Population and human development since 1700

Romola Davenport and Osamu Saito

Introduction

This chapter explores complex relationships between population and economy since c.1700. The West in the 1700-1870 period saw mortality declining in most countries, setting off a fertility decline after 1870. The demographic transition is comprised of these two sequential changes. While the chronology is different for non-Western countries, the transition from a situation of high mortality and high fertility to that of low mortality and low fertility is universal. As Massimo Livi-Bacci puts it, it was a series of passages from disorder to order and from waste to efficiency (2012, ch. 4). It is one of the fundamental transformations humans experienced and is believed to have been a change accompanied by another fundamental change, modern economic growth (Thompson 1929; Notestein 1945). Across the world regions, however, it seems that there is no straightforward correlation between population change and economic growth.

Falls in mortality meant that a number of European populations in long-settled areas achieved sustained growth rates approaching 1% per annum or more across the nineteenth century (Chesnais, 1992). These rates were historically unprecedented, but were modest compared with rates of 1.5-5 % per annum amongst developing countries in the mid-twentieth century (Wilson and Airey, 1999). In Britain most of the rapid population growth of the late eighteenth and nineteenth centuries was absorbed into cities, and the urban population (proportion living in settlements of population 2,500 or more) increased from 18% to 30% between 1700 and 1800, and to 65 % by 1870 (de Vries, 1984; Law, 1967). Urbanisation was probably also rapid in Belgium; however, elsewhere
urban growth was slow before the mid-nineteenth century, and there was substantial growth in rural areas (Bairoch & Goertz, 1986, p. 288; Floud et al., 2011, p. 277; Wrigley, 2014a). Thus, while falls in especially urban mortality were required for modern levels of urbanisation, they did not drive urbanisation (contra Dyson, 2011).

Population growth slowed in western Europe and the Anglophone neo-Europes (or western offshoots) at the end of the nineteenth century. Between 1870 and 1930 all countries in these areas experienced significant and historically unprecedented fertility declines, such that crude birth rates fell from a range of c.25-45 births per 1,000 population to less than 20 in all cases, and total fertility rates (TFRs) from around 5 births per woman to less than 3 by the 1930s (Chesnais, 1992, p. 117-28). These declines were achieved largely by control of fertility within marriage (rather than by falls in marriage rates), and by traditional means of contraception, especially withdrawal and abstinence. The causes of this relatively synchronised and profound transition remain hotly debated. Much of the debate is focussed on explaining the exact timing of fertility declines (Mason, 1997). Common to all populations, however, was the progressive impact of rising life expectancy and associated population growth.

1. Pre-transition demography and the demographic transition

According to a stylised depiction of the term, ‘In traditional societies, fertility and mortality are high. In modern societies, fertility and mortality are low. In between, there is the demographic transition’ (Demeny 1972, p. 153). The traditional societies were ‘inefficient’ because given the high levels of infant and child mortality and frequent catastrophes a good deal of births was need to maintain a low level of population growth. The traditional economy was governed by a Malthusian mechanism, expressed in observed correlations between mortality and indices of agricultural performance such as the frequency of famines. Thus, according to the classical theory of demographic transition, the demographic transition – starting with the attenuation of mortality crises, followed by a decline in infant and child mortality, and finishing with fertility decline – was causally related to the transition to modern economic growth, i.e. to the emergence of an industrial and urban economy, in which population increase was no longer incompatible with output and income growth. However, this interpretation of the link between economy and demography has been called into question in recent historiography (see Davenport 2015; and Saito 2014a, 2016).

The first phase of mortality decline took the form of the attenuation of mortality fluctuations. For this phase there is evidence on seventeenth- and eighteenth-
century England and Sweden. In England, for example, nationwide famines disappeared very early: after 1600, no years experienced an annual crude death rate exceeding 50 per cent above trend (Wrigley and Schofield 1989, table 8.11). On the agricultural side there is consensus that much of agrarian progress in England was made before 1750: it is demonstrated that by that date English labour productivity reached a level higher than any other European counties for which estimates are available (Allen 2000), while an alternative estimate suggests that the country may have achieved that position by 1700 (Burnette 2014). According to even more recent estimates, the period between 1550 and 1700 saw a doubling of agricultural production; since the country’s population in 1700 was two-third higher than in 1550, it implies that agricultural output per total population reached a level one-third higher than at the beginning of the 150-year period (Broadberry et al. 2015, table 3.21). The Swedish crude death rate exhibits a similar narrowing of its oscillation with the last high peak in the first decade in the nineteenth century (Livi-Bacci 2012, figure 4.3 and table 4.2).

It is notable that the timing of a weakening of the mortality response to short-term wage fluctuations coincided with an agricultural revolution in the 1800s, although the landless remained vulnerable to such short-term fluctuations (Bengtsson and Ohlsson 1985; Bengtsson 2004). At the other end of Eurasia, Japan’s last nationwide famine took place as late as the 1830s. But the frequency of famine occasioned by drought had already begun to decline from the early eighteenth century onwards, thus making rice harvest much less unpredictable. During the 125-year period from 1721, while population remained stationary, primary-sector output is estimated to have increased by one-third (Saito and Takashima 2016, table 2). Clearly, therefore, much of the attenuation of mortality volatility can be accounted for by the stabilisation of harvests. This is a result not inconsistent with the Malthusian story.

