Stature and Sibship: Historical Evidence

By

Timothy J. Hatton

(University of Essex and Australian National University)

ABSTRACT

This paper examines historical evidence for a quality-quantity trade-off between sibship size and height as an indicator of health. The existing literature has focused more on education than on health and has it produced mixed results. Historical evidence is limited by the lack of household level data with which to link an individual's height with his or her childhood circumstances. Nevertheless a few recent studies have shed light on this issue. Evidence for children in interwar Britain and for soldiers born in the 1890s who enlisted in the British army at the time of WW1 is reviewed in detail. Both studies support the idea of a significant trade-off, partly due to income dilution and partly because, in these settings, large families were a conduit for infection. Evidence from country-level time series is consistent with this view. The fertility decline that began in the late nineteenth century made a modest but nevertheless significant contribution to the overall increase in heights during the following half-century.

JEL Codes: I12, I38, N24.

Keywords: Health and height; family size; quality-quantity trade-off

Acknowledgements

I am grateful for comments from participants at the Third CEPR Economic History Symposium, Oslo and for detailed comments from Stefan Öberg and two anonymous referees. I am also grateful for funding from the Australian Research Council Discovery Project DP140103603.

Address for correspondence

Timothy J. Hatton
Department of Economics
University of Essex
Wivenhoe Park
Colchester CO4 3SQ, UK
Email: hatton@essex.ac.uk
Tel: +441206872182

1. Introduction

It is often suggested by historians that large families such as those observed in nineteenth century Europe, or more recently in developing countries, produce negative outcomes for children (Dribe et al., 2012). If so, then long run declines in fertility may be one reason among many for improved outcomes. These outcomes include education and cognitive ability as well as a range of different dimensions of health. Inspired by the early work of Becker and Lewis (1973) and Becker and Tomes (1976) economists and other social scientists have investigated the tradeoff between the quantity and average quality of children. The results have been mixed and they seem to be specific to the particular setting, the outcome variable of interest, and the way in which the model is specified and estimated. Most of the focus has been on the educational outcomes for children in terms of grade levels or academic attainment. Here I focus on health outcomes as measured by height. This has obvious historical relevance and its links with economic growth have recently been highlighted (Floud et al., 2011; Weil, 2007).

Height has been widely investigated in historical settings, at first as a proxy for absent data on the standard of living, and later in its own right as an indicator of health and physical well-being. Most of the analysis has focused on long run trends in heights in order to chart progress in particular countries and regions (see Steckel, 1995, 2009; Baten and Blum, 2012). Many of the datasets used come from the records of army recruits, institutions such as prisons, or records of transported convicts or slaves. These often provide details of where the person was born and perhaps details of his previous occupation or that of his father. It is therefore possible to measure the effects of locality, whether rural or urban, and to include other indicators of local conditions. But the records rarely include demographic detail of the household in which the individual grew up. Thus, with a few exceptions, it has not been possible to examine the relationship between an individual's height and the number of siblings in sufficient detail to assess the quality-quantity tradeoff with any precision.

This paper is divided into four main sections. The first gives an overview of the quality-quantity trade off using various measures of child quality. I argue that for several reasons one might expect to find a stronger tradeoff between the number of siblings and height, as compared with other outcomes such as education. One might also expect the effects to be stronger where families are large and incomes relatively low. The next two sections focus on historical evidence from Britain. The first examines the heights of individual children in poor families in the 1930s, as presented by Hatton and Martin (2010a). It considers the possible biases in the estimation of the effect of sibship size on height. It also examines the effects of the child's birth order on his/her height. The following section examines the heights of soldiers enlisting in the British army around the time of the First World War, drawing on Bailey et al. (2015). In order to capture the structure of the household in which they grew up, these servicemen are identified as children in the 1901 census. The effects of local conditions are also examined and the interaction between sibship size and the local disease environment is tested. The final section assesses the degree to which the fertility transition contributed to improving health and increasing height from the late nineteenth century to the mid-twentieth century. It also presents national-level evidence on the effects of the fertility transition for a range of European countries, drawing on Hatton (2014). Overall the results indicate that sibship size had significant effects on height but that its contribution to the long run increase in stature is modest.

2. The quality-quantity tradeoff

The basic model

The idea of an inverse relationship between the quality and the quantity of children proposed by Becker (1960) was explored in pioneering studies by Becker and Lewis (1973) and Becker and Tomes (1976) and it has been examined by many others since. The basic idea is that forward-looking households decide on the number of children that they will have and how much to invest in the 'quality' of each child. The parents value both quality and quantity but family resources are limited. So in the light of the cost of additional children and cost of investment in quality they choose the optimal combination, given the overall resource constraint. A formal derivation of the basic theoretical model and a number of variants of it can be found in Ermisch (2003, Ch. 6).

This trade-off is depicted in Figure 1. The curve Z_1, derived from the overall budget constraint, represents that part of family resources devoted to children, including the time of the parents, and it illustrates that the larger the number of children, *N*, the lower the average quality of those children, $\bar{q}$. Its slope reflects the relative 'price' of quantity to quality. The family's preferences are reflected by indifference curve, U_1, and so the optimal combination for that family is given at point A. A different family devoting the same resources to children but with a steeper indifference curve might choose the combination A', which would mean more children and lower average quality.

The comparison of points like A and A' could be used to measure the quality-quantity trade-off, which should in principle reflect their relative price. However another family, that might otherwise have chosen the combination A', could choose to spend a larger share of its resources on children and locate at point B on the higher budget line, Z_2. Comparing points A and B, the trade-off between quality and quantity will appear less steep. Alternatively the family devoting a larger share of resources to children might locate at B' in which case the correlation between quantity and quality would appear to be positive. It is important to note that quality and quantity are chosen jointly, even though the investment in quality takes place subsequent to the child's birth. This means that the number of children cannot be taken as an independent 'cause' of child quality. In general, differences in the combinations of quality and quantity chosen by different families could arise from a variety of unobserved characteristics that influence both.

Much of the literature on the quality-quantity trade-off has been devoted to finding an exogenous source of variation in the number of children in order to identify (or at least test for) an underlying inverse relationship between sibship size and the outcomes for children. The reason that this is such a challenge is that in most settings it is necessary to look across different families. A number of studies have followed Rozenzweig and Wolpin (1980) in using the incidence of twin births on the grounds that this is an exogenous source of variation to family size. The idea is that a twin birth is a random event that increases family size but is not correlated with family choice or circumstance.[1] Other family-level instruments that have been used include the advent of miscarriages, the sex composition of lower parity children and the mother's sibship size. An alternative approach is to focus on external influences that affect fertility, such as family planning programmes, most notably China's one-child policy. [2]

Different outcomes and different settings

There is now a large and diverse literature that seeks to identify the quality-quantity trade-off and it focuses on various outcomes for the child in a variety of different settings. Most of the attention has been lavished on the child's education as the outcome. This has been motivated by the strong and universal link between education and earnings, as well as by theories of economic growth and development that have placed increasing weight on the role of education

[1] Two qualifications are in order. The first is that the greater the number of births that a family chooses, the greater the likelihood of having a multiple birth; therefore the probability of a twin birth depends on sibship size. One solution is to utilise twin births occurring only at a particular parity such as first birth. The other point is that twin births are rare (about one in eighty) and they may have a limited impact on final sibship size as families adjust subsequent fertility. So the overall leverage exerted by twins on sibship size may be slight.

