Contribution of Health to Economic Development:
A Survey and Overview

Muhammad Jami Husain
Keele University, United Kingdom

Please cite the corresponding journal article:
http://www.economics-ejournal.org/economics/journalarticles/2010-14

Abstract
The policies for better health, poverty reduction, and less inequality, throughout the world, require thorough understanding of both the processes and causal paths that underlie the intricate relationship between health and wealth (income). This is deemed difficult, contingent, and only partially understood. The adage ‘health is wealth’ is still, primarily, an intuitive proposition. A vast majority of researchers instead present theoretical and empirical arguments of the reverse proposition, i.e. ‘wealth is health’. A recent strand of the literature, however, reflects changes in the perceptions: improvements of health and longevity are no longer viewed as a mere end- or by-product of economic development; but argued as one of the key determinants of, and therefore means to achieve, economic development and poverty reduction. Hence, better health does not have to wait for an improved economy; rather, measures to reduce the burden of disease, to give children healthy childhoods, to increase life expectancy etc. will in themselves contribute to creating richer economies. Drawing on the traditional and emerging perspectives on the health-income relationship, this literature review presents a non-exhaustive survey of existing methodological approaches and their results that are applied to track and measure how health influences economic outcomes.

Submitted as Survey and Overview

JEL: I10, O40, J11
Keywords: Health; income; economic growth; life expectancy; mortality; causality

Correspondence
M. Jami Husain, Research Institute for Public Policy & Management (IPPM), Claus Moser Research Centre (Room: CM 2.02), Keele University, Staffordshire, ST5 5BG.
e-mail: m.j.husain@ippm.keele.ac.uk; ecojami@yahoo.com

The author gratefully acknowledges the comments and constructive suggestions from Gauthier Lanot, Peter Lawrence and Tim Worrall. The responsibility of any error rests with the author

© Author(s) 2009. Licensed under a Creative Commons License - Attribution-NonCommercial 2.0 Germany
Contribution of Health to Economic Development: A Survey and Overview

1. INTRODUCTION

The adage ‘health is wealth’ is still, primarily, an intuitive proposition. A vast majority of researchers (e.g. Mckeown, 1976; Prichett and Summers, 1993) instead presents theoretical and empirical arguments of the reverse proposition, i.e. ‘wealth is health’. The latter view also attributes credit for recent improvement of health status to ‘higher incomes’ in promoting technical progress and dissemination of new health technologies (Easterlin, 1999; Cutler et al., 2006). Furthermore, the thrust of contemporary discussions on health reforms typically sees interventions that promote health and the delivery of health care as costs that need to be contained - implying that income is the main instrument and health outcomes the end-points of “development” objectives (Suhrcke et al., 2005). This may underestimate the role that health plays in economic development. Notwithstanding widespread recognition that population health is an important factor in strengthening economies and reducing poverty, health and longevity have remained prominently as a subject in the field of epidemiology and demographics - with intermittent attempts by the economists to link it to the mainstream of economics.

A recent strand of the literature (see for instance Suhrcke et al., 2005; Alsan et al., 2006; Bloom and Canning, 2003a; Bloom, Canning and Sevilla, 2004; CMH, 2001; WHO, 2002) reflects changes in the perceptions: improvements of health and longevity are no longer viewed as a mere end- or by-product of economic development; but argued as one of the key determinants of, and therefore means to achieve, economic development and poverty reduction. Hence, better health does not have to wait for an improved economy; rather, measures to reduce the burden of disease, to give children healthy childhoods, to increase life expectancy etc. will in themselves contribute to creating richer economies (Alsan et al., 2006).

What is still required from the research agenda is to justify this proposition with a sound theoretical and empirical basis, and measure quantitatively the ways health contributes to the improvement of economic outcomes. The economic valuation of health improvements provides new insights, and improves our understanding on why and how devoting resources to health improvements can be a powerful means of sustaining economic growth and mitigating poverty.

Despite remarkable improvements with regards to various population health indicators since the second half of the 20th century, the fight against poor health and diseases continues. In particular, the common goals for the developing countries are set within the world’s overarching framework for reducing poverty and deprivation- the UN’s eight Millennium Development Goals (MDGs). Nearly half of these goals concerns different aspects of health: to cut in half the proportion of people who suffer from hunger between 1990 and 2015; to reduce child mortality (the under-five mortality rate) by two thirds by 2015 from its 1990
level; to reduce by three quarters, between 1990 and 2015, the maternal mortality ratio; to halt by 2015 and begin to reverse the spread of HIV/AIDS, Malaria and other diseases; and to cut in half the proportion of people without sustainable access to safe drinking water by 2015 (Wagstaff and Claeson, 2004; Sachs, 2004). Are these goals the end-point of the development objectives, or pave the way for enhanced growth performance by setting the countries on permanent growth trajectories with a virtuous cycle? The policies for better health, poverty reduction, and less inequality, throughout the world, require thorough understanding of both the processes and causal paths that underlie the intricate relationship between health and wealth (income). This is deemed difficult, contingent, and only partially understood. Drawing on the traditional and emerging perspectives on the health-income relationship, this literature review presents a non-exhaustive survey of existing methodological approaches and their results that are applied to track and measure how health influences economic outcomes.

2. PERSPECTIVES ON THE HEALTH - WEALTH RELATIONSHIP

The traditional view describes the nexus between health and wealth (or income) as a ‘gradient’ to express the gradual association between the two. Causality is assumed to run from income to health, i.e. describes how health improves with income throughout the income distribution. In particular, the common finding is that the absolute reduction in mortality for each dollar of income is much larger at the bottom than at the top of the income distribution (Deaton, 2002, p. 14). Health differentials are also frequently expressed in terms of other correlates, e.g. race, geography, occupational grade, rank, socio-economic status etc., which are often described under the caption of ‘inequalities in health’ (Kawachi et al., 1999).

What causes the gradient? One line of argument highlights the primacy of income (and socioeconomic status) in determining the health conditions; others emphasize the effect of health on income determination - often referred to in the literature as selection or reverse causation (Goldman, 2001; Deaton, 2002). Accordingly, economists used contesting captions in emphasizing their line of thoughts - e.g. “Wealthier is Healthier” (Prichett and Summers, 1993) as opposed to “Healthy Bodies and Thick Wallets…” (Smith, 1999). Many studies also propose an explanation based on ‘access’ – if better-educated, richer, or lighter skinned people have better access to health care, and if health care has a major effect on mortality and morbidity, then education, wealth, or race will predict health outcomes (Deaton, 2002, p.17). Other arguments highlight the role of health-related behavior (e.g. use of tobacco, alcohol, drugs, sex etc.) and the effect of life-saving technology in determining the gradient. While the evidences on the gradients may be numerous, the underlying mechanisms of such correlations are not yet disentangled. Indeed the mirage of the health-wealth relationship invoked numerous researches, with diverse conclusions in terms of ascribing primacy to various factors.