With more empirical evidence accumulated, however, both the stage-theoretic conceptualisation and the Malthusian interpretation of the transition process have been called into doubt. One area in which the supposed scenario does not work is fertility. Agricultural progress may well have exerted a positive effect on the level of fertility. In England, there is evidence that marital fertility increased in the eighteenth century (Wrigley et al. 1997, ch. 7). Given another finding that its co-variant was a decline in stillbirth rates (Woods, 2009; Wrigley 2004), both changes are likely to have reflected an improvement in living standards, especially in the availability of foodstuffs. Fertility also rose as a consequence of higher marriage rates, in Britain, and it is likely that fertility generally tended to rise rather than decline before the transition to modern fertility. And this pre-transition rise in fertility took place, not just in eighteenth-century England, but also across the world regions – from nineteenth-century Asia, including Japan, to
twentieth-century African countries (Dyson and Murphy 1985). It was particularly the case when there was an ‘excess demand’ for children, and it is likely that ‘fertility conditions in premodern societies more nearly approximate an excess demand situation than one of excess supply’ (Easterlin 1978, p. 131).

In the past, it was often assumed, fertility levels were at a biological maximum. However, this assumption is no longer tenable. While most of pre-modern societies were unplanned, ‘natural’ fertility populations, the actual levels of observed marital fertility varied considerably across world regions. For example, the average of pre-transitional total marital fertility rates (TMFRs) for three East Asian countries of Japan, Korea and rural China was 5.6, with the range of 5.8-5.3, while that for five European counterparts of England, France, Germany, Spain and Sweden stood at 8.1 with the range of 8.7-7.4 (Saito 2014a, figure 1; note that the lowest in Europe, i.e. England, was higher than the highest in East Asia, i.e. Japan). There is no guarantee that marital fertility rates were correlated with the total fertility rates (TFRs), for the age at marriage and, hence, the age-specific proportions married may vary according to nuptiality regimes. As argued by Hajnal (1965, 1983) and demonstrated for England (Wrigley and Schofield 1989, ch.10), it is expected that in north-west European countries where the nuclear family system prevailed the gap between TFR and TMFR is likely to have been wide whereas in societies where more complex family systems were dominant both tended to come closer. But among west-European countries, there existed room for substantial variation in the level of ‘natural’ fertility on the eve of fertility transition.

It has been argued that patterns of marriage and household formation promote or hinder economic development in three main ways – via effects on fertility, on female labour force participation, and on the status of women. Hajnal argued that the north-west European system was characterised by late and non-universal marriage for both sexes, neo-localism (establishment of a new household upon marriage), and a particular kind of labour market institution – the circulation of young adults in the labour market (especially as female live-in servants) before marriage; all this was in sharp contrast to Eastern Europe, where family form was joint and marriage young and universal (Hajnal 1983). Indeed, where marriage age was late, the spousal gap in marriage age small, and marriages were arranged by the couples themselves, then the potential for women and couples to

---

1 See Tomobe (2001) for Tokugawa Japan. Drixler (2013) puts forward a strong argument that low fertility levels in the period after c.1700 were accounted for by widespread practices of infanticide and abortion. However, his claim would be justifiable only when seventeenth-century births and deaths are assumed to have been accurately recorded. For this see Saito (2014b).
negotiate reproductive decisions was necessarily higher than in situations where familial and social pressures predominated. Since we know that north-west European populations did well in long-term economic development, it is often argued that the system and its correlate, a culture of personal autonomy, were conducive to modern fertility decline, on the one hand, and tended to promote female labour force participation and mobility, and hence, economic growth, on the other (Hajnal, 1965, p. 132; Wrigley, 2014b; de Moor & van Zanden, 2014; Carmichael et al., 2016).

However, the association between marriage systems and economic performance is not straightforward. First, fertility of north-west European countries was not unusually low, despite late and non-universal marriage. Rather, evidence from a range of historical societies indicates that most populations exercised norms and behaviours that served to regulate family size and realised fertility (including abortion, infanticide, neglect, prolonged breastfeeding, and sexual abstinence), suggesting that the supposed European continuum from the east to the north-west cannot be used as a measure for areas outside Europe and European offshoots. Second, the level of female labour force participation appears to have varied widely, depending on the extent to which women could combine childbearing with other forms of non-domestic work, which was constrained by the type of work available for females outside the home. In nineteenth century England, for example, employment of married women was concentrated in a limited number of districts where textiles, straw-plaiting and lace-making offered jobs for female labour (Shaw-Taylor & You, unpubl.). In the non-European world, by contrast, most women in family farms tended to work along with men; but while many African and Asian married women did provide the bulk of agricultural labour (e.g. Boserup, 1970; Goody, 1976), there are cultures in which women's participation in productive work itself was discouraged, be it domestic or paid work. In the Indian sub-continent, the south represents the former and the north the latter type (Dyson and Moore 1983). A similar contrast may be found between Japan and Korea in the East Asian region; but it should be noted that the family system in north India is what Hajnal called ‘joint’ type whereas Korea’s is said to have moved closer to the stem family type (Saito 2014a).