[2] Although not considered in detail here, the issue of timing may also be important. At what point in the child's lifetime is sibship size observed? Do younger siblings matter more than older siblings? And at what point do we observe the measure of quality? While it might be possible to look at what happened over time as family size increases, few studies have done this.

in fostering enterprise and innovation (Galor, 2011). Other measures include test scores, cognitive ability, disease incidence and subjective or objective measures of overall heath at various points in the lifecycle and the variable of most interest here, height. The settings also vary with some focusing on relatively developed countries where families are small and others on poor countries where families are typically large.

In a landmark study of low income families in Gary Indiana, Hanusheck (1992) found that children's test scores were inversely related to the number of children in the family but not to their birth order. Among the more prominent studies is that of Black et al (2005) which examined a very rich dataset for the entire population Norway, using twins as the instrument and years of education as the outcome. By contrast, these authors found no evidence of a sibship size effect but some evidence that those with higher birth orders had significantly less education.[3] Similarly Angrist et al. (2010) using twins and sex composition as instruments found no evidence of an effect of sibship size on education or other adult outcomes in Israel. For the United States others have found evidence that sibship size had a negative effect on the probability of attending private school but not on outcomes such as grade retention (Conley and Glauber 2005; Cáceres-Delpiano, 2006).

A number of studies have focused on developing countries. For Brazil, using twins as an instrument for sibship size, Ponczek and Portela Souza (2012) find that larger sibship reduces child labour and increases school attendance. In Korea where family sizes are typically small Lee (2007) finds, using the sex of the first born as an instrument for family size, that sibship size reduces the family's per-child expenditure on education, especially for the largest families. Exploiting the proximity to family planning facilities, Dang and Rogers (2012) find that larger sibship reduces the incidence of private tutoring. Particular attention has been paid to China, using as instruments either twins or variations in the application of the one child policy. Thus Li et al. (2008) and Rosenzweig and Zhang (2009) find negative effects of family size on children's education, especially in locations with poor public education systems. On the other hand Qian (2009) finds that, in rural areas, the addition of a second child has a positive effect on the school enrolment of firstborns.

[3] Although widely cited, this result presents something of a puzzle: if the underlying model is one of resource constraints, then family size should matter whereas birth order may or may not matter depending on whether or not scarce resources are unequally allocated between siblings. But it is hard to understand why the allocation between siblings should matter while the overall resource constraint does not.

Studies of educational inputs and outcomes have produced mixed results. Apart from the particular methodology used, it seems likely that there are three reasons for this. One is that in developed countries with small families (such as Norway) the household budget constraint is less binding and there is more scope for substitution away from other expenditures towards children, or perhaps adjustment on other margins such as labour supply. Second, education, if not completely free, is heavily subsidised and cost to the family may be minimal. Hence the results are stronger when the focus is on discretionary expenditures. Third, education is acquired over a long period much of which may post-date the child's early life experience, and which therefore becomes more a decision of the individual and less a decision of his or her parents. By contrast health outcomes, and particularly height, are largely influenced by conditions within the household and particularly in the first few years of life (Cole 2003). It may therefore depend more on inputs that are less easily substitutable, such as parental time,.

Increasing attention has been paid to the effect of sibship size on height and other measures of health. Early studies by Horton (1986, 1988) suggested negative effects for both sibship size and birth order on the heights of children in the Philippines. Desai (1992) found negative sibship size effect on child heights in Latin America but not in Africa, but Alderman (1990) found some evidence of such effects for Ghana. As with the education literature, the early studies did not account for the endogeneity of sibship size. Focusing on Romania under the Ceausescu regime, and using twins at first birth as the instrument, Glick et al. (2007) found a negative relationship between sibship size and the heights of children under five. For East Germany Baten and Böhm (2010) find large negative sibship effects on the heights of 5 and 6 year-olds but not using instrumental variables. A particularly interesting study is that of Liu (2014) using different dimensions of the one child policy in China as instruments for sibship size. He finds that sibship size has a strong negative effect on height but not on educational attainment. The former may be explained in the context of modest living standards (in 1993) even though families are small; the latter by the fact that post-compulsory education is heavily subsidised.

Historical studies

We might expect the results from today's developed countries in the past to resemble more closely those of developing countries in the present, as families were larger and living standards much lower than today. And all the more so for times when public provision of education and health services was largely absent. That might be especially true for health prior to the advent of modern medical advances and when public understanding of the consequences for children of basic hygiene and nutrition was in its infancy. As in present-day developing countries the interaction between household behaviour and community infrastructure might also be important.

Not surprisingly, there are fewer historical studies that attempt to estimate the quality-quantity trade-off, especially as we move further back in time. However, using family reconstitution data for males born in England between 1690 and 1814, Klemp and Weisdorf (2012) find that larger sibship size is associated with lower skills, as measured by occupation, and lower literacy, as observed by the (in)ability to sign marriage registers.[4] Sibship size is instrumented using the interval between marriage and first birth, which is interpreted as a proxy for fecundity. Using county-level data from the Prussian census of 1849, Becker et al. (2010) find an inverse relationship between sibship size and school attendance. Studies by Bleakley and Lange (2009) for the US South in 1910 and by Fernihough (2011) for Belfast and Dublin in 1911 find causal effects on school attendance that support the quality-quantity trade off. Interestingly, Parman (2015) finds that the education of World War II servicemen is negatively associated with family size and positively related to height. Thus the few studies that do exist suggest strong historical links between sibship size and education, perhaps because the families concerned were poor by modern standards and, with little public provision, education and skills were costly to acquire. In the absence of compulsory schooling beyond a minimal level families exercised a wide range of educational choice.

There are even fewer historical studies focusing on height as the outcome. Weir (1993) and Schneider (1996) explored variations over time in national or regional data on height and fertility. But analysis at the household level has been hampered by data limitations. On one hand census and vital registration data does not generally include measures of health and especially not height. On the other hand data on heights from sources like military or prison records generally provide very little evidence of childhood circumstances such as sibship size. An important exception is Öberg (2015), in which demographic data for five parishes from the Scanian Economic Demographic Database are linked to adult heights from conscript inspection records of men born between 1797 and 1950. The results indicate that the negative association

[4] At first sight this contrasts with the well-known finding of Clark (2005) that, in preindustrial times, family size was positively correlated with wealth and human capital. However this can be interpreted as a comparison between points A and B' in Figure 1, where the shift of the budget line from Z_1 to Z_2 is interpreted as an increase in lifetime income. Hence it may not be inconsistent with a quality-quantity trade-off if differences in income and the endogeneity of tastes are taken into account.

between sibship size and height was strongest in the late nineteenth century but became insignificant after the early 20th century. In the following sections I elaborate on findings such as these with evidence from Britain for cohorts born in the interwar period and in the 1890s.