In any case, health plays an obvious role in determining physical capacities (e.g. strength, stamina, and endurance) and mental capacities (e.g. cognitive functioning, reasoning ability, intelligence). Health is therefore an important form of human capital, the improvement of
which, *ceteris paribus*, should enhance workers’ productivity. This form of health-productivity relationship at the microeconomic level has been established in different contexts (see for example, Schultz, 2005; Strauss and Thomas, 1998; Suhrche et al., 2005, Savedoff and Schultz, 2000). If productivity and income are positively correlated and if good health increases productivity, then it should also lead to higher income. The extent to which the positive correlation between health and economic prosperity reflects a causal effect of health on economic outcome at the macroeconomic level is debatable, and has remained as a vexing issue. Researchers traditionally see these correlations as reflective of a causal link running from income to health, and consequently advocate the sanguine view that increases in national income will improve population health (see, e.g. McKeown, 1976; Pritchett and Summers, 1993; World Bank, 1993; Marmot, 2002; Deaton, 2002; Goldman, 2001). For instance, Marmot (2002) identifies two ways in which income can influence health: through a direct effect on the material conditions that have a positive impact on biological survival and health, and through an effect on social participation, the opportunity to control life circumstances, and the feeling of security. Above a certain threshold of material deprivation, income may be more important because of its link with these social and psychological factors, particularly in societies where social participation depends heavily on individual income (Marmot 2002, p.16). Acemoglu, Johnson and Robinson (2002) in a retrospective analysis of historical international epidemiological transition, remained doubtful to ascribe credit to improved health status as the first order determinant of economic growth, and discard the proposition that unfavourable health conditions are the root cause of the poverty of some nations. However, these views have been challenged by the likelihood that the income-health correlation is also explained by a causal link running the other way, i.e. from health to income.

An early and influential empirical work by Preston (1975) established strong cross-country correlations between measures of aggregate health, such as life expectancy or child mortality, and per capita income. Preston (1975) plotted life expectancy against income per capita for a cross section of countries for the years 1900, 1930 and 1960. The plots, subsequently known as “Preston Curves”, show an increasing and concave relationship - i.e. among the poorest countries small changes in income are associated with large increase in life expectancy and among the richer countries increase in income are associated with small, albeit positive, increase in life expectancy. The conventional view postulates that as GDP per capita increases the countries should move approximately along the contour of the Preston-curve. If, however, the Preston-curve shifts upward or rightward over time, then the conventional conjecture based on the cross sectional relationship becomes fallacious. An upward move of the curve means the countries being able to achieve increment in the life expectancy without increases in the GDP per capita. A rightward shift on the other hand would imply the countries not being able to increase their life expectancy despite increments in the GDP per capita. While the cross-sectional correlation is more or less stable over time, Preston (1975) found the curves shifting upward over time. Preston’s estimate attribute only between 10-25% of the overall improvement in the life expectancy between 1930s and 1960s to the national income
per head. About 75-90% of the growth in life expectancy for the world as a whole between the 1930s and the 1960s are credited to the factors exogenous to a country's current level of income (e.g. public health interventions, including diffusion of health sector innovations).

Similar assertion is found in WHO (2002) referring to a study by the World Bank (Wang et al., 1999), which presents an analysis of data from 1952–1992. The study compares the increase of per capita income and the infant mortality rates and finds that if the relationship between income and infant mortality rate had remained as it was in 1952, IMR would have dropped from 144 per thousand to 116 per thousand by 1992. In reality, however, it dropped much more sharply to 55 per thousand. Furthermore, their estimate shows that 40% of differential mortality improvements between countries may be accounted for by differences in their income growth rates, which leaves a significant proportion of health gains to be unaccounted for. The research suggests that 45% of the reduction of child mortality can be accounted for by the technological progress and knowledge diffusion, 38% is due to the educational achievements of female adults, and only 17% can be attributed to the effect of income (WHO, 2002, p.15).

The role of health as a decisive component of human capital and, hence, as a potential determinant of economic growth has recently been the subject of major research attention (Suhrcke et al, 2005, p.9). Grossman’s (1972) influential work makes the case for health as human capital, and introduces the idea of investing in human capital (health and education) to improve outcomes in both the market (work) and non-market (household) sectors. Grossman (1972) provides new insight on the demand for health within a household production function model, where health is considered as a durable capital good that produces a flow of services over time in terms of "healthy time", or "sickness-free time", which are inputs into work, consumption and leisure activities. People inherit an initial stock of health capital, which depreciates with age and can be increased with investment that include child care, nutrition, clothing, housing, medical services, and the use of one's own time. In his prize lecture Nobel Laureate Theodore W. Schultz (1979) emphasized the merits of investing in health postulating population health as the decisive factor of production. The central thrust of Schultz (1979) is that population quality and knowledge matter – “Longer life spans provide additional incentives to acquire more education, as investments in future earnings. Parents invest more in their children. More on-the-job training becomes worthwhile. The additional health capital and the other forms of human capital tend to increase the productivity of workers. Longer life spans result in more years participation in the labor force, and bring about a reduction in ‘sick time’. Better health and vitality of workers in turn lead to more productivity per man hour at work”. Nobel Laureate Robert W. Fogel, in his seminal works on Great Britain and France covering the period 1790-1980 (Fogel, 1993, 1994, 1997), elucidates how the changes in health, in the composition of diet, in clothing and shelter, and reductions in the incidence of infectious diseases can significantly affect the efficiency with which ingested energy is

---

1 Although the impact of human capital formation on economic productivity has been well demonstrated, earlier empirical work attributed much of the human capital development to educational attainment. See, e.g., Schultz (1961), Mincer (1958), Andren and Palmer (2001).
converted into work output. His analysis of the distribution of body stature (height and weight) across the population in Great Britain and France and their association with food supply and calorie intake suggests that the latter contributed to the secular decline in the morbidity and mortality, and enhancing long-term productivity. He estimated that improvements in health and nutrition contributed to about one-third of income growth in Britain between 1790 and 1980. Subsequently, a recent strand of literature provides powerful underpinnings of the argument that good health is an essential determinant of economic growth, and that, conversely, bad health is a significant brake on economic and social development.