Finally, a few words on the culture of personal autonomy. The importance of female autonomy is clear with respect to demographic outcomes. Female autonomy, variously defined, is widely regarded as one of the most powerful predictors of child survival and small family size in contemporary poor populations. Female status has been associated with precocious declines in fertility and mortality, and with a lack of gender discrimination in child mortality (e.g. Caldwell, 1986; Dyson & Moore, 1983; Jejeebhoy, 1995). Where a woman has more control over decision-making, and where she can choose to take up paid
work or education outside the home, it is argued that these capabilities will empower her to seek medical and reproductive information and services for herself and her children without requiring approval of her partner and other household members, and will also increase the opportunity costs of high fertility for the household. While the pathways of influence of female autonomy on fertility and mortality identified in the late twentieth century were unlikely to operate where the causes of mortality were less obvious and recourse to effective medical and contraceptive services largely unavailable, it seems inescapable that where women were strongly discriminated against then they suffered health consequences and barriers that would have negatively influenced the health of their children. Conversely, it is likely that women who were free to work and live as independent adults in other households before childbearing were in a position to invest more in their offspring, and to promote the health of both male and female children. Relatively high female autonomy, therefore, was probably a pre-condition of early declines in child mortality and fertility. However, this autonomy variable seems to cut across the family formation patterns. As we have seen, women’s autonomy was high in south India as well as in Japan compared with their northern counterparts. Such comparison led two East Asian specialists to a conclusion that ‘the contrast between Western and Eastern Europe may have an East Asian parallel. From the global perspective necessary to see such broad patterns, it appears that China is to Japan as Eastern is to Western Europe’ (Wolf and Hanley 1985, pp. 3-4). But each of the pairs was rather unique in the sense that there was no apparent correlation with marriage or household formation variables (Saito 2014a).

Turning back to the mortality side, pre-transition urban mortality scenes were dominated by migration-related and density-dependent factors. Mortality rates were much higher in urban than in rural populations, and urbanisation therefore acted to raise average mortality rates (the so-called urban graveyard effect), both through the redistribution of population to high mortality environments and, in the early stages, through increased circulation of infectious diseases via transport and trade networks. Critically, the onset of mortality decline and the demographic transition was marked by dramatic reductions in urban mortality, falls that have been argued to have been essential to modern rates of urbanisation (de Vries 1984; Dyson 2011). In towns and cities generally, infant and child mortality was very high, but recent immigrants of all ages also suffered high mortality because they encountered relative lethal diseases against which they had acquired no immunity in childhood. Urban mortality rates therefore varied substantially depending on cultural habits (especially with respect to infant feeding practices) and prevailing pathogens. However, what was at work in the urban society was in sharp contrast with the rural sector before the transition: urban mortality tended to rise when trade was brisk. As Richard Smith has
suggested for an extremely ‘open’ society and economy like England in the first half of the seventeenth century, openness meant a greater impact of infectious diseases imported from abroad, for their domestic circulation was ‘greatly intensified by the burgeoning urban population and by the willingness or ability of the population to move geographically’ (Smith 2001, p. 204).

This pattern of rising mortality with increasing trade and urbanisation is well-described in the influential work of William McNeill (1976). However McNeill and others also argued that the rises in mortality following the introduction of new diseases were followed by a gradual process of accommodation between host and pathogen that involved a decline in pathogen virulence. Thus, as once rare epidemic diseases became childhood diseases they also became less lethal (McNeill 1976; Kunitz 1983; Walter and Schofield 1989). This assumption was based on the recognition of an evolutionary trade-off between virulence and transmission, because pathogens that disable or kill their hosts are less likely to be transmitted to a new host. However, over the last 30 years this evolutionary paradigm has been overturned with the recognition of the complexity of selective forces operating on the determinants of virulence. In fact, for the most lethal diseases of the early modern period, including smallpox, plague, typhus, typhoid, cholera, malaria, and yellow fever, there was no necessary selection for avirulence, because they did not depend on ambulatory human victims for transmission. Crucially though, these particularly lethal diseases are also relatively susceptible to the disruption of transmission by human interventions (including quarantine, isolation, vaccination, control of insect vectors, and clean water supplies). Because these diseases were relatively lethal, crude but successful preventative measures against them produced disproportionately large gains in longevity (Davenport 2015).

Reductions in mortality from the most virulent diseases therefore tended to precede reductions in mortality from less lethal but highly infectious ‘childhood’ diseases. These less lethal diseases (‘childhood’ infections, respiratory and diarrhoeal pathogens) proved more difficult to control. They flourished in high density populations and were fatal mainly to the young, the elderly and the malnourished. Their high infectiousness and person-to-person transmission made prevention difficult, and therefore reductions in mortality required improvements in nutritional status (to increase resistance), sanitation and personal hygiene. It is therefore no surprise that late nineteenth- and early twentieth-century reformers of public health focused on slum conditions as well as child and maternal health. Therefore the early stages of mortality decline involved a transition from a regime of relatively high adult mortality, with little distinction by wealth, to one where mortality was increasingly concentrated at the extremes of life, and amongst the poor. Control of famine-related mortality contributed to
this transition, because young adults were also at highest relative risk during these types of crises (Dyson and Ó Gráda 2002, p. 10).

A key consequence of the mortality transition we have described here is that the risks of migration to cities altered profoundly as a consequence of both disease endemicisation and prevention. For example before the late eighteenth century young adult migrants to London were at considerable risk of smallpox, because many came from rural areas where smallpox was infrequent (Davenport et al., 2011). The advent of inoculation and then vaccination eliminated this risk. Similar arguments apply to other acute diseases that either endemicised (measles, whooping cough) or were controlled (plague, typhoid, cholera, malaria). Therefore the risks of rural-urban migration were substantially reduced by the transition from acute diseases of wide effect to chronic and childhood diseases. This transition also had important implications for the impact of migration on urban mortality levels. In the case of tuberculosis, a chronic infection that rose in importance, and in some cases, prevalence, in the nineteenth century, migrants could make behavioural adjustments that were not possible in response to acute diseases like smallpox. Those already clinically infected with tuberculosis were unlikely to migrate to towns, but if healthy migrants to towns contracted the disease there then the slow progression of the disease made it possible to return, or be sent, home to die (and infect others), deflating measures of urban mortality, and increasing tuberculosis transmission and mortality in rural populations (a ‘salmon bias’ effect: Welton, 1872; Johansson & Mosk, 1987).

