
Pre-shrinkage height			-0.0198 (0.01240)	-0.0202 (0.01240)
Height shrinkage				0.0119 (0.01062)
Observations	2714	2714	2714	2714
R-square	0.107	0.107	0.107	0.107
<b>Wald tests</b>				
Adult SES variables			14.527	13.490
P value			0.0242718	0.0358841
Age category dummies			157.251	142.195
P value			0.000	0.000
Childhood SES variables			18.072	18.265
P value			0.080	0.076
Birth province dummies			29.289	29.601
P value			0.082	0.077
Current county dummies			709.213	708.750
P value			0.000	0.000
<b>T-test: MH + PH = 0</b>				0.512
P value				0.609
County dummies	Yes	Yes	Yes	Yes

Notes: Data source is CHARLS 2011. Columns (1), (5), (9), (13) and (17) show OLS results with robust standard errors are in parentheses. Columns (2), (6), (10), (14) and (18) show 2SLS results with instrumental variables being upper arm length, lower leg length, quadratics and the interaction between upper arm and lower leg length. Robust standard errors are in parentheses. In other columns, standard errors adjusted for predicted variables are in parentheses. All regressions include adult SES variables, age category dummies, childhood SES variables, birth province dummies and current county dummies. Missing dummies are added, if available. In Wald tests, adult SES variables include urban, married, education levels and log expenditure per capita; childhood SES variables include living in urban area before 16 year-old, childhood health status, parents' education, parents' death before 18 years old.

**Table 5.3: Height shrinkage, Pre-shrinkage height and Health**

[illegible]

[illegible]