3. Evidence from Interwar Britain

Several studies have looked at the quality-quantity trade off in education in postwar Britain. A study of a 1946 birth cohort found that, in the presence of other socioeconomic variables, birth order and the number of younger siblings were negatively associated with adult height (Kuh and Wadsworth, 1989).[5] Similar results were found in a 1958 birth cohort and their offspring but the effects were weaker for the younger generation than for their parents (Li and Power, 2004). So the effects seem to have become weaker over the postwar period but until fairly recently there has been very little evidence for earlier periods.

The Boyd Orr Cohort

The survey directed by Sir John Boyd Orr in the 1930s is the only one dating from that period (or earlier) that contains evidence on the heights of children together with other household characteristics (Rowett Institute, 1955). It covered 1343 households with school-aged children in 16 locations in England and Scotland during the years 1937-9. These were intended to be geographically representative and the survey was targeted to over-represent low income households. It appears that these households were contacted through schools and, probably as a result, the survey contained a disproportionate number of large families. The survey contains a variety of information about the household's circumstances and living conditions, although some of the information is sketchy. In addition, the survey included details of a medical examination of the children, which was conducted at the schools. The medical survey did not cover all households, nor did it cover each child in the families that were included in the household survey.

Overall there are useable observations from medical examinations of 2946 children in 1131 households. The means from these data are presented in Table 1, where the first column is the average across all individuals and the second column is the average of the mean values for each household (Hatton and Martin 2010a, p. 167). The average age of children in the sample is just under 8, the average birth order is 2.8 and a little over half are female.[6] The number of children

[5] Earlier studies of height at age seven in the 1946 cohort include Goldstein (1971) and Fogelman (1975), both of which found some evidence of a negative effect for the number of siblings.

[6] These children range in age from 2 to 14; they are fairly evenly distributed across the age range, with 33 percent aged 2-5, 37 percent aged 6-9 and 30 percent aged 10-14.

in the average family is 3.7 while the average child in the survey came from a family with more than 4.5 children present, which is substantially larger than the average for the 1930s. Family income was reported only as a categorical measure of income per household member. About 60 percent of these households (and more that 70 percent of children) had family income of less than 10 shillings per capita per week. This benchmark is close to the poverty line used by Rowntree in his 1936 survey of York where 37 percent of working class households and 43 percent of children under 14 were found to be in poverty (Rowntree, 1941, pp 42, 114-9).

The original data from the Boyd Orr survey was recovered by a group of epidemiologists who conducted extensive research on the data. Their results indicate that economic and demographic factors were important influences on the height of children (Gunnell et al, 1998; Martin et al. 2002). One of these studies found that that child heights were negatively associated with the number of children in the household but there was no consistent relationship between height and birth order. They also found that per capita expenditure on food had a positive effect on height while the degree of crowding had an inconsistent but mainly negative effect. They also explored other anthropometric measures finding that socioeconomic variables mattered more for leg length than for sitting height, especially for younger children. Other measures such as foot length and shoulder width were studied by Whitley et al. (2008) revealing rather weaker effects. Interesting though these studies are, they do not adequately distinguish between direct and indirect effects on height and they do not account for the potential endogeneity of family size.

Birth order effects

One of the strengths of the Boyd Orr cohort is that it covers several children in the typical household. Thus we can observe a number of children in the same family whose birth orders cover a wide range: from first- to eleventh-born. One reason that this is important is that it is possible to base the estimate on within-family variation rather than relying on variation across different families. Accordingly, families where just one child was measured are omitted, leaving 2560 children in 843 families. Regressions with family fixed effects absorb all the effects common to the family, in particular the number of children and family income.

The dependent variable in this analysis is the z-score for height, which is the deviation of height from the median for each year of age by sex divided by the standard deviation of height by age and sex. This is calibrated from within the Boyd Orr dataset so it is not based on an external standard. The results are reported in Table 2 (see Hatton and Martin 2010a, p. 169) and the coefficients can be interpreted as units of standard deviations of height. In the first column of the table only the individual's birth order is included and this provides very little evidence of a birth order effects.

The second column includes date of birth in years, which is positive and significant. Given that the z-scores adjust for age this must be interpreted as a cohort effect, with later cohorts being taller than those born earlier. This effect is surprisingly large, amounting to an increase of 2.88 cm per decade. However it is consistent with the data assembled from school medical inspectors' reports by Harris (1994, 1995). These data show increases in height between 1920 and 1939 of 2.72 cm per decade for boys at ages between 7 and 12 and 2.68 cm per decade for girls aged 7 to 12. These trends were due, at least in part, to improvements in the external environment as captured by the infant mortality in the locality during early childhood (Hatton, 2011, p. 969). In the presence of these cohort effects, there is now a significant and negative birth order effect. This implies that moving up the birth order by one reduces height by 0.6cm. Ignoring the offsetting cohort effect, the difference in height between the first and sixth child in a family would be about 3cm--a substantial amount. These birth order effects seem to be consistent with the notion that younger children got squeezed, especially in early childhood, as the family expanded.

The third column shows that neither the child's sex nor being first-born has any appreciable effect on the height z-score. However twins are shorter by more than half a standard deviation, equivalent to about 3 cm for an eight year-old. Other hypotheses are possible, like a squeeze on those in the middle, but there is little evidence of that, or of differences in the effect of birth order between girls and boys (Hatton and Martin 2010a, p. 170).

Sibship size effects

In order to study the effect of family size we must compare one family with another. And in order to sidestep birth order effects we focus on the average height z-score of the children in each family (excluding twins). Sibship size is simply the number of children in the family at the date of observation. Although these families may be incomplete, siblings yet to be born are irrelevant because the focus is on current height of those present, not their height sometime in the future. The first two columns of Table 3 show the effect of sibship size and average date of birth (from Hatton and Martin 2010a, p. 172). The first column reveals a highly significant negative effect of sibship size on height.[7] The coefficient of -0.13 implies that one additional sibling reduces height by 0.7 cm. Date of birth gives a negative coefficient here but it is insignificant and removing it has no effect on the coefficient on sibship.

In order to address concerns about possible endogeneity, twins at last birth is used as an instrument for family size. This is likely to better predict family size than twins at lower parity, such as first birth would, but on the other hand it applies to very few families (only 21). Not surprisingly therefore, the IV coefficient in the second column is much weaker although it remains significant. The IV coefficient is larger in absolute size, implying that the 'true' effect of adding one sibling is to reduce average height by 1.2 cm. One reason for the difference in these estimates could be that families that are intrinsically unhealthy may also have fewer children, imparting a positive bias to the OLS coefficient.

If the effect of sibship size reflects resource scarcity within the family then the resource dilution effect should be better captured by income per capita. Accordingly we add a variable for income per capita in the household. This is a crude categorical variable (5 categories) taken directly from the survey. As column (3) of Table 3 shows the income variable produces a strong positive coefficient, which implies that a shilling of per capita income increases height by 0.32 cm. Nevertheless the number of children in the household remains negative and significant indicating that sibship size has direct effects on height, for example where more siblings increase the probability of growth-inhibiting infections. [8]

In the last column of Table 3 both sibship size and per capita income are instrumented. The additional instruments include predicted household income based on the head's occupation and the characteristics of the county in which the household is situated. Although the significance of both variables falls, the coefficient on income per capita is little changed and that on sibship size increases. This suggests that there is some additional negative effect on height of growing up in a large family that is not simply accounted for by spreading available income more thinly.