All this implies that the classical formulation of demographic transition has interpretational problems. First, it has been assumed that the pre-transition level of mortality was invariably high and the initial phase of population increase was accounted for solely by reductions in death rates, but we now know that the process of mortality decline tended to be much longer than previously thought. Second, there were cases in which fertility increased before the onset of modern fertility decline, suggesting that in those countries the burden of population pressure was heightened although its impact would never have outweighed that of mortality decline. Third, the prolonged process of mortality decline meant that its effect on the burden of population pressure could have been a complex one. The ‘pressure’ is comprised of two components – one comes from the rising rate of growth and the other from the increasing dependency ratio. While the immediate effect of a delay in decline of pandemic and/or famine mortality was to lower the rate of population increase, it implied that adult mortality remained relatively high and, depending on how the death rates in other age groups would change, the direction and magnitude of change in the dependency ratio could vary. Consider, for example, two diagonally different cases. One is England in the industrial revolution period, whose agricultural progress had already led to
an early attenuation of mortality crises and also to a modest rise in marital fertility; moreover, while the level of urban death rates remained high, the initial process was followed by a gradual decline in childhood – if not infant – mortality. In this case, the rate of population growth and the dependency ratio both rose. The other case is India in the pre-1920 period, where its economy was characterised by no agricultural development and its demography by high levels of infant and child mortality as well as frequent famine and epidemic outbreaks, and this represents a low population growth case with a moderate dependency ratio. The former, on the face of it, would have meant a tremendous difficulty when the country was to enter the phase of modern economic growth, while the population pressure in the latter case must have been lower. Despite all this, as we all know, England’s economic performance was far superior to India’s. The issue of population pressure in the transition to modern economic growth is far more complex than we tend to assume.

2. Population change and economic growth 1850-1950

This consideration raises a question about the link between economy and population. Livi-Bacci has graphically shown relationships between the two demographic indicators and real GDP per capita for 16 advanced countries in five benchmark years between 1870 and 2000 (2012, figures 4.4 and 4.8). One is a rising but non-linear curve for life expectancy and the other a similarly non-linear but declining curve for TFR. The non-linearity means that the curve in each graph becomes flatter after 1950, suggesting that as people become affluent, further income growth will make virtually no impact on the demographic indicators. On the other hand, it is expected that in the first phase of the transition, improved material well-being did lead to increased order and efficiency in demography. However, a close look at figures 1 and 2, which we have drawn from a different set of data to focus on the period before c.1950, reveals that in the nineteenth century the correlation seems not quite apparent.

[Figures 1 and 2 about here]

Thus, one may question if the causal link was invariably from income to demography; we believe that there is an alternative interpretation. The two are shown in figure 3 as simplified process charts, in which the boxes represent

---

2 The curve for life expectancy and real GDP per capita is often called the ‘Preston curve’, named after the demographer Samuel Preston who first showed that there was a non-linear, cross-sectional relationship between the two. His aim of fitting such a curve to country data was to explore the relative contribution of economic factors to increases in life expectancy, not the other way round. See Preston (1975).
processes (i.e. mortality decline, fertility decline and economic growth) and the arrows indicate functional relations between the boxes with the plus sign being an increasing function and the minus one a decreasing function. Interpretation A is a version of the classical theory of demographic transition, which focuses on economic growth influencing both mortality and fertility processes. According to this interpretation, it was the growth of GDP per capita that acted as an initiator in the transition to both improved survival and fertility decline. Interpretation B reverses the direction of the arrow between the mortality and economic growth boxes, with other connecting arrows kept unchanged. Given the premise that longevity is an important element for the formation of human capabilities and hence a component of human development index (HDI), a composite measure created by the United Nations (UNDP 1990), one can argue that in the period under review the causal link between mortality and material well-being was from life expectancy to GDP per capita. Since improved survival is also one of the pre-conditions for fertility decline (Cleland 2001), it was the improvement of people’s health and longevity, not the rise of their material well-being, which set off the dual process of transformation in economy and demography.

[Figure 3 about here]
benchmark years; in our analysis, on the other hand, 1850 is added but the examination is limited until 1950 in order to focus on the period for which per capita GDP growth is supposed to have been critically important. Countries covered include developing as well as developed countries as far as 1913 and 1950 are concerned, totalling 24 for each of these two sub-sets. The number of countries is reduced to 18 in 1870 and to 8 in 1850; the 1870 sample includes India and Japan, but the 1850 sub-set does not include any non-European countries (for the complete list of countries covered, see figure 1 above).

In the first of four graphs (figure 4), gains in life expectancy at birth are plotted against the corresponding real, log-transformed GDP per capita estimates. Clearly there occurred an upward shift in the average gain in longevity from 1850-1870 to 1870-1913, then to 1913-1950, but for each of the three periods life expectancy seems to have been insensitive to a change in the initial level of material well-being. The upward shift is likely to have been driven by forces other than GDP per capita (this statement remains unaffected even if the log form is replaced by the measure of percentage change). Then, the vertical and horizontal axes are transposed with life expectancy at the initial date on the horizontal and percentage change in GDP per capita on the vertical axis (figure 5). This graph indicates a loosely positive relationship, but a closer look at its scatter plot reveals that the looseness is caused by a rightward shift of a trend line from the pre-1913 to the post-1913 period. The ‘other’ forces at work in figure 4 may have also accounted for this shift, but in figure 5 the positive relationship between the two variables proves to be statistically significant if the effect of the rightward shift is controlled for: the higher the level of life expectancy initially, the greater its positive impact on the rate of increase in GDP per capita.