Also, owing to the rapid mortality and fertility decline, the latter half of the twentieth century has witnessed changes in the size, density and age composition of population on a dramatic scale, in both developing and developed countries. As countries move through the demographic transition of falling mortality followed by falling fertility, they face first a period of increasing child dependency ratios as larger proportion of the population moves through the working ages, and eventually of increasing old-age-dependency ratio. The effect of fertility decline in the second intermediate stage, through which virtually all developing countries have passed and will be passing in the early twenty first centuries, is a one-off ‘demographic bonus’ or ‘window of opportunity’ – a period of almost 50 years during which an initially high ratio of the working age to the dependent population gradually declines (Bloom and Canning, 2001, 2003b; Bloom, Canning, and Sevilla, 2003; Bloom and Williamson, 1998). This window of opportunity generates more workers to produce more total output, provided that they are productively employed; accumulates greater wealth, given that savings occur and are productively invested; and makes available a larger amount of human capital, when appropriate investments are made in its formation (see for example, Birdsall and Sinding, 2001; Bongaarts, 2001; Kelley and Schmidt 2001; Williamson 2001; Lee, Mason and Miller 2001; Bloom and Canning, 2001). Therefore, the emergent view on the contribution of health to economic growth brings, along with various economic aspects, the demographic attributes in the analysis. As a result, improvement of health and longevity is increasingly being considered as the driver of economic growth, development and poverty reduction.2

This shift in views on the health-income relationship has critical policy implications. For instance, if causation is perceived to be running from income to health, the policy implication is to support the poor, fight poverty, redistribute income etc. Alternatively, if causation is perceived to be running from health to earnings, the policy implication is to prioritize investments in public health, insure income against poor health through disability insurance or sickness benefits, so that sick people do not get poor.

2 After all that has been said, the conventional view that wealthier implies healthier is undeniable and still the dominant. Apparently, a more plausible proposition entails that the two views (i.e. health promotes wealth and vice versa) are compatible, and that causality running in each direction can give rise to cumulative causality, a so-called virtuous circle in which health improvements promote economic growth, which in turn promotes health (Bloom and canning, 2001; Alsan et al. 2006)
3. METHODOLOGICAL APPROACHES TO TRACK THE CHANNELS OF INFLUENCE

Construction of a theoretical framework postulating the potential role of health in economic growth requires identifying the main channels and causal links between health and economic outcome. Researchers assert that health contributes to economic prosperity via their effects on labor supply and labour market participation; worker productivity (and hence on wages and earnings); investments in human capital (e.g. higher skills as a result of greater education and training), savings available for investment in physical and intellectual capital, fertility choices faced by women; and population age structure (see for instance Alsan et al., 2006; Bloom, Canning and Jamison, 2004; Bloom, Canning and Sevilla, 2003; Bloom and Canning, 2004; Bloom, Canning and Malaney, 1999; CMH, 2001; Easterlin, 1999; Hamoudi and Sachs 1999). The impact transmission mechanism differs between improvements in the adult health and that of the child health. For instance, healthier adults are more likely to be in the workforce and with greater productivity, to live longer and to save more for retirement; while healthier children are more likely to have better schooling outcome and cognition, to lead to lower child mortality and reduced fertility, and in the end generate a demographic dividend. Healthier nations are also likely to attract larger amounts of foreign direct investments, since investors avoid environments where labour force are likely to be weakened by heavy disease burden, and where access to health care is limited (Alsan et al., 2004, 2006). The next section highlights different methodological approaches that the researchers applied in analyzing and establishing various aspects of the transmission mechanism.

3.1 MICROECONOMIC APPROACH

One widely used methodological approach, based on individual or household data, provides evidence that health matters for a number of economic outcomes: wages and earnings, the amount of hours worked, labour force participation, early retirement, saving decision, and the labour supply of those giving care to ill household members (Suhrcke et al., 2005). This microeconomic approach to establish the health-economy link has mainly involved experimental and quasi-experimental studies, which intend to provide direct scientific evidence on the causal effect of health on the outcome studied; and observational studies, which are typically based on survey data and are estimated in conjunction with a model of behavior that seeks to provide a plausible argument for interpreting the evidence in a causal framework (Thomas and Frankenberg, 2002). Examples of microeconomic approaches in different context can be found, inter alia, in Strauss and Thomas (1998), Suhrcke et al. (2005), Thomas and Frankenberg (2002), Ruger et al. (2006), Behrman and Rosenzweig (2001), Miguel and Kremer (2004), and Schultz (2002).

Together with the hypothesis that work-output and other labour outcomes are influenced by both macro-nutrients (i.e. energy, or protein intake) and micro-nutrients (i.e. iron, iodine, zinc, calcium and several key vitamins), the experimental and quasi-experimental studies have mainly involved establishing the link between nutrition and productivity. Thomas and Frankenberg (2002) present a survey of studies that examine the causal effect of iron
deficiency on work capacity by comparing the output and productivity of the treatment group (recipient of iron nutrients) vis-à-vis the control group (non-recipient or lacking iron nutrients). The studies (both clinical and field studies) show that iron deficiency affects an individual’s aerobic capacity, endurance, energy efficiency, and work output. For example, Li et al. (1994) conducts a randomized treatment-control study with 80 iron-deficient non-pregnant Chinese female cotton mill workers and finds that the workers who received iron supplementation for 12 weeks had a 5% increase in gross and net energetic efficiencies relative to the controls who received a placebo. Furthermore, treatments had significantly reduced heart rates and a 17% increase in production efficiency. There was no increase in the work output, perhaps because output was constrained by the conveyor belt-based production technology and was dependent on the output of the co-workers. However, the gains in the energetic and production efficiency enabled the treated women to spend more time and expend more energy on non-work activities.

Similar kind of studies examined the effect of food or calorie supplementation on productivity, and comparing the results between the treatment and control groups. While explanations that are based on the underlying biological mechanisms that link micro-nutrients (e.g. iron) with cell functioning are powerful, the explanations of the underlying mechanisms, and the findings, are not convincingly clear in the studies that examines the effect of food or calorie supplementation on productivity (Thomas and Frankenberg, 2002). Although numerous studies (e.g. Strauss, 1986; Sahn and Alderman, 1988; Thomas and Strauss, 1997; Fuentes et al., 2001) confirmed the positive impact of calorie intake on productivity and earning, several studies (e.g. Deolalikar, 1998; Haddad and Bouis, 1991) highlighted the importance of unobserved heterogeneity and showed that, when controlled for other potential effects, the positive association disappears.

The difficulties involved in measuring nutrient intakes in the household surveys led the researchers to use anthropometric variables or physiological proxies for health such as height, body-mass-index (BMI - a composite indicator measured by weight/height\(^2\)), age of menarche, adult stature etc, which are perceived to be largely determined by the nutrition intakes during the childhood, and thereby capture the longer and shorter run dimensions of nutrition. For instance, albeit height may be partly linked to genotypes and family background, better height may reflect a cumulative measure of the absence of poor diets and infection in early childhood including the foetal period, and thereby is viewed as a longer-run indicator of nutritional status. In this line, height may be perceived as an indicator of human capital much along the lines of education (Thomas and Strauss, 1997). That a taller person is likely to command higher wage is demonstrated in several studies (e.g. Strauss and Thomas, 1998; Thomas and Stauss, 1997; Haddad and Bouis, 1991; Foster and Rosenzweig, 1993). Experiments also indicate that several domains of health care other than nutrition have a causal impact on economic prosperity. For example, a study in the United Kingdom (see Moffett et al., 1999) randomly assigned men with back pain to an exercise program (the treatment group) or to the usual primary care management (the control group). The treatment group, after a year, was found to have reported fewer days of absenteeism compared to the
controls. Other instruments in experimental studies include changes in the price of health care (Newhouse, 1993), changes in the prices of health services (Dow et al., 2001), differential health insurance schemes (Gruber and Hanratti, 1995) etc. All of the instruments used in the studies confirm positive impact of health on economic outcome.