It is possible to explore the channels of influence a little further. Hatton and Martin (2010a, b) find that food expenditure and overcrowding (more than two persons per room) have significant effects, suggesting that the separate sibship size effect works, at least in part, through crowding

[7] It is worth noting that the R^2 in the regressions is low even though the heights are family averages. According to a Finnish study, genetic factors account for about 80 percent of the variance in height across individuals (Silventoinen et al., 2000). See also McEvoy and Visscher (2009) for a survey of genetic influences on height.

[8] This is despite the fact that, as income per capita is not adjusted by an equivalence scale, this could impart some positive bias to the coefficient on the number of children.

within the household. Consistent with this, poor quality housing characteristic of back-to-back slums is negatively associated with height. There is further corroboration that housing quality and overcrowding affected height through its impact on the disease environment. In particular poor cleanliness (assessed by interviewers) was strongly associated with overcrowding and housing quality (Hatton and Martin 2010a, p. 178). These conditions also affected the incidence of medical conditions, such as respiratory infections, which are known to inhibit growth during childhood. In the Boyd Orr cohort these infections were negatively associated with height and positively associated with sibship size (Hatton and Martin 2010b, p. 514).

4. Pre-World War I

To move further back in time we can look at servicemen who enlisted in the British army around the time of the First World War. We focus on men who were born in the 1890s and in order to capture their family circumstances during childhood we locate them in the 1901 census (see Bailey et al., 2015). Before discussing these data in more detail it is worth emphasising that here we are linking the heights of men as adults with their circumstances when observed as children, something that few historical studies have been able to do.

First World War army records

Details of the army service records and of the procedure for identifying the servicemen as children in the 1901 census are detailed in full in Bailey et al. (2015). A few points are worth noting. The characteristics recorded upon enlistment are taken from the attestation forms and the medical inspection reports. Although the information is sometimes incomplete, it typically includes next of kin, address and birthplace. Because most of the servicemen were unmarried, their next of kin is usually a parent, which helps in identifying the family in the 1901 census. As a result we are able to match (with varying degrees of certainty) 85 percent of those that we searched for. We focus on those living in England and Wales, recording details of the household, and adding characteristics of the locality. For locality we use registration districts, thus enabling a much finer classification of local conditions than would be possible at the county level.

As always with military data, sample selection issues need to be considered. About two thirds of men born in the 1890s enlisted in the armed forces during World War 1 and the imposition of conscription in 1916 ensured that they were broadly representative. Under intense pressure for recruits, the army took a substantial number that were below the original height standard of 160 cm (9.5 percent in our sample). Nevertheless it seems likely that the shortest and least fit are underrepresented as are the tallest, given that commissioned officers are not included in the records.

As Table 4 shows, the average serviceman in our matched sample enlisted fairly early in the war, at an average age of 20.5, and had an average height of 168cm (5 feet 6 inches). The data from the 1901 census reveal that these servicemen grew up in households containing an average of 6.5 persons with three other children in the family when. Their mothers had an average age of 35 when observed in 1901 and only 13 percent lived in households where the head was in a white collar occupation (social class 1 or 2). The locality in which they lived had an average population of about 150,000 with population density of 5138 per km^2 and in which about 5 percent of households were overcrowded (more than two persons per room). Other information includes whether the district was predominantly industrial or agricultural (as classified by the Registrar General), the infant mortality rate and the literacy rate of the parents' generation (as recorded in marriage registers of the early 1880s).

Household conditions and height

The effects of household level variables are shown in Table 5, where the dependent variable is the height (in cm) recorded upon enlistment. As each of these servicemen comes from a different family it is not possible to assess the within-family birth order effects as was possible with the Boyd Orr cohort (when entered in these regressions birth order was insignificant).[9] Hence we focus principally on the number of children in the family. All the regressions include dummy variables for age at enlistment. Those under the age of 20 were somewhat shorter than those aged 20 or above (the reference group), evidently because they were still growing.

The first column of Table 5 shows that, when entered without other household variables sibship size gives a significant negative coefficient. However, sibship is the number of children observed in the 1901 census, rather than completed family size, which would be a more appropriate measure for heights observed in adulthood. One way of adjusting for this is to use the mother's age to estimate the number of additional siblings and then add these to the

[9] It is important to recognise that birth order and sibship size are naturally correlated—one cannot be birth order five in a family of three. Birth order is a ranking and its 'true' effect is the hypothetical effect of moving up the birth order by one while holding family size constant. In order to identify the true birth order effect (when looking across families) one has to somehow purge the birth order variable of the sibship size effect. One way of doing this is to use instead the deviation of birth order from the sibship average birth order (which is (N+1)/2) (see Hatton, 2014, p. 162). For WW1 servicemen this index did not yield a significant coefficient. One reason may be that we have only an estimate of completed sibship size (which is relevant for final heights); alternatively there may simply be too much heterogeneity when estimating across families as compared with estimating within families.

observed sibship in order to predict completed sibship size.[10] As column (2) shows, predicted sibship size gives a larger and more significant coefficient. This implies that an additional sibling reduces height by 0.3cm, an effect that is about half the size of that estimated in the Boyd Orr cohort. This is likely to be a lower bound (as suggested by Table 3), possibly because the effect is attenuated by catch-up or retarded growth in later childhood. It might also be due to a correlation between unobservable components of family health and sibship size, but unfortunately there are insufficient twins at any given parity to use as an instrument.

Other household variables are added in the third column of Table 5. If the head of household was in a white collar occupation, and therefore middle class, height was greater by 1.2 cm. For female-headed households the coefficient is negative but it is not significant. These variables reflect social status but are likely also to capture income, which is not recorded directly in the census. Finally the number of rooms (up to five) in the dwelling has a significant positive effect on height; one additional room adds half a centimetre. This is consistent with the notion that crowding had a negative effect on height through the spread of infection.[11]

Effects of the locality

Conditions in the locality where the person was born and grew up are often found to influence height, and such effects are sometimes interpreted as reflecting unmeasured conditions in the household. But they can also be interpreted as ‘true’ neighbourhood effects rather than as crude proxies for conditions in the average household. The most powerful of these is infant mortality which is a sensitive indicator of the local disease environment affecting children. Infant mortality could have two opposing effects, scarring and selection. On one hand, as a proxy for the risk of infection, it may result in shorter stature—the scarring effect. On the other hand, higher infant mortality might leave healthier and taller survivors—the selection effect. In most settings the scarring effect dominates, see Bozzoli et al. (2009); Hatton (2011), but in extreme cases, such as the Great Chinese Famine, selection effects may be large enough to offset scarring effects (Gørgens et al. 2012).