The second issue is whether or not the initial level of life expectancy had an expected impact on fertility decline. There are two main mechanisms. First, in accordance with orthodox economic accounts of fertility decline, gains in child survival increased incentives to invest in the nurture and education of children, which raised the costs of childbearing and provided higher motivation to restrict their numbers. Second, fertility declines may represent a delayed response to the increases in average family size caused by improving child survival rates. This explanation has often been dismissed in the literature as easily disproven (which is why the arrow connecting mortality decline to fertility decline is bordered with a dotted line in figure 3). However, it has rarely been properly tested. In
most studies current or lagged infant mortality has been used as a measure of child survival. This approach presupposes that parents are aware of current levels of infant mortality, and that only infant survival is taken into account in parental considerations of child survival. In England, infant mortality only fell in the aggregate after 1900, although child mortality fell significantly after c.1870, coincident with fertility declines. Cleland (2001) argued plausibly that rises in the average numbers of children per family eventually saturated traditional avenues for the redistribution of excess children (via adoption, fostering, apprenticeship and other means), forcing parents to devise novel means of limiting the numbers of child dependents. Thus, figure 6 shows life expectancy at the initial date on the horizontal axis and the difference between TFRs in the beginning and ending years of each period on the vertical. The slope the scatter diagram suggests is an expected one but the relationship appears rather weak. However, if controlled for the initial level of TFR, then the effect of life expectancy on the degree of decline in TFR becomes statistically significant.

The third, related issue concerns how well economic growth explains fertility decline. As suggested above, orthodox economic theory predicts that the ‘price’ of children and the ‘shadow price’ of women’s time will lead to fertility decline. Figure 7 shows that the decline in TFR (measured on the vertical axis) was rather loosely correlated with the level of GDP per capita at initial date (on the horizontal axis), but again if the initial level of TFR is controlled for, then the significance of the estimated regression coefficient for log-transformed GDP per capita becomes statistically significant at the 5 per cent level.

By way of concluding this section, three points may be made. First, mortality decline was indeed the initiator of the dual process of transition: first, reductions in children’s mortality paved the way for a fertility transition yet to come; and second, improved survival, despite its immediate effect being to enlarge the denominator of GDP per capita, did help to get modern economic growth started and, moreover, to generate further growth through its enhancement effect on people’s physical and mental capabilities. The mortality transition, second, was largely a product of two mutually independent processes, i.e. agricultural progress and the disappearance of the urban graveyard effect. Both, in many countries, started in an early modern setting – well before the rise of modern medicine.

On the fertility side, third, its pre-transitional levels mattered when accounting for the fertility transition. Decades ago, having summarised the evidence on fertility patterns in historic Europe, Knodel and van de Walle (1979) made it clear that the level of marital fertility individual countries exhibited before the onset of fertility decline was highly variable and seems to have had no
consistency in the level of economic development. Cultural factors must have played a role in accounting for the diversity in the fertility transition. One such cultural factor is the status of women, which, with other variables, is believed to have affected the course of fertility change ‘independently’ of economic and social conditions.

3. Demographic contributions to economic growth: two issues

(i) Dependency ratios and the demographic window of opportunity

The adjustment from high to low rates of fertility and mortality took over a hundred years in those societies that experienced mortality declines in the nineteenth century, but was accomplished in under fifty years in many countries that entered the transition in the twentieth century. One consequence of the relatively slow changes in vital rates in European and neo-European populations is that age structural changes were also modest, suggesting that the increase of young dependents may not have been particularly rapid compared with the working-age population.

[Figures 8 and 9 about here]

Figure 8 compares the dependency ratios, a simple measure of the number of potential dependents per 100 adults of working age (here ages 15-59), for selected countries in the nineteenth and twentieth centuries. Dependency ratios are a crude means of gauging the potential impact of demographic conditions on productivity, on the assumption that relatively low ratios of dependents to adults of working age favour savings and investment, and stimulate productivity. Whenever such a ‘demographic dividend’ accrues, the period may be called a ‘demographic window of opportunity’ (Bloom, Canning and Sevilla 2003).

Clearly, of the countries that experienced early industrialisation, only France had favourable dependency ratios in the nineteenth century. Most European populations experienced historically high dependency ratios mainly reflecting the burden of young children. The first industrialiser, England, was not an exception: the classical industrial revolution took place not only in the face of mounting population pressure, but also ‘in despite of a changing age structure which, as it grew younger, became progressively less favourable to a surplus of production over consumption and to the generation of a demand for non-agricultural goods’ (Wrigley and Schofield 1981, p. 449). On the other hand, the relatively modest growth rates of most countries in this period meant that youth dependency ratios did not reach the extremes experienced by many developing countries in the twentieth century, with the exception of the United States (Figure 9). Estimates
available for England and Sweden allow us to go back to earlier centuries. Figure 10 reveals that while the relatively high ratios of children to adults in the nineteenth century were not historically exceptional, there was a long phase of upswing before reaching the nineteenth-century peak. In other words, England and Wales in the second half of the seventeenth century enjoyed a combination of a low rate of population growth and an unusually favourable dependency ratio, suggesting that population pressure, often portrayed as Malthusian, was not a situation characteristic of the entire early modern period.