Analyses of the bearings of poor health on the earnings capacity, primarily due to the withdrawal from the labour market, are found in numerous studies (see, e.g. Andren and Palmer, 2001; Hansen, 2000; Contoyannis and Rise, 2001; Gambin, 2004; Pelkowski and Berger, 2004; Gustman and Steimeir, 1986; Bartel and Taubman, 1979; Luft, 1975; and Chirikos and Nestel, 1985). The point of departure for such analyses is usually the models of human capital pioneered by the economists like Schultz (1961), Mincer (1958), Becker (1962), Ben Porath (1967), and Grossman (1972). In particular, Grossman’s (1972) model of the demand for health that articulated the intricate interrelation among work-time, wages, and health motivated many studies focusing on work, wage, and health.3

The question of whether sickness history affects annual earnings and hourly wages is addressed, for instance, by Andren and Palmer (2001). They estimated both annual earnings and hourly wage equations, and found that people who are healthy in the current year, but who have had long-term sickness in the previous five years have lower earnings in the following years than persons without long-term sickness, even if they did not experience a new spell of long-term sickness. The empirical part is based on a longitudinal survey conducted in Sweden during 1983-1991 for people in working age 16-64. Studies dealing with similar research questions and findings, include inter alia, Hansen (2000), Contoyannis and Rise (2001), Gambin (2004), Pelkowski and Berger (2004), Gustman and Steinmeir (1986) etc.4

Several micro-level studies have focused on the extent to which less than full health reduces supply of labour. Three main aspects of the potential labour supply effects of health are: labour force participation, early retirement, and the labour supply of caregivers. A study by Chirikos and Nestel (1985) finds strong evidence of the impact of health problems on labour supply by estimating the net effects of poor health on wages and annual hours of work. Analyzing the subjects grouped by gender, race, and health status, they found sickness history to be detrimental to current earnings.5 Albeit the general consensus that health increases the

3 For a comprehensive review of literatures examining the impact of health on productivity, earnings and wages, and labour supply, see Andren and Palmer (2001) and Surcke et al. (2005).


5 They distinguish the ‘direct effect’ of health on the annual hours of work (due to changes on the preferences between leisure and market work) from the ‘indirect effect’ through the impact of health on wages. Their analysis is based on data from the US National Longitudinal Surveys (NLSs) on older men in 1976 and mature women in 1977. Information on health over a 10-year period was combined to construct four health categories under which all individuals could be classified: ‘continuously healthy’; ‘continuous poor health’; enjoying ‘improving health’; and ‘deteriorating health.’ Estimates were made separately for women and men, and for black and white people, giving rise to four sex-race groups.
probability of participating in labour force, there is no consensus about the magnitude of this
effect; and the comparison of results from different studies is difficult, as they use different
measures of health, model forms and estimation techniques (Suhrcke et al., 2005, p.45).

With regards to the correlation between education and health, prior research revealed large
and positive association (see for example, Grossman, 2004; Lleras-Muney, 2005; Grossman
and Kaestner, 1997). Several points of view explain this correlation as: (a) the causal relation
runs from increases in schooling to betterment of health, because better education makes
people better decision makers with better information and conscience in terms of adopting
prudent use of medical care, healthy habits and caution in the choice of occupation;\(^6\) (b) the
direction of causality runs from better health to more schooling via better cognitive ability and
less absenteeism; and (c) the correlation does not imply direct causal relation, instead,
differences in one or more ‘third variables’, such as physical and mental ability, genetic
characteristics and parental background, affect both health and schooling in the same
direction. However, the causality has not been established in a definitive manner - there is
little agreement among economists concerning the mechanisms involved.

Turning to the potential impact that health has on education, it is argued that good health
enhances cognitive functions and reduces school absenteeism and early drop-outs (Bleakley,
2007). Hence, children with better health can be expected to reach higher educational
attainments and be therefore more productive in the future. Moreover, healthier individuals,
with a longer lifespan in front of them, would have greater incentives to invest in education
and training as they can harvest the associated benefits for a longer period (Kalemli-Ozcan et
al., 2000). Also, reduction of adult mortality reduces the number of orphans, who receive less
schooling than children with living parents (Case et al., 2004). The relationship between
health in childhood and cognitive development has been the subject of several studies in the
USA. Edwards and Grossman (1980) find a significant correlation between two indicators of
cognitive development (an IQ measure, and a measure of school achievement) and several
health indicators. The authors use cross-sectional data from the Cycle II of the US Health
Examination Survey (HES) which covers children aged 6 to 11 years over the period 1963–
65. They estimate the effect of 13 health-related dimensions (four related to past health status
and nine related to current health status) on two measures of the children’s cognitive
development. The impact of a series of non-health variables is controlled for. Their results
show a significant correlation between the two measures of cognitive development used and
the following health-related variables: low birth weight (negative), being breast-fed (positive),
having a ‘significant abnormality’ (negative), height (positive effect), and the number of
decayed teeth (negative). The measure of school achievement was also found to be
significantly and negatively correlated with having hearing problems and with health being
assessed as poor or fair by the child’s parents. Finally, the IQ measure was lower for children
whose mother was younger than 20 years at birth and higher for those whose mother was
older than 35 years compared with those whose mother was between 20 and 35 years old.

\(^6\) One line of thought emphasises the role of education over (instead) income in the process of health
determination (Deaton, 2002).
Smith’s (1999, 2003, and 2005) study examines the dynamic interplay between health and economic prosperity. The objective here is to pin down some of the mechanisms through which the two are associated. He attempts to unravel the causality issue with the gradients that exist in the relationship between health and the socio-economic status (SES). He analyses the nature of the SES-Health gradient by age-specific household income quartile, and reveals that: (i) until the end of life, at each age every downward movement in income is associated with increased incidence of poor health. Moreover, these health differences by income class are large; (ii) there exists a strong nonlinearity in the relation between income and health, with the largest health differences taking place between the lowest income quartile and all the others, (iii) there is a distinct age pattern to the SES health gradient, with health disparities by income class expanding up to around 50 years, after which health gradient slowly fades away. Subsequently, Smith (1999, 2003, and 2005) addresses the following issues: (a) to what extent does health matter for SES outcomes? (b) do the households’ SES attributes predict the onset of new health events? If the direction of causation is that SES mainly affects health, what dimensions of SES actually matters – financial aspects such as income or wealth; or the non-financial dimension like education? (c) is there a life course component to the health gradient? And (d) Can the non-linear characteristics of the SES-health gradient over ages during which the gradient is withering away be generalised to the whole life-course, especially to those ages during which it is emerging?