[10] The prediction is based on a regression coefficient of the number of children on mother’s age for mothers aged 20-40 of 0.191 (‘t’ = 20.4). To calculate predicted sibship, e.g. for a mother aged 30 we add (40-30)*0.191 = 1.91 additional children to the observed number. For families where the mother is over 40 or there is no mother present, no further children are added. This gives an average predicted sibship of 5.19 as compared with the unadjusted average of 4.16.

[11] If the number of rooms is replaced by a dummy variable for the number of persons greater than the number of rooms, this takes a negative coefficient with a ‘t’ value of 2.33. However, if both variables are included the latter becomes insignificant.

Table 6 adds to the regression the average infant mortality rate (as a percentage of births) over the decade of the 1890s in the local registration district. Here, the age dummies and the constant term are included but not reported. In column (1) infant mortality gives a highly significant coefficient, which is consistent with the notion that infections inhibit growth. It implies that children growing up in a locality where infant mortality was 20 percent would be around 0.4 cm shorter that those growing up where the infant mortality rate was 10 percent.

Infant mortality provides a summary measure of the disease environment but it clearly has deeper causes. Important among these are sanitary conditions, the fabric of houses and streets and industrial pollution--influences that are often hard to measure directly.[12] Even more obscure are the customs and practices of households and their knowledge of the effects of hygiene and nutrition. To capture some of these effects, column (2) adds two variables, the rate of adult female illiteracy in the district (based on ability to sign the marriage register) and a dummy variable for districts dominated by heavy industry. Both of these variables are significant and in the expected direction. In their presence the coefficient on the infant mortality rate is almost halved and its significance is greatly reduced.[13]

While female illiteracy and the industrial character of the district have direct effects on height they are also important determinants of infant mortality (Bailey et al. 2015, p. 17). The third column of Table 6 shows that when the latter is omitted the coefficients female illiteracy and the industry dummy become larger and more significant. A difference between districts of ten percentage points in female illiteracy is associated with a 0.8cm reduction in height and growing up in an industrial district is associated with a substantial decrease of 1.5cm. It is worth also noting that, even though locality variables have substantial effects, including them has very little effect on the coefficient on sibship size.

5. Trends over time and across countries

Trends in Britain

Fertility fell dramatically from the late nineteenth century through to the 1930s and it is worth asking how much the fall in family size could have contributed to the increase in height. One

[12] Studies that directly link reductions in infant mortality to sanitary reforms include Cutler and Miller (2005) for the US, Macassa et al. (2006) for Sweden, and Newell and Gazeley (2012) for the UK.

[13] Several studies point to the importance of the supply of protein, especially milk, as proxied by the density of cattle in the locality (see Baten, 2009). Adding to the column (3) regression the share employed as cattlemen as a proxy for access to protein gave a positive but insignificant coefficient with very little effect on the coefficient on sibship size. This probably reflects the fact that by 1901 the railways had provided good access to farm products for urban populations, and milk was far less adulterated than thirty years earlier.

approach is to use the cross sectional estimates to assess the contribution of falling family size to the increase in height. From surveys of working class budgets it is estimated that average sibship size fell from 4.5 in 1886 to 2.0 in 1938, while average family income (in constant 1938 prices) increased from 55.7 shillings to 69.6 shillings. Rising income and falling family size would both have contributed to increasing height.

One calculation is based on an estimated coefficients from the Boyd Orr cohort. This suggests that in the half century after 1886 the joint effect of rising per capita income and falling family size was to increase the height of eight year-olds by about 3.8 cm. More than 2cm of this is accounted for by declining family size, partly due to reducing the denominator of family income per capita and partly due to the direct effect of sibship size (Hatton and Martin, 2010b, p. 516). Falling sibship size added about 0.3cm per decade between 1886 and 1906, increasing to about 0.5cm per decade in the ensuing thirty years.

For adult heights we can use the result for World War 1 servicemen. However, the estimated effects differ for at least two reasons. First, time series evidence indicated that the heights of children increased more rapidly in the first half of the twentieth century than the heights of the same cohorts as adults.[14] Part of this difference is due to the fact that children were reaching maturity earlier, at which time they stopped growing.[15] In the late nineteenth century they were still growing after age 18 (as illustrated in Table 5); by the middle of the twentieth century they were not. Thus we might expect the effect of sibship size on adult height to be less than that for children. The second is that the estimated coefficient underestimates the sibship effect because it fails to account for endogeneity. Adjusting the latter upwards by the ratio of IV to OLS coefficients in the Boyd Orr cohort gives: $-0.3 \times (1.2/0.7) \approx -0.5$. Over the century from the 1870s British adult male heights increased by about 10cm (see below). Falling sibship size would account for $(2.0 - 4.5) \times -0.5 = 1.25$cm or about one eighth of the total. But if we take just the period from the late 1880s to the late 1930s, when heights increased by about 5cm, then it would account for about a quarter of the increase.

Trends in European heights and fertility decline

[14] In Britain, from 1910 to 1950 the heights of schoolboys at ages 6, 8, 10 and 12 increased by between 1.7 and 2.3 cm per decade (Hatton, 2011, p. 963). By contrast the heights of adult males increase by about 1 cm per decade, see Table 7 below.

[15] The clearest evidence is for the age at menarche in girls. In Europe this declined from an average age of 14 in 1900 to 12.8 in 1947 (Wyshak and Frisch, 1982); there is also evidence that earlier menarche is associated with smaller sibship size (Morris et al., 2010).

Trends in the height of adult males by birth cohort have been assembled for 15 European countries. For the postwar period these are based mainly on height-by-age from cross-sectional surveys and these are carried back to the birth cohorts of the 1870s using data for the heights of army recruits (for details see Hatton and Bray, 2010). As Table 7 shows, for birth cohorts over the century from the 1870s to the 1970s, height increased at about a centimetre per decade. There is evidence of some acceleration in the first half of the twentieth century among most of the countries of north and middle Europe with some subsequent slowing down, while in southern Europe the spurt occurred after the Second World War. It seems likely that the fertility decline contributed something to these trends, although there are many other influences.

The database on five year averages of heights in the 15 countries can be used to explore the correlates of height at the macro level by regressing adult height on conditions around the time of each cohort's birth. The explanatory variables are inevitably crude and the details of their construction are provided in Hatton (2014). In particular there are no consistent series on family size and instead this is represented by the ratio of children aged 0-14 to married females aged 20-44 (similar results are obtained with ages 20-54 in the denominator). Because of the strong trends in many of the variables, I use deviations of each variable from its linear trend, obtained from regressions with fixed country effects. The regressions in Table 8 include dummies for country and for breaks in the series. While we must be cautious about inferring causal effects, the 20-year time lag between the explanatory variables and observed heights, the elimination of trends, and the use of country fixed effects provides some reassurance.

The first column of Table 8 shows that the log of GDP per capita has a strong positive effect on height as might be expected. It implies that a ten percent increase in GDP per capita around the time of birth increases height by about 0.18 cm. The proxy for sibship size is significantly negative. However, a decline of one child per married women of childbearing age adds only 0.5cm to average height. Column (2) adds average years of education for the parents' generation. This variable is calculated from the number of children of the previous generation attending school. It takes a positive coefficient as predicted, and it implies that one additional year of parental education increases height by 0.2cm.