To conclude, first, for early industrialisers the transition to modern economic growth began in a period when demographic conditions remained unfavourable in terms of sheer rates of growth and age structure. But, second, in many countries of the eighteenth and nineteenth centuries, the demographic onus was not extreme. On the other hand, there were moments in which countries enjoyed a demographic bonus: England in the seventeenth century, France in the nineteenth, and many European countries in the early twentieth as well as Japan in the mid-twentieth. However, in the twentieth-century cases the demographic window of opportunity seems to have been short-lived, while whether or not England and France in earlier environments were in a position to make effective use of that demographic dividend for future development is another question (for the seventeenth-century English case, see comments by Wrigley and Schofield 1981, p. 449).

Finally, it should be remembered that dependency ratios calculated in this way are a poor measure of the potential of populations to work, save and invest. While estimates of labour force participation and productivity are beyond the scope of this chapter, we will consider here the implications of the demographic transition for population health and labour productivity.

Health status and labour productivity

Falls in mortality are often but not always associated with improvements in population health. Where health did improve, initially through reduced exposure to disease and dearth, then such improvements could have increased productivity, through greater strength and endurance, enhanced cognitive capacities, and less time spent in ill health. Fogel and colleagues have gone so far as to argue that the process of economic development was one of ‘technophysio-evolution’, where gains in human physical and cognitive capacities drove gains in output that promoted further improvements in health in a virtuous spiral of increasing productivity (Floud et al., 2011). However health need not improve in tandem
with life expectancy, if the causes of mortality are acute diseases with few after-effects, or if increasing survival leaves more of the population in a state of debility.

The evidence for changes in population health derives mostly from anthropometric studies of adult male heights, data that become more abundant from the second half of the nineteenth century. While individual heights are strongly affected by genetic factors, for large populations, inter-population variation in average adult heights are a (complex) outcome of maternal health and nutritional status in childhood. Nutritional status is determined by the quality and quantity of food intake, and energy demands arising from work and disease. In addition to sheer calorific intake, growth is sensitive to the intake of specific nutrients, especially protein and calcium, and heights may therefore differ also between populations because of ecological and economic factors affecting cattle:human ratios and access to milk, or as a result of dietary preferences (Baten & Blum, 2014). The mean heights of populations may also vary according to their migrant composition, because heights of adult migrants reflect conditions in the area where they spent their childhood. Migrants may also be positively selected for tallness (Humphries & Leunig, 2009; Baten et al., 2010).

In European populations where heights data derive from conscript armies then average adult male heights appear to have risen fairly steadily from at least the mid-nineteenth century (Figure 11a). However in Anglophone populations, where heights data derive mainly from volunteer armies, slaves (for the U.S.) and prison populations, an initial advantage in stature in cohorts born before c.1830 was reversed for cohorts born between c.1830 and c.1850 (for the U.K.) or c.1830-1890 (for the U.S.) (Figure 11b). This reversal coincided with the onset of secular increases in real wages, and has been argued to indicate a trade-off between income and health during the Industrial Revolution, a phenomenon dubbed the ‘industrialisation puzzle’ or the ‘antebellum paradox’. In this scenario urbanisation and industrialisation were associated with increased exposure to infectious diseases, higher workloads and reduced quality of diets. A compelling alternative explanation is that the composition of military volunteers fluctuated with alternative opportunities in the labour market, making military service or crime relatively unattractive when wages were buoyant (Bodenhorn et al., 2017). If this were indeed the case then the relatively tall stature of Anglophone recruits and criminals, despite their low social status, would suggest that the Anglophone populations from which military volunteers were drawn were indeed substantially taller than continental European populations throughout the nineteenth century. Such a conclusion would strengthen the observed relationship between GDP per capita and male heights in the nineteenth century, in which Anglophone samples otherwise present an anomaly (Baten, 2000).
Few comparable data exist for non-European populations. Figure 11a includes data only for populations where the sampling process used to obtain height measurements was considered relatively unaffected by potential selection biases. Heights of the Chinese male population of Taiwan, measured under Japanese rule, showed some deterioration amongst cohorts born in the 1890s, and no overall trend. The population consisted largely of relatively recent migrants from Southern China, who may have been taller than the populations which they left. The Japanese data derive from military conscripts and indicate that the Japanese population was markedly shorter than European populations in the same period, but also experienced rapid growth during the period of rapid economic development after 1870.

Taken together, these data indicate that heights of conscripted males in a number of national populations rose in tandem with life expectancy, urbanisation and economic development in the second half of the nineteenth century, suggesting synergistic relationships between these variables (e.g. Steckel & Floud, 1997). However whether these positive relationships held before the mid-nineteenth century is unclear. Sources for the investigation of heights of cohorts born before the early to mid-nineteenth century are relatively fragmentary and liable to selection biases, which may explain some of the apparently trendless fluctuations in height in eighteenth century populations. If we accept that the male populations of Britain and America were taller than their contemporaries by the early nineteenth century then it remains unclear when this advantage emerged. Americans enjoyed greater and more equitable access to food and land, and suffered lower exposure to infectious diseases as a consequence of isolation and low population densities. Britain however experienced rapid population growth and urbanisation, factors expected to promote disease exposure, and to dampen rises in per capita consumption. Nevertheless as we have argued, mortality fell especially in cities, and this together with increasing agricultural output and the safety net provided by the English poor laws may have improved population health despite apparently unpromising conditions. Two changes in particular are likely to have promoted gains in stature and health, before the mid-nineteenth century.