Smith (1999, 2003, and 2005) uses the HRS and AHEAD panels to examine the consequences of new health on a series of SES related outcomes: out-of-pocket labour supply, labour force activity, household income and wealth. He also explores the pathway from SES to health examining whether the onset of new chronic conditions is related to levels of household income, wealth, and education. Since HRS and AHEAD samples capture the SES-health gradients for the cohorts aged 51-61 and above 70 respectively, the life course attributes are explored using the Panel Study of Income Dynamics (PSID), which spans all age groups, allowing for examining behaviour over the complete life cycle. He finds that the onset of unanticipated new health shocks (not induced by SES) lead to significant downward-cascading ripple like impact on four aspects of SES. Also, for each of the SES outcomes, new severe health events have a significant effect, although most of the impact on income and wealth arises through changes in labour supply and not necessarily because of medical expenses. On the impact of SES on Health, education, not individual’s financial resources, is

---

7 HRS is a national sample of about 7,600 American households (12,654 individuals) with at least one person in the household aged 51–61 years originally interviewed in the fall of 1992 and winter of 1993.

8 A US sample of respondents who were at least age 70 at baseline.

9 All the models include a vector of baseline attributes, including baseline measures of birth cohort (or age), marital status, race, ethnicity, education, region of residence, quintiles of family income, and an extensive vector of measures of baseline health. These health measures include dummies for the self reported health status, the presence at baseline of each chronic condition, a set of behavioural risk factors (smoking, exercise, BMI, drinking), and a scaled index of functional limitations based on the answers to the questions about activities of daily living.

10 PSID gathered almost 30 years of extensive economic and demographic data on an nationally representative sample of approximately 5000 (original) families and 35000 members of those families.
the primary factor in predicting the health outcomes. However, Smith (2005) appreciates the growing evidence, including some reported in his paper, that measures of economic circumstances during childhood have a bearing on health outcomes later in life. Parental income appears to be central correlates of the onset of some critical childhood diseases, which then set the stage for the adult SES health gradient.

Smith (1999) also analyses the dual relation between health and economic status in terms of the impact that health bears on family saving decision. It is highly plausible that savings (or savings rate) increases with the prospect of a longer and healthier life. The idea of planning and, hence, saving for retirement would be expected to occur only when mortality rates become low enough for retirement to be a realistic prospect. A recent study by Kinugasa and Mason (2007) analyze the effects of the transition in adult longevity on the national saving rate using historical data and international panel data, and an overlapping generation model; and finds that the rise in adult life expectancy has a large and statistically significant effect on aggregate saving. They find the effects to be especially pronounced in East Asia because its mortality transition was very rapid. Similarly Lee, Mason, and Miller (2000) tracked the life cycle saving pattern in the context of the demographic transition in Taiwan, and obtained similar conclusions.

While the micro estimates establish that improved health leads to better individual economic outcomes, it would be incorrect simply to scale up the individual level estimates to derive the effects at the macro level. Although the micro-estimates reveal that individual health differentials lead to differences in earning (e.g. due to the healthier being successful in competing against less healthy individuals in the labor market)\footnote{For example, Persico, Postlewaite and Silverman (2004) suggest the role of “competitive advantage” in adolescence in explaining the effect of height on economic outcomes.}, such effects would be much smaller were all individuals to become healthier. On the other hand, differences in the income-health relationship at the micro and macro level may also be due to the positive externalities of health that are not taken into account at the micro level. Acemoglu and Johnson (2006, 2007) assert that micro estimates do not directly answer the question of how important differences in disease environments and health conditions are in accounting for cross-country income disparities, because they do not incorporate general equilibrium effects. In particular, they mention about the diminishing returns to effective units of labour when land and/or physical capital are supplied inelastically. “In the presence of such diminishing returns, micro estimates will exaggerate the aggregate productivity benefits from improved health, especially when health improvements are accompanied with population increases” (Acemoglu and Johnson, 2006, p.4).
3.2 MACROECONOMIC APPROACHES

3.2.1 Production Function Approach

Corollary to the micro level evidences of the contribution of health on economic outcomes, there are attempts to corroborate the micro evidences by the macro evidences of the effect of population health on economic growth. This macro-approach has been basically an adaptation of the firm-level production function at the national level that seeks explanations for variations in per capita income. The attempts in this regard entail analyzing the components of the growth process of the economy and identifying the appropriate entry points for the ‘health’ component in the growth analysis. The two main approaches are the augmented Solow (and) neo-classical approach and the ‘new growth theories’, with their respective empirical counterparts, as discussed in the next section, of growth accounting exercises and macro regressions. Models of economic growth therefore have been extended to include the importance of health as a human capital input.

The inclusion of human capital in the neo-classical, or exogenous, growth theories, mainly put forward by Solow (1956) and Swan (1956), enables us to embed all three of the main proximate sources of income differences: Physical capital, human capital, and technology. The human capital (e.g. health, education, skills, training etc) has been incorporated in an aggregate production function of the Solow-model set-up in two different ways: as a separate factor of production (see, e.g Mankiw et al., 1992; Knight, Loyaza, and Villanueva, 1993; Knowels and Owen, 1995) or as a determinant of technological progress (e.g. Benhabib and Spiegel, 1994; Knowels and Owen, 1997). However, the incorporation of health component into the growth models transcends the traditional Solow type (and its extended versions) growth models. The positive effect of health on economic growth is identified either in exogenous growth model during the transition to the steady state or in endogenous growth models, each within the context of intertemporal optimization.\(^{12}\)

The standard neoclassical growth model that introduces household optimization and endogenizes savings has contributed by paving the way for further analysis of capital

\(^{12}\) The two archetypes of growth models that have been more frequently considered in empirical growth applications are the Solow model and the AK model, with the distinction between the two depending essentially on the hypothesis made on the production function (decreasing or constant returns to variable factors, respectively). The A-K model refers to one of the simplest endogenous growth model, which is derived from the simplification of the original Solow (1956) model: \(Y = F(K,L) = K^\alpha L^{1-\alpha}\). Modifying the production function so that \(\alpha=1\), one gets \(Y=AK\), where \(A\) is some positive constant. It is the production function that gives the ‘A-K’ model its name. The assumption of \(\alpha=1\) refers to the case, in which the transition dynamics extends infinitely. In this way, The A-K model generates growth endogenously. (Jones, 2002, pp.157-160). The AK model implies a permanent effect of the investment rate on long run growth, and growth is therefore endogenous, while in the Solow model the impact of investment rate is only transitory, the only driver of growth being therefore exogenous technological change. When a distinction between physical and human capital is made, the natural extensions of these models are, respectively, the augmented Solow model (Mankiw et al., 1992) and the two-sector AK (or Uzawa-Lucas) model (Uzawa, 1965; Lucas, 1988). Consistently, in the latter the process of accumulation of human capital is the main driver of growth, while in the former it has only a transitory impact on growth, permanently affecting only the level of steady state.
accumulation, human capital investment, and endogenous technological progress. Also, by relaxing the representative household assumption, there has been a major departure from the baseline neo-classical growth model with the advent of the overlapping generation (OLG) models. The study of interactions between economic development and longevity has been enriched by OLG models with endogenous longevity (see, for example, Chakraborty, 2004; Bhattacharya and Qiao, 2005; Wang and Leung, 2003). In these models the health care and its interactions with other forms of investment choices has been modelled.