In column (3) the percentage infant mortality rate is introduced in quadratic form in order to capture the possible non-linear effect on height. As noted above, Bozzoli et al. (2009) argue that, at high levels of infant mortality, the selection effect is larger relative to the scarring effect and so the overall negative effect will be smaller and could even become positive.[16] The result in column (3), where the linear term is negative and the squared term is positive, is consistent with that argument, although there could be other interpretations. Not surprisingly when infant mortality is added to the regression the coefficient on GDP per capita declines but there is also some reduction in the effect of family size.

How much did these variables account for the long run increase in height? Their contributions can be calculated using the coefficients in column (3) of Table 8. In northern and middle Europe the growth of GDP per capita contributed 1.6 to 1.8 centimetres to height over the whole period under review, while the dramatic fall in infant mortality added 4.4 to 4.8 cm (Hatton, 2014, p. 362). The contributions of the other variables are modest; the rise in education accounted for 0.8 to 1.0 cm while the fall in family size accounted for 0.5 to 0.6 cm. However, this is undoubtedly a downward-biased estimate as we have only a poor proxy for average sibship size. Comparison with the results from micro-data from Britain would suggest doubling this effect.

6. Conclusion

There is a growing literature that provides evidence on the existence of a trade-off between the number of children in a family and various dimensions of child quality. Although most of the focus has been on education as the outcome, health as measured by height is also important. A review of the existing literature suggests that such effects might be stronger in settings where families are poorer and larger and when the focus is on height rather than on education. They may also be stronger where public provision is limited, where the state of knowledge is poor, and where the external disease environment is harsh.

Although heights have been studied in a wide range of contexts, few such studies have directly assessed their links with sibship size at the family level. In this paper I have reviewed the results from two micro-level datasets for Britain both of which support the idea of a quality-quantity trade-off. In the Boyd Orr cohort the results suggest that sibship size had a substantial negative effect and not only through the dilution of income. For World War I soldiers measured as adults, the sibship size effect is smaller but the negative effect of the surrounding disease

[16] The turning point implied by the estimated quadratic is an infant mortality rate of 21.6 percent, which is within the range observed in the data.

environment is even clearer. These effects can also be discerned in country-level time series although the variable used for sibship size is less than satisfactory.

These results imply that the fertility decline in European countries that began in the late nineteenth century and progressed through the first half of the twentieth century improved the health of successive cohorts as children and as adults. It made a modest but nevertheless significant contribution to the overall health gains, at least as measured by height. For Britain, falling family size accounted for up to a quarter of the increase in adult male heights for cohorts born between the 1880s and the 1930s. But improvements in the sanitary environment, in housing conditions, in public health systems, as well as in education and basic knowledge of nutrition and hygiene were more important.

References

Alderman, H. (1990). Nutritional Status in Ghana and its Determinants. World Bank SDA Working Paper No. 3.

Angrist, J., Lavy, V. and Schlosser, A. (2010). Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, 28, 773-824.

Bailey, R. E., Hatton, T. J. and Inwood, K. (2015). Health, Height and the Household at the Turn of the 20th Century. *Economic History Review* (forthcoming).

Baten, J. (2009), "Protein Supply and Nutritional Status in Nineteenth Century Bavaria, Prussia and France," *Economics and Human Biology*, 7, pp. 165-180.

Baten, J. and Blum, M. (2012). Growing Tall but Unequal: New Findings and New Background Evidence on Anthropometric Welfare in 156 Countries, 1810–1989. *Economic History of Developing Regions*, 27 (Sup1), S66-S85.

Baten, J. and Böhm, A. (2010), "Children's Height and Parental Unemployment: A Large-Scale Anthropometric Study on Eastern Germany, 1994–2006," *German Economic Review*, 11, pp. 1–24.

Becker, G. S. (1960). An Economic Analysis of Fertility. In G. S. Becker (ed.), *Demographic and Economic Change in Developed Countries*. Princeton NJ: Princeton University Press.

Becker, G. S. and Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, 81, S279-S288.

Becker, G. S. and Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, 84, S143-S162.

Becker, S., Cinnirella, F. and Woessmann, L. (2010). The Trade-off between Fertility and Education: Evidence from before the Demographic Transition. *Journal of Economic Growth*, 15, 177-204.

Black, S.E., Devereaux, P.J., and Salvanes, K.G. (2005). The More the Merrier? The Effect Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, 120, 669–700.

Bleakley, H., and Lange, F. (2009). Chronic Disease Burden and the Interaction of Education, Fertility, and Growth. *Review of Economics and Statistics*, 91, 52–65.

Bozzoli, C., Deaton, A. S. and Quintana-Domeque, C. (2009). Adult Height and Childhood Disease. *Demography*, 76, 647-669.

Cáceres-Delpiano, J. (2006). The Impacts of Family Size on Investment in Child Quality. *Journal of Human Resources*, 41, 738-754.

Clark, G. (2005), "Human Capital, Fertility, and the Industrial Revolution," *Journal of the European Economic Association*, 3, pp. 505-515.

Cole, T. J. (2003). The Secular Trend in Human Physical Growth: A Biological View. *Economics and Human Biology*, 1, 161-168.

Conley, D. and Glauber, R. (2006). Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility. *Journal of Human Resources*, 41, 722-737.

Cutler, D. and Miller, G. (2005). The Role of Public Health Improvements in Health Advances: The Twentieth Century United States. *Demography,* 42, 1-22.

Dang, A-H. Rogers, H. (2013). The Decision to Invest in Child Quality over Quantity: Household Size and Household Investment in Education in Vietnam. World Bank Policy Research Working Paper 6487.

Desai, S. (1992). Children at Risk: The Role of Family Structure in Latin America and West Africa. *Population and Development Review,* 18, 689-717.

Dribe, M., Van Bavel, J. and Campbell, C. (2012). Social Mobility and Demographic Behavior: Long term Perspectives. *Demographic Research*, 26, 173-190.

Ermisch, J. F. (2003). *An Economic Analysis of the Family.* Princeton NJ: Princeton University Press.

Fernihough, A. (2011). Human Capital and the Quantity-Quality Trade-Off during the Demographic Transition: New Evidence from Ireland. University College Dublin Working Paper 11/13.

Floud, R., Fogel, R.W., Harris, B., and Hong, S.C. (2011). *The Changing Body: Health, Nutrition, and Human Development in the Western World since 1700*. Cambridge: Cambridge University Press.

Fogelman, K. R. (1975). Developmental Correlates of Family Size. *British Journal of Social Work,* 5, 43-57.

Galor, O. (2011). *Unified Growth Theory*. Princeton: Princeton University Press.

Glick, P. J., Marini, A. and Sahn, E. (2007). Estimating the Consequences of Unintended Fertility for Child Health and Education in Romania: An Analysis using Twins Data. *Oxford Bulletin of Economics and Statistics,* 69, 667-691.

Goldstein, H. (1971). Factors Influencing the Height of Seven Year Old Children: Results from the National Child Development Study. *Human Biology,* 43, 92-111.