First, although early declines in mortality benefited adults more than children, improvements in maternal health also conferred large benefits on very young infants. In England adult mortality, maternal mortality, stillbirths and early neonatal mortality all fell in tandem across the eighteenth century, consistent with strong improvements in maternal health (Wrigley et al., 1997; Smith & Oeppen, 2006; Woods, 2009). This had important implications for stature and child development, because gestation is one of the most critical developmental periods, when insults in utero can result in premature birth or growth retardation with enduring negative consequences for height (Luo & Karlberg, 2000; Cole,
The most important environmental effects on foetal development are maternal infection and severe nutritional deprivation, factors that changed markedly in the initial stages of the mortality transition. The recession of famines reduced the incidence of severe maternal nutritional deficits. Endemicisation and the control of the more lethal diseases reduced the risk of infections in pregnancy, and also increased maternal immunity and therefore the transfer of passive acquired immunity against childhood diseases to the foetus, with protective effects in early infancy. The risk of infections in adulthood was highest to young adult rural-urban migrants, and reductions in rural-urban disease gradients contributed to a marked convergence in neonatal mortality between rural and urban populations (Smith & Oeppen, 2006). Increases in maternal breastfeeding, evident in London after 1750, were also key to improving infant nutrition and reducing excess urban mortality (Landers, 1993).

The second relevant effect of early mortality declines on adult heights is via their impact on survival prospects of rural-urban migrants. What slim historical evidence we have suggests that rural to urban migrants were generally positively selected for skills and probably height and robustness. However before the dramatic reductions in urban mortality that commenced in the mid-eighteenth century in north-western Europe, migration to towns entailed a high mortality risk from diseases against which height and good nutritional status offered no protection (including plague, smallpox, typhus and typhoid). Migration to towns therefore operated selectively to eliminate the most robust adults from the population. Conversely, declines in urban mortality reduced the trade-offs between longevity and wealth, and reduced negative selection against tallness and other positive migrant attributes, allowing the emergence of a positive association between health and longevity, with positive consequences for human capital formation.

The early gains in population health that we have outlined here were probably enough to counterbalance the malign effects of urbanisation, at least in the early stages of industrialisation in Britain. However there may have been periods where the rapidity of urbanisation outweighed improvements in health. A very common finding in research on stature is that urban populations were shorter than rural ones. This was the case for Britain, Sweden, the U.S., Australia and Japan in the nineteenth century, although apparently not for parts of Germany and Spain (Baten, 2009; Floud et al., 1990 chap. 5; Shay, 1994; Sandberg & Steckel, 1997; Twarog, 1997; Whitwell et al., 1997). At the national level, the level of urbanisation has been argued to explain much of the cross-sectional variation in height between countries (Steckel & Floud, 1997, chap. 11). However urbanisation was no impediment to height gains in the late nineteenth century, and rapid rates of urbanisation were associated with concomitant increases in stature, and a convergence (or crossover) between rural and urban
heights. This waning of an urban height penalty is consistent with the progressive disappearance of excess urban mortality in these populations by the early twentieth century. According to our argument, the disappearance of the urban height penalty represents the final stage in a progressive reduction in rural-urban differentials in health and mortality, that began for north-western European populations in the mid-eighteenth century. There may nonetheless have been periods of reversal. Life expectancy and infant mortality rates were relatively stagnant in England between 1820 and 1860, and worsened in the U.S., in association with rapid rates of urbanisation in the early nineteenth century. Woods (1985) attributed the British pattern to the extensive redistribution of the population from relatively healthy rural to unhealthy urban areas, against a backdrop of gradually improving mortality in most settlement types. In this case rapid population redistribution could have resulted in lower average stature even if rural-urban differentials in height were narrowing. This contrasts with fairly continuous improvements in life expectancy and infant mortality after 1750 in Sweden, where urbanisation was negligible before the late nineteenth century (and where heights of conscripts rose monotonically for birth cohorts born after 1840).

Finally, adult heights reflect maternal health and living conditions in childhood, and may be an inadequate guide to adult health. However more direct evidence of trends in the health of adults is elusive. Sources that report sickness are difficult to interpret both because the sources are selective, and because ill health is subjective and may display what Johansson termed a ‘cultural inflation of morbidity’ (Johansson, 1991). Drawing on evidence from working men’s friendly societies (cooperative associations that provided sickness benefits and sometimes healthcare), Riley argued in his book ‘Sick not Dead’ (1997) that mortality improvements in nineteenth century Britain were associated with increasing levels of morbidity. Illnesses became less frequent, but the average length of sick spells increased. More recent work on similar sources, adjusted for age, has not supported Riley’s (or Johansson’s) conclusions (e.g. Harris et al., 2012).

Taken together, the evidence from adult male heights suggests both wide variation in nutritional status between populations, and substantial improvements in the economically most advanced countries over the course of the nineteenth century. It is likely that males in Britain and the Anglophone neo-Europes enjoyed a substantial advantage in height over other European populations by 1800. While heights should not be used as an absolute indicator of adult capacity in cross-country comparisons, the growth in adult male stature of nineteenth century European populations suggest significant improvements in population health and therefore in human capital in this period. Improvements in urban mortality would have increased the returns to height and health, by reducing the risks of migration.
4. Conclusion

Not all economies that achieved economic growth during the period in question were characterised by invariably high mortality and fertility. The initial levels of the two varied considerably from culture to culture. But what was common to all those economies was the progressive impact that rising life expectancy and associated population growth exerted on the growth process, on the one hand, and on the demographic transition process from disorder to order and from waste to efficiency, on the other.