Endogenous growth theory departs from neo-classical growth theory in emphasizing that technological progress itself is an economic process, with economic determinants, much like the process of capital accumulation. Therefore, the endogenous growth theories suggest that when human capital is taken into account in the process of capital accumulation, there is no reason to think that diminishing returns will drag its marginal product down to zero (as postulated by neo-classical growth theories) because part of that accumulation is the very technological progress needed to counteract diminishing returns. The endogenous growth theories paved the way for researchers to incorporate primarily education, and subsequently, health as important determinants of technical progress (see Lucas, 1988; Barro, 1996; Barro and Sala-i-Martin, 1995; Rebelo, 1991; Mankew et al., 1992;).

Bloom, Canning and Sevilla (2004) estimate a production function model of aggregate economic growth including two variables: work experience and health. Their results show that health has a positive and statistically significant effect on economic growth. It suggests that a one-year improvement in a population’s life expectancy contributes to an increase of 4% in output, and that the life expectancy effect in growth regressions appears to be a real labor productivity effect, and is not the result of life expectancy acting as a proxy for worker experience.

Howitt (2005) sets up a model of innovation-based Schumpeterian growth theory, an extension of the $A$-$K$ endogenous growth model, which contains different channels through which an improvement in a population’s health affects a country’s long-run growth performance. He analyses six different channels through which an improvement in a country’s population health will impact its long-run growth performance: (i) health induced productive efficiency; (ii) life expectancy affecting skill-adjusted death rate; (iii) creativity and research intensity; (iv) coping skill with regards to R & D; (v) learning capacity; and (f) inequality affecting the school attendance rate. The model shows that increase in life expectancy have a direct effect on the steady-state average skill level of the population. However, the sign of the effect depends on the demographic incidence; i.e. whether the life expectancy prolongs the life-span of the productive workers, or working primarily through a reduction in infant mortality.

Zon and Muysken (2005) proposes an endogenous growth model where the provision of health services influence an economy’s rate of growth. Since growth is produced using labour services that have alternative uses, increasing health activities may also imply lowering output and possibly growth. Hence there is a clear trade-off between the health state of the
population and growth performance. The authors combine a stylized demographic and epidemiological model with Lucas (see, Lucas, 1988) model that distinguishes between care activities and cure activities. Their model illustrate how changes in rate of morbidity and the rate of mortality influence the optimum allocation of scarce labour resources over its various uses that include care, cure, final output production, and human capital accumulation activities (Lopez-Casasnovas et al., 2005, p.9). Their model underlines the negative consequences of the exclusive focus on cutting current health costs (rather than focusing on intertemporal effects of health activities) on the long-run growth performance.

3.2.2 Economic Growth Regression Approach

The common empirical approach toward examining the impact of health on economic growth has been to focus on data for a cross-section of countries and to regress the rate of growth of income per capita on the initial level of health (typically measured by life expectancy, or survival rate), with controls for the initial level of income and for other factors believed to influence steady-state income levels. These factors might include, for example, policy variables such as openness to trade; measures of institutional quality, educational attainment, and rate of population growth; and geographic characteristics. The application of growth regression approach can be seen, inter alia, in Barro (1996); Barro and Lee (1994); Barro and Sala-I-Martin (1995); Bhargava et al. (2001); Bloom, Canning, and Malaney (2000); Bloom and Malaney (1998); Bloom and Sachs (1998); Bloom and Williamson (1998); Caselli, Esquivel, and Lefort (1996); Gallup and Sachs (2000); Hamoudi and Sachs (1999). Almost all studies have found evidence of a positive, significant, and sizable influence of life expectancy (or some related health indicator) on the subsequent pace of economic growth. The estimated effects of an increase in the life expectancy (expressed in log terms) economic growth are the followings: 4.2% in Barro, 1996; 7.3% in Barro and Lee (1994); 5.8% in Barro and Sala-I-Martin (1995); 6.3% Bloom, Canning, and Malaney (2000); 2.7% in Bloom and Malaney (1998); 3.7% in Bloom and Sachs (1998); 4% in Bloom and Williamson (1998); 0.1% in Caselli, Esquivel, and Lefort (1996); 3% in Gallup and Sachs (2000); and 7.2% in Hamoudi and Sachs (1999). Although these studies differ substantially in terms of country samples, time frames, control variables, functional forms, data definitions and configurations, and estimation techniques, the parameter estimates of the effects of life expectancy and age structure on economic growth have been reasonably comparable across studies.

Growth regressions, however, are prone to the problem of multicollinearity leading to a large variation in the significance and size of the coefficient estimates. Also, the critics (see for example Acemoglu and Johnson, 2006) of cross-country regression studies see these as establishing merely a strong correlation between measures of health and both the level of economic development and recent economic growth without being able to establish a causal effect of health and disease environments on economic growth. The critics’ scepticism with the results of these studies lies in their concern that such macro studies may be capturing the effects of many omitted variables (Acemoglu and Johnson, 2006; Bloom, Canning and Sevilla, 2004).
Acemoglu and Johnson (2006, 2007) provide an empirical analysis based on the international epidemiological transition, apparently led by the wave of international health innovations and improvements that began in the 1940s, and finds that there is no evidence that the large exogenous increase in life expectancy led to a significant increase in per capita economic growth. They construct an instrument, referred to as predicted mortality, which is based on the pre-intervention distribution of mortality from 15 major diseases around the world and dates of global interventions (e.g. innovation and widespread use of dichloro-diphenyl-trichloroethane to curb malaria, and other disease related vaccines). This instrument appears to have a large and robust effect on changes in life expectancy starting in 1940, but no effect on changes in life expectancy before the interventions. Then, they show that instrumented changes in life expectancy have a large effect on population; a 1% increase in life expectancy leads to an increase in population of about 1.5-2%, but a much smaller effect on total GDP both initially and over a 40-year horizon. Accordingly, the impact of increase in life expectancy on per capita income is insignificant or negative. The interpretation of the insignificant or negative impact of life expectancy on GDP per capita is that, in the face of population growth, the other factors in the production function, capital and land in particular, did not adjust. As population increased, the capital-labour ratio decreased, which eventually led to a decrease in the per capita income. However, the results of many micro-level studies (see e.g. Strauss and Thomas, 1998; Suhreke et al., 2005; Thomas and Frankenberg, 2002; Ruger et al., 2006; Behrman and Rosenzweig, 2001; Miguel and Kremer, 2004; Schultz, 2002) and cohort-level studies (e.g. Bleakley, 2006a; 2007) obtained large positive effects of health on productivity. In that case, these latter effects, assuming that they exist and of partial equilibrium nature, were fully offset by the crowding-out effect of population growth (Bleakley, 2006). While it is customary to consider physical capital to be fixed in the medium run, Bleakley (2006) raises question on the assumption of inelastic land and capital, particularly in the long term.