County dummies

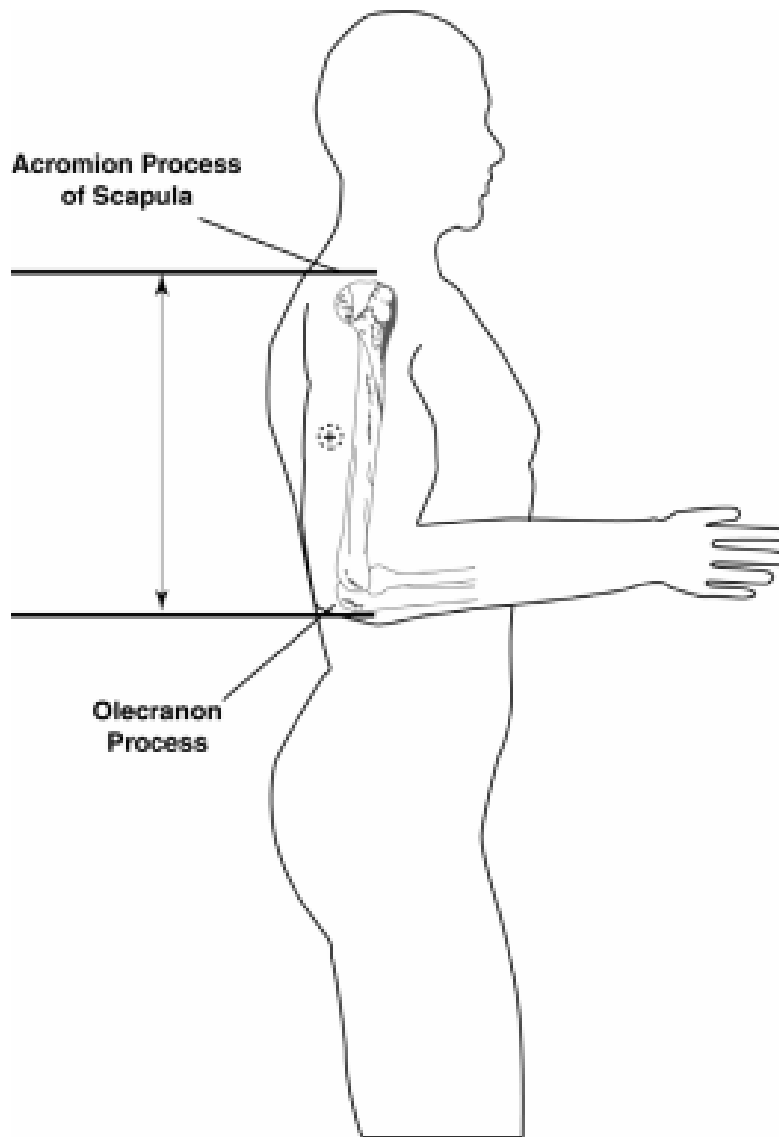


Table 5.3 (continued): Height shrinkage, Pre-shrinkage height and Health

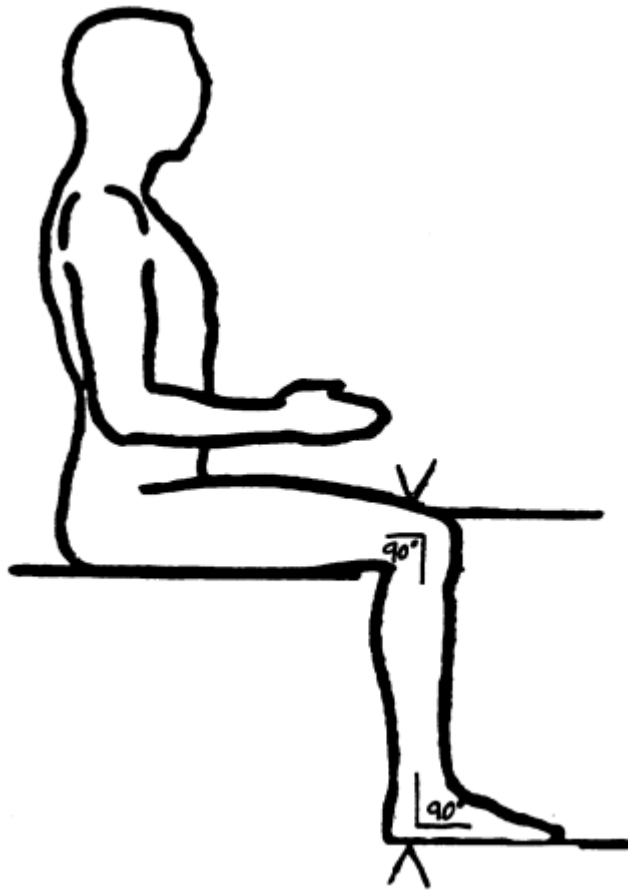
Dependent variable	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
	CESD (0 - 30)		Life expectation pessimistic (0 - 1)					
	OLS	2SLS	OLS	OLS	OLS	2SLS	OLS	OLS
<b>Panel A: Male Sample</b>								
Measured height	-0.0456** (0.01748)	-0.0409 (0.02681)			-0.0023 (0.00147)	-0.0002 (0.00215)		
Pre-shrinkage height			-0.0337 (0.02731)	-0.0346 (0.0273)			-0.0005 (0.00223)	-0.0006 (0.00223)
Height shrinkage				0.0547** (0.02480)				0.0037* (0.00202)
Observations	2773	2773	2773	2773	2599	2599	2599	2599
R-square	0.046	0.046	0.045	0.046	0.078	0.077	0.077	0.078
<b>Wald tests</b>								
Adult SES variables			63.839	58.753			31.081	27.999
P value			0.000	0.000			0.000	0.000
Age category dummies			7.206	6.341			94.309	87.783
P value			0.206	0.274			0.000	0.000
Childhood SES variables			7.357	7.132			10.608	10.563
P value			0.770	0.788			0.477	0.481
Birth province dummies			23.468	23.301			23.580	23.670
P value			0.266	0.274			0.261	0.257
Current county dummies			301.109	300.828			225.756	224.294
P value			0.000	0.000			0.000	0.000
<b>T-test: MH + PH = 0</b>				0.548				1.061
P value				0.584				0.289
<b>Panel B: Female sample</b>								
Measured height	-0.0018 (0.02144)	0.0396 (0.03317)			-0.0009 (0.00180)	0.0034 (0.00265)		
Pre-shrinkage height			0.0379 (0.03358)	0.0362 (0.03365)			0.0035 (0.00255)	0.0032 (0.00256)
Height shrinkage				0.0297 (0.02901)				0.0038* (0.00218)
Observations	2669	2669	2669	2669	2451	2450	2451	2451
R-square	0.041	0.040	0.042	0.042	0.065	0.062	0.066	0.067
<b>Wald tests</b>								
Adult SES variables			28.776	27.606			13.282	11.460
P value			0.000	0.000			0.039	0.075
Age category dummies			1.881	1.569			72.261	62.713
P value			0.865	0.905			0.000	0.000
Childhood SES variables			20.490	20.241			15.126	14.956
P value			0.039	0.042			0.177	0.185
Birth province dummies			29.943	29.884			22.615	22.588
P value			0.071	0.072			0.308	0.309
Current county dummies			343.272	342.989			239.585	237.877
P value			0.000	0.000			0.000	0.000
<b>T-test: MH + PH = 0</b>				1.521				2.153
P value				0.128				0.031
County dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Data source is CHARLS 2011. Columns (1), (5), (9), (13), (17) and (21) show OLS results with robust standard errors are in parentheses. Columns (2), (6), (10), (14), (18) and (22) show 2SLS results with instrumental variables being upper arm length, lower leg length, quadratics and the interaction between upper arm and lower leg length. Robust standard errors are in parentheses. In other columns, standard errors adjusted for predicted variables are in parentheses. All regressions include adult SES variables, age category dummies, childhood SES variables, birth province dummies and current county dummies. Missing dummies are added, if available. In Wald tests, adult SES variables include urban, married, education levels and log expenditure per capita; childhood SES variables include living in urban area before 16 year-old, childhood health status, parents' education, parents' death before 18 years old.

## Appendix 2- Measurement of upper arm and lower leg lengths



UPPER ARM LENGTH MEASUREMENT



LOWER LEG MEASUREMENT