Gørgens, T., Meng, X. and Vaithianathan, R. (2012). Stunting and Selection Effects of Famine: A Case Study of the Great Chinese Famine. *Journal of Development Economics*, 97, 99-111

Gunnell, D. J., Davey Smith, G., Frenkel, S. J., Kemp, M. and Peters, T. J. (1998). Socio-economic and Dietary Influences on Leg Length and Trunk Length in Childhood: A Reanalysis of the Carnegie (Boyd Orr) Survey of Diet and Health in Prewar Britain. *Paediatric and Perinatal Epidemiology*, 12, 96-113.

Hanusheck, E. (1992). "The Trade-off between Child Quality and Quantity. *Journal of Political Economy*, 100, 894–917.

Harris, B. (1994). The Height of Schoolchildren in Britain, 1900-1950. In J. Komlos (ed.) *Stature, Living Standards and Economic Development* (pp. 25-38). Chicago: University of Chicago Press.

Harris, B. (1995). *The Health of the Schoolchild: A History of the School Medical Service in England and Wales.* Buckingham: Open University Press.

Hatton, T. J. (2011). Infant Mortality and the Health of Survivors: Britain, 1910–50. *Economic History Review*, 64, 951–972.

Hatton, T. J. (2014). How have Europeans Grown so Tall? *Oxford Economic Papers*, 66, 349–372.

Hatton, T. J. and Bray, B. E. (2010). Long-run Trends in the Heights of European Men, 19th–20th centuries. *Economics and Human Biology*, 8, 405–413.

Hatton, T. J. and Martin, R. M. (2010a). The Effects on Stature of Poverty, Family Size and Birth Order: British Children in the 1930s. *Oxford Economic Papers*, 62, 157-184.

Hatton, T. J. and Martin, R. M. (2010b). Fertility Decline and the Heights of Children in Britain, 1886-1938. *Explorations in Economic History*, 47, 505-519.

Horton, S. (1986). Child Nutrition and Family Size in the Philippines. *Journal of Development Economics*, 23, 161-176.

Horton, S. (1988). Birth Order and Child Nutritional Status: Evidence from the Philippines," *Economic Development and Cultural Change*, 3, 341-354.

Klemp, M. P. B. and Weisdorf, J. (2012). Fecundity, Fertility and Family Reconstitution Data: The Child Quantity-Quality Trade-Off Revisited. London: CEPR Discussion Paper 9121.

Kuh, D. and Wadsworth, M., (1989). Parental Height, Childhood Environment, and Subsequent Adult Height in a National Birth Cohort. *International Journal of Epidemiology*, 18, 663-668.

Lee, J. (2007). Sibling Size and Investment in Children's Education: An Asian Instrument. *Journal of Population Economics*, 21, 855–875.

Li, H., Zhang, J and Zhu, Y. (2008). The Quantity-Quality Trade-off of Children in a Developing Country: Identification using Chinese Twins. *Demography*, 45, 223–243.

Li, L. and Power, C. (2004). Influences on Childhood Height: Comparing Two Generations in the 1958 British Birth Cohort. *International Journal of Epidemiology,* 33, 1320-1328.

Liu, H. (2014). The Quality–quantity Trade-off: Evidence from the Relaxation of China's One-child Policy. *Journal of Population Economics,* 27, 565–602.

Macassa, G., Ponce de Leon, A. and Burström, B. (2006). The Impact of Water Supply and Sanitation on area Differentials in the Decline of Diarrhoeal Disease Mortality among Infants in Stockholm 1878—1925. *Scandinavian Journal of Public Health,* 34, 526-533.

Martin, R. M., Smith, G. D., Mangtani, P., Frankel, S. and Gunnell, D. (2002). Association between Breastfeeding and Growth: The Boyd Orr Cohort Study. *Archives of Disease in Childhood: Fetal and Neonatal Edition,* 87, F193-F201.

McEvoy, B. P. and Visscher, P.M. (2009). Genetics of Human Height. *Economics and Human Biology,* 7, 294–306.

Morris, D. H., Jones, M. E., Schoemaker, M. J., Ashworth, A. and Swerdlow, A. J. (2010). Determinants of Age at Menarche in the UK: Analyses from the Breakthrough Generations Study. *British Journal of Cancer,* 103, 1760-1764.

Newell, A. and Gazeley, I. (2012). The Declines in Infant Mortality and Fertility: Evidence from British Cities in Demographic Transition. Bonn: IZA Discussion Paper No. 6855.

Öberg, S. (2015). Sibship Size and Height before, during, and after the Fertility Decline: A Test of the Resource Dilution Hypothesis. *Demographic Research*, 32, 29-74.

Parman, J. (2015). Childhood Health and Human Capital: New Evidence from Genetic Brothers in Arms. *Journal of Economic History*, 75, 30-64.

Ponczek, V. and Portela Souza, A. (2012). New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country. *Journal of Human Resources*, 47, 64-105.

Qian N. (2009). Quantity-quality and the One Child Policy: The Only-child Disadvantage in School Enrolment in Rural China. Cambridge MA, NBER Working Paper 14973.

Rosenzweig, M.R. and Wolpin, K.I. (1980). Testing the Quantity-Quality Fertility Model: the use of Twins as a Natural Experiment. *Econometrica*, 48, 222–40.

Rosenzweig, M. R. and Zhang, J. (2009). Do Population Control Policies Induce more Human Capital Investment? Twins, Birth Weight, and China's "One Child" Policy. *Review of Economic Studies,* 76, 1149-1174.

Rowett Institute (1955). *Family Diet and Health in Pre-War Britain: A Dietary and Clinical Survey.* Dunfermline: Carnegie U. K. Trust.

Rowntree, B. S. (1941). *Poverty and Progress: A Second Social Survey of York.* London: Longmans Green.

Schneider, R. (1996). Historical Note on Height and Parental Consumption Decisions. *Economics Letters,* 50, 279-283.

Silventoinen, K., Kaprio, J., Lahelma, E. and Koskenvuo, M. (2000). Relative Effect of Genetic and Environmental Factors on Body Height: Differences across Cohorts among Finnish Men and Women. *American Journal of Public Health*, 90, 627-630.

Steckel, R. H. (1995). Stature and the Standard of Living. *Journal of Economic Literature,* 33, 1903-1940.

Steckel, R. H. (2009). Heights and Human Welfare: Recent Developments and New Directions. *Explorations in Economic History,* 46, 1-23.

Weil, D. N. (2007). Accounting for the Effect of Health on Economic Growth. *Quarterly Journal of Economics*, 122, 1265-1306.

Weir, D. R. (1993). Parental Consumption Decisions and Child Health during the Early French Fertility Decline, 1790-1914. *Journal of Economic History,* 53, 259-274.

Whitley, E., Gunnell, D., Davey Smith, G., Holly, J. M. P. and Martin, R. M. (2008). Childhood Circumstances and Anthropometry: The Boyd Orr Cohort. *Annals of Human Biology,* 35, 518-534.

Wyshak, G. and Frisch, R. E. (1982). Evidence for a Secular Trend in Age of Menarche," *New England Journal of Medicine*, 306, 1033-1035.