Bleakley (2006) also remains critical about the instrument used by Acemoglu and Johnson (2006), because (a) it disproportionately affects early-life mortality and does not capture, apparently, the adult and/or late life Mortality; (b) the instrument is weighted towards mortality without taking into account the morbidity issue. The diseases under consideration in the Acemoglu and Johnson (2006, 2007) are lethal diseases with high case-fatality rates. Some diseases tend to have low case-fatality rates but lead to the significant reduction of morbidity (e.g. vivax Malaria, hook-worm etc.) for the population cohorts affected. In a study on Columbia, which was affected by both the vivax malaria (a parasite associated with high

---

13 In another instance, a possible explanation for the negative impact of health capital formation on GDP growth in the process of economic development is offered by the model of ‘unbalanced growth’ by Baumol (1967). This model postulate that imbalances in productivity growth between a ‘progressive’ (e.g. manufacturing) sector and a ‘nonprogressive’ sector (e.g. health care sector) of the economy lead to perpetual expenditure shifts into the latter, and thus causing an adverse structural change effect and a decline in overall GDP growth. Papers by Nordhaus (2008) and Hartwig (2008) support this hypothesis. However, Pugno (2006) offers a theoretical explanation with an endogenous growth model by hypothesizing that the growth-stimulating effect of health capital formation is likely to overcompensate the adverse sectoral shift effect.
morbidity) and the falciparum malaria (a parasite associated with high mortality), Bleakley (2006a) found that eradication of vivax malaria produced substantial gains in human capital and income. On the other hand, estimates indicate no such gains from eradicating falciparum. Bleakley (2007) studies several eradication campaigns against the hookworm parasitic disease in the Americas and finds that cohorts exposed to these campaigns as children had higher productivity and human capital. As these cohorts entered the workforce the average incomes rose above and beyond that due to the changing cohort composition of workers, which suggests the presence of positive spillovers from health capital.

Again, the estimates of the global growth regressions that take rich and poor countries together may differ significantly from the ones conducted by taking the rich or poor countries separately. Examination of the role of health for regions of different levels requires the use of health indicators that are better able to discriminate between levels of health among these countries. For example, it may be expected that the life-expectancy variable can hardly have much explanatory power for the analysis of the rich countries, because life expectancy and adult mortality vary more widely among low- and middle-income countries than in high-income countries. For instance, Weil (2007) suggests that health’s positive effect on GDP is strongest among poor countries. For rich countries, the existing empirical evidence on whether health capital formation stimulates GDP growth is mixed. For example, while Rivera and Currais (1999a, 1999b, 2003) find that the growth of health expenditure imparts positive effect on productivity growth for OECD countries, findings in Hartwig (2008) do not suggest that health capital formation fosters economic growth in rich countries. The studies by MacDonald and Roberts (2002) as well as Knowles and Owen (1995, 1997) did not find ‘life expectancy’ to be a statistically significant explanatory variable for productivity growth in high-income countries. Also, the results obtained by Bhagrava et al. (2001) indicate a negative impact of the adult survival rate on economic growth for the US, France, and Switzerland.\(^1\)

### 3.2.3 Cost of Illness (COI), Disease Burden Approaches

The output of COI studies is an estimate, expressed in monetary terms, of the total burden of a particular disease from either a societal or (if a narrower set of costs is included) sector-specific perspective (Suhrche et al., 2005). COI studies separate the costs of illness into three components: (i) Direct costs referring to costs falling on the health sector in relation to prevention, diagnosis and treatment of disease; (ii) Indirect costs typically measuring the lost productivity potential of patients who are too ill to work or who die prematurely (i.e. the ‘human capital approach’). The measurement of indirect costs is a matter of much debate. Some COI studies consider the loss of future earnings, discounted to take account of the fact that the income will arise in the future. Others use the willingness to-pay method, in which individuals are asked to choose between different scenarios with the objective of assessing

\(^{14}\) A negative, but not significant, coefficient of health investment for income per capita growth is found by Easterly and Rebelo (1993).
how much they are willing to pay to be in a particular state of health; and (iii) Intangible costs capturing the psychological dimensions of the illness to the individual (and their family), i.e. the pain, bereavement, anxiety and suffering. This is the cost category that is typically hardest to measure. Suhrche et al. (2005) reports the estimates of the economic burden of numerous diseases in the developed countries; e.g. cardiovascular disease, mental illness, obesity and tobacco, diabetes, and alcohol.

While the economic dimension of ill health by using estimates based on the COI are important to make policy-makers aware of the seriousness of a given illness, these studies are limited by certain methodological challenges and by their failure to determine the direction of causality in the relationship between health and economic outcomes. The costs measured in cost-of-illness studies do not reflect the full economic costs associated with a given disease or risk factor; nor does the total or average size of the cost of illness necessarily imply a macroeconomic impact in terms of reduced economic development. The marginal contribution of health to production and to the incentives to invest in human capital is of more relevance for such deductions (Suhrche et al., 2005, p.38). The difficulties involved in comparing the existing studies across countries and disease or risk factors and the assumptions involved in estimating the indirect costs makes it difficult to propose generalised version of a COI. Also, the widespread use of the present value of future labour earnings to determine the foregone economic value caused by mortality or morbidity implicitly assumes that people who are not part of the workforce are ‘unproductive’ and therefore have no (economic) value. This problem has, however, been addressed by methods that seek to measure the value of life, and in some COI studies by assigning the minimum wage to individuals outside the labour force.15

In a broader application of COI, Bloom and Malaney (1998) examine the macroeconomic consequences of the Russian mortality crisis during the first half of 1990s. They carry out economic assessment of the Russian mortality crisis by estimating first both the size and the structure of the demographic shock. Then, they evaluate the costs of the mortality crisis in Russia using the human capital method of estimating the cost of illness (COI), which accounts for both the direct costs (e.g. personal medical care costs such as expenditures on diagnosis, treatment, and care, and non-personal medical care costs such as government expenditures on education and basic research) and the indirect costs (e.g. estimate of the value of the productivity forgone because of the morbidity and mortality associated with the illness and some measure of pain and suffering). They found that the capitalized value of the income forgone due to the crisis represents a value between 1.8 and 2.7% of Russia’s 1990 GDP. Considering the pitfalls in the COI methodology, Bloom and Malaney (1998) provide an alternative estimate of the Russian mortality crisis by using an empirical neo-classical growth model incorporating population age structures into it. Due to data limitations, they did not estimate the parameters of a Russian specific growth model; instead they use the estimates from a cross-sectional analysis of 78 other countries to evaluate the average contribution of

changing life expectancy and age structures to economic growth. Subsequently, they use the resulting estimates to evaluate the impact of Russia’s mortality crisis on economic growth. Their calculations, based on the growth model, suggest that the net effect of the drop in life expectancy, the decline in the rate of total population growth, and the even larger decline in the rate of growth of the working age population have been to lower the annual rate of growth of income per capita in Russia by roughly one-third of one percentage point.