The immediate effect of the decline in mortality must have been a pressure on the standard of living since the resultant increase in total population tended to increase food prices and, hence, to decrease real wages while the worsened dependency ratio was another burden to the working population. Eventually, however, increased longevity resulted in better growth performance through its positive effects on population health and human capital accumulation. Moreover, the mortality transition was accompanied by shifts in major causes of death. When comparatively lethal infectious diseases were major killers, nutritional intake did not afford much protection. It was only when they were under control that a positive feedback between longevity and living standards started working – through increased nutrition and better public health services the state began to provide. In his seminal article, Preston (1975) suggested that an increase in income accounted only for 10-25 per cent of the rise in life expectancy, but this positive effect is likely to have become a little clearer in later stages of the transition. Also likely is that the decline in mortality in general, and probably that in under-5 mortality in particular, led to fertility decline in later stages, although only after 1870.

We have characterised mortality declines before 1870 as a function of two overlapping processes. First, famine control and the consequent stabilisation of mortality was often accompanied by increasing disease exposure (as a consequence of economic and epidemiological integration). Second, control of the most lethal diseases made towns in particular less dangerous and facilitated rapid urbanisation. These two processes were sequential in England but overlapped elsewhere. Critically, these processes were a ubiquitous feature of demographic transition, occurring virtually everywhere before 1950. In parts of India, Africa and Oceania mortality appears to have risen in the nineteenth century as a consequence of increasing trade and the impacts of colonisation, and then underwent a very marked stabilisation and decline in mortality after c.1920.
as a consequence of control measures against both famine and the most lethal epidemic diseases (especially plague, smallpox and cholera) (Doyle, 2013, chap. 2; Dyson, 1989; Dyson & das Gupta, 2001). Similar measures were implemented slightly earlier in parts of Latin America. These improvements in survival were often associated with increases in fertility (and fecundity), and account for the very rapid rates of population growth evident in most developing country populations by 1950. Fertility declines were also more rapid in these populations than amongst the forerunners in the demographic transition (with some marked mainly sub-Saharan exceptions), and show similar variety with respect to levels of economic development.
References


Davenport, R.J. 2015. The first stages of the mortality transition in England: a perspective from evolutionary biology, Centre for Competitive Advantage in the Global Economy (CAGE) working paper, University of Warwick.
Dyson, T. 2011. The role of the demographic transition in the process of urbanization, Population and Development Review 37, S1: 34-54.


Law, C. M. 1967. The growth of urban population in England and Wales, 1801-1911. Transactions, Institute of British Geographers, 41, 125-43


Tomobe, K. 2001. The level of fertility in Tokugawa and Meiji Japan, c.1800-


Figure 1. The Preston curve. Real GDP per capita (1990 international dollars) and life expectancy at birth ($e_0$): four cross-sections combined, 1850-1950

Sources: Crafts (1997, 2002) and van Zanden et al. (2014), except for Japan’s per capita GDP estimates, which are taken from Settsu, Bassino and Fukao (2016), table A.9, and Saito and Takashima (2016), table 2. Names of countries for which data are available, listed by benchmark year: 1913 and 1950: Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, Denmark, Finland, France, Germany, India, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, UK, USA (24 countries). 1870: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, India, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, UK, USA (18 countries). 1850: Belgium, Canada, Denmark, France, Netherlands, Norway, Sweden, UK (8 countries).
Figure 2. Total fertility rate (TFR) and real GDP per capita (1990 international dollars): four cross-sections combined, 1850-1950

Sources: See figure 1 above.
Figure 3. Two interpretations of the relationship between economic growth and the demographic transition

Interpretation A

Interpretation B
Figure 4. Real GDP per capita at initial date and the gain in life expectancy at birth \((e_0)\): three time-periods combined, 1850-1950

Sources: See figure 1 above.
Note: The coefficient of \(\ln\) (GDP per capita) at initial date is statistically not significant.
Figure 5. Life expectancy at birth ($e_0$) at initial date and change in real GDP per capita: three time-periods combined, 1850-1950

Sources: See figure 1 above.
Note: The coefficient of $e_0$ at initial date is statistically significant at the 5 per cent level (with the effect of time periods controlled for).
Figure 6. Life expectancy at initial date ($e_0$) and the decline in TFR: three time-periods combined, 1850-1950

Sources: See figure 1 above.
Note: The coefficient of $e_0$ at initial date is statistically significant at the 1 per cent level (if controlled for TFR at initial date together with time periods).
Figure 7. Real GDP per capita at initial date and the decline in TFR: three time-periods combined, 1850-1950

Sources: See figure 1 above.

Note: The coefficient of Ln (GDP per capita) at initial date is statistically significant at the 5 per cent level (if controlled for TFR at initial date together with time periods).
Figure 9. Youth dependency ratios (children aged 0-14 per 100 adults aged 15-59)

Sources: see Figure 8.

Figure 10. Youth dependency ratios, Sweden, and England and Wales.

Sources: see Figure 8.
Figure 11. Estimated average adult male height, national populations.

Sources: Drukker & Tassenaar, 1997, Tables 9A.1, 9A.2; Floud et al., 1990, Table 4.1; Fogel, 1986, Table 9.A.1; María-Dolores & Martínez-Carrión, 2011, Table 1; Olds, 2003, Table 3; Sandberg & Steckel 1997: Table 4.1; Shay, 1994; Weir, 1997, Table 5B.1.