Table 1: Characteristics of Children in the Boyd Orr Survey

	By individual	By household
No of cases	2946	1131
Individual characteristics		
Average age	7.92	7.85
Average birth order	2.78	2.35
Percent female	52.7	53.9
Household characteristics		
Number of children in family	4.56	3.74
Total number in family (adults + children)	6.75	5.96
Percent with family income per capita < 10s	71.56	59.8

Source: Hatton and Martin, 2010a, p. 167.
Note: The average across households is calculated as the average of the mean value for each household. Due to missing data, income per capita is the average over 2911 children in 1112 households.

Table 2: Effect of Birth Order on Height z-scores in the Boyd Orr Cohort

	(1)	(2)	(3)
Constant	-0.042	-1.042	-1.203
	(0.34)	(3.35)	(3.69)
Birth order	-0.004	-0.100	-0.105
	(0.31)	(3.12)	(3.24)
Date of birth (years)		0.042	0.048
		(3.24)	(3.69)
First born			0.033
			(0.73)
Sex (F = 1)			0.003
			(0.09)
Twin			-0.535
			(4.07)
R^2 (within)	0.00	0.006	0.015
No. children	2560	2650	2650
No. families	835	835	835

Source: Hatton and Martin (2010a) p. 169.
Note: Regressions with family fixed effects; *t*-statistics in parentheses.

Table 3: Effect of Income and Family Size on Height z-scores in the Boyd Orr Cohort

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Constant	0.988	1.403	-0.455	-0.160
	(3.32)	(2.48)	(1.35)	(0.24)
No of children	-0.134	-0.233	-0.071	-0.162
	(9.61)	(2.06)	(4.78)	(2.35)
Income per capita			0.053	0.057
			(8.41)	(3.54)
Date of birth (years)	-0.011	-0.013	0.009	0.009
	(1.21)	(1.31)	(0.96)	(0.82)
R^2	0.083	0.059	0.145	0.123
No. Families	1102	1102	1102	1102

Source: Hatton and Martin 2010a, pp. 171, 175

Note: *t*-statistics in parentheses computed from robust standard errors. For the IV regression in column (2) the p-value for the endogeneity test is 0.66. For the IV regression in column (4) the p-value for the endogeneity test is 0.08 and that for the overidentification test is 0.09.

Table 4: Descriptive statistics of WW1 servicemen sample

	Mean	**Std. Dev.**
Individual serviceman characteristics		
Height (cm)	167.78	6.53
Attestation date (year)	1915.5	1.36
Age at attestation (years)	20.53	1.96
Household characteristics in 1901		
Persons in household	6.54	2.14
Sibship size	4.16	2.08
Mother's age (if mother present)	35.6	6.70
Female household head (%)	6.03	23.8
While collar household head (%)	13.2	33.8
Households with 4 rooms or less (%)	51.7	50.0
Locality characteristics in 1901		
Population (000s) 1901	148.6	126.9
Population density (000s per km^2) 1901	5.1	9.0
More than 2 per room (%) 1901	5.6	6.3
Infant mortality rate (%) 1891-1900	15.2	2.9
'Industrial' district (%)	27.0	44.4
Female illiteracy (%) 1881-4	15.3	9.2

Source: Bailey et al. (2015), p. 7.

Table 5: Effects of childhood household variables on height of WW1 servicemen

	(1)	(2)	(3)
Constant	169.260	170.078	167.984
	(502.06)	(411.33)	(242.26)
Age <18	-2.919	-2.755	-2.769
	(4.83)	(4.60)	(4.62)
Age 18	-1.988	-1.905	-1.866
	(4.95)	(4.79)	(4.74)
Age 19	-1.178	-1.092	-1.071
	(3.77)	(3.48)	(3.46)
Observed sibship size	-0.169		
	(2.61)		
Predicted sibship size		-0.302	-0.308
		(4.15)	(4.23)
Female head of household			-1.066
			(1.73)
White collar head of household			1.237
			(3.10)
No of rooms			0.486
			(3.53)
R^2	0.020	0.025	0.039
No. of individuals	2236	2236	2236

Notes: *t*-statistics in parentheses from robust standard errors clustered at the registration district level. These specifications are variations on those reported in Tables 3 and 4 of Bailey et al. (2015).

Table 6: Effects of local infant mortality on height of WW1 servicemen

	(1)	(2)	(3)
Predicted sibship	-0.272	-0.274	-0.285
	(3.83)	(3.89)	(4.01)
Female head of household	-0.969	-0.922	-0.934
	(1.62)	(1.54)	(1.56)
White collar head of household	1.332	1.314	1.277
	(3.43)	(3.40)	(3.28)
No of rooms	0.329	0.384	0.455
	(2.38)	(2.84)	(3.42)
District Infant mortality (%)	-0.388	-0.200	
	(7.81)	(2.95)	
District female illiteracy rate		-0.056	-0.079
		(2.71)	(4.32)
Industrial district		-1.109	-1.569
		(2.51)	(3.95)
R^2	0.068	0.078	0.074
No. of individuals	2236	2236	2236

Notes: *t*-statistics in parentheses from robust standard errors clustered at the registration district level. These specifications are variations on those reported in Table 6 of Bailey et al. (2015). Constant terms and dummy variables for age under 18, 18 and 19 are included but not reported.

Table 7: Increase in heights of male birth cohorts in centimetres per decade

	1871-75 to 1976-80	1871-75 to 1911-15	1911-15 to 1951-55	1951-55 to 1976-80
Austria	1.11	0.59	1.50	1.32
Belgium	1.08	0.41	1.59	1.32
Denmark	1.24	0.58	1.83	1.37
Finland				0.84
France	0.91	0.57	1.10	1.16
Germany	1.25			1.20
Great Britain	0.93	1.14	0.99	0.50
Greece				1.55
Ireland	0.80			1.00
Italy	1.06	0.72	1.14	1.50
Netherlands	1.41	1.34	1.32	1.67
Norway	0.93	0.79	1.49	0.26
Portugal			0.94	1.72
Spain	1.19	0.74	0.79	2.53
Sweden	0.97	0.68	1.25	1.00
Average	1.08	0.76	1.27	1.26
Standard Deviation	0.18	0.28	0.31	0.54

Source: Hatton and Bray, 2010, p. 407.

Table 8: Proximate determinants of adult male height in 15 European countries

	(1)	(2)	(3)
Log GDP per capita	1.811	1.704	0.887
	(5.77)	(5.49)	(2.43)
Family Size	-0.492	-0.537	-0.365
	(2.96)	(3.50)	(2.39)
Years of education		0.214	0.205
		(2.34)	(2.50)
Infant mortality rate (per 100 births)			-0.562
			(7.78)
Infant mortality rate squared			0.013
			(5.66)
R-squared	0.175	0.194	0.383
Countries	15	15	15
Observations	267	267	267

Note: Estimated with country fixed effects; robust *t*-statistics in parentheses. This is a modified version of Table 3 in Hatton (2014), where the insignificant inequality variable has been dropped.

Figure 1: The Child Quality-Quantity Trade-off