3.2.4 Analysis of Historic Health and Wealth Evolutions

The global transformation in human health over the last two centuries has encouraged researchers to focus increasingly on the links between population health and economic growth. Consequently, researchers analyzed the historical decline in mortality, mortality differences between rich and poor countries, differences in mortality within countries, and juxtaposed these evolutions with the national output and wealth trends.

Fogel in his Nobel Prize lecture in 1993 (see Persson, 1997) emphasizes that useful insights can be achieved by analyzing the long-run dynamic processes through a study of history. Fogel (1993, 1994, 1997) presents an impressive historical account of the joint evolution of the population body-size (e.g. height and weight) and food supply, and shows that the latter to be critical for long-term labour productivity. His empirical investigation on Great Britain and France suggests that the increase in the amount of calories available for work over the past 200 years (i.e. 1780-1980) have made important contribution to the secular decline in the morbidity and mortality, and subsequently to the growth rate of the per capita income. Increased calorie intake not only increased the size of the productive labour force, but also allowed those in the workforce to be more productive. The estimates show that incorporation of the bottom fifth of the population (who had the lowest calorie intake and highest morbidity profile) into the workforce through the provision of adequate calories would have contributed 0.11% to the UK annual growth rate between 1780 and 1980. Additionally, the productivity gains due to sufficient calorie intake contributed 0.23% to the annual growth rate. British per capita income during this period grew at an annual rate of about 1.15% - suggesting that the combined effects of the increase in dietary energy available for work, and of increased human efficiency in transforming dietary energy into work output contributed about 30% of the British per capita growth.

Cutler et al. (2006) provides a detailed synthesis on the determinants of mortality reductions with a retrospective of cross-country historic health and wealth evolutions since the early eighteenth century. The study identifies three phases of mortality reductions. The first phase, from the middle of the eighteenth century to the middle of the nineteenth century, is the one where improved nutrition and economic growth is perceived to have played the major role in health determination, along with incipient public health measures. In the closing decades of the nineteenth century and into the first decades of the twentieth, the second phase occurred, in which public health mattered more - first negatively, because of high mortality in cities, then positively thanks to the delivery of clean water, removal of wastes, and advice about
personal health practices. The third phase, from the 1930s onwards, has been the era of large scale medical intervention, starting with vaccination and antibiotics, and moving on to the expensive and intensive personal interventions that characterize the medical system today (Cutler et al., 2006).

Arora (2001) investigates the influence of health (using five measures of long-term health) on the growth paths of ten industrialized countries over the course of 100 to 125 years, and finds that changes in health increased their pace of growth by 30 to 40 percent, altering permanently the slope of their growth paths. However, the stipulated impact of health on economic growth has remained contentious. For example, Acemoglu, Johnson and Robinson (2003) argue that health differences are not large enough to account for much of the cross-country differences in incomes, and that the variation in political, economic, and social institutions is a more central factor.

3.2.5 Changes in Demography and Its Economic Impacts

Changes in the demography is the result of change in both of fertility and mortality decline, and allowing for the migration patterns, the aggregate change in population growth brings over time changes in the size of population, its density, and its age composition. Besides the impact of health on several behavioral aspects with regards to economic variables, including savings patterns, labour force participation, investment and skill formation, incorporation of these demographic aspects has enriched the analysis that tracks the health-wealth relationship.

In recent years opinion on the effect of demography on economic performance has swung between the Malthusian views of Coale and Hoover (1958) that raise concerns on the adverse effect of high birth rates on economic growth, and the cornucopian views of Julian Simon (1981) that refers to the revisionist approach and asserts that over the long run the different components of demographic change can have offsetting, and thus moderating effects (Birdsall and Sinding, 2001; Kelley, 2001). Foley (2000) suggests a more cheerful view of demography, i.e. the existence of a stable demographic equilibrium with a high standard of living and relatively low total population. Foley’s (2000) theory of demographic equilibrium is based on Smith’s (1776) assumption of increasing returns to population due to a widening division of labour. The theory entails the existence of a demographic equilibrium at which population is stabilized because fertility falls with rising household income and household income rises with increasing population due to an increase in the social division of labour. Therefore, in contrast to the Malthusian equilibrium of a low standard of living at which high mortality balances high fertility, Foley (2000) finds an equilibrium point with a high standard of living at which low fertility balances low mortality (Foley, 2000, p.310).

Due to medical, nutritional and lifestyle changes, the demographic situation of many countries has improved: mortality has declined, life expectancy has increased, fertility has declined\textsuperscript{16},

\textsuperscript{16} As couples realize their children are more likely to survive, they need, and eventually have, fewer of them to attain their desired family size. In addition, desired fertility tends to decline as earnings opportunities improve.
and morbidity has diminished. As a consequence, the latter half of the twentieth century has witnessed changes in the age composition of population on a dramatic scale, in both developing and developed countries. Birdsall and Sinding (2001) states - “…though the changes have come slowly in terms of the short scholarly life of individual researchers, they have been stunningly rapid in historical terms, and highly differentiated across countries – making it possible to assess the impact of those changes over time and across countries.” Subsequently, the concept of ‘dependency burden’ paved the way for the researchers to investigate and juxtapose the patterns of demographic transitions in different regions (e.g. East Asia, Sub-Saharan Africa, Ireland) with that of economic performances (see among others, Bongaarts, 2001; Bloom and Canning, 2003c; Birdsall and Sinding, 2001; Kelley and Schmidt 2001; Williamson, 2001; Lee, Mason and Miller, 2001; Mason, 2005; Jimenez and Mamta, 2006). Bongaarts (2001) traces the shift in dependency ratios that accompanies the demographic transition across the major regions of the developing world, showing that the shift occurred earliest in East Asia, followed shortly thereafter by Latin America, and considerably later by Africa. The Middle East and South Asia are at intermediate points between Latin America and Africa.

Several studies show that East Asia's remarkable economic growth in the past half century coincided closely with demographic change in the region (Bloom and Canning, 2006; Bloom and Williamson, 1998; Bloom, Canning and Malaney, 2000). The sharp fall of infant mortality between 1950 and 2000 (i.e. from 181 to 34 per 1,000 births) was followed by fertility reduction from six to two children per woman. The lag between falls in mortality and fertility created a baby-boom generation. Between 1965 and 1990, the region's working-age population grew nearly four times faster than the dependent population. This is deemed to be a welcome demographic dividend, which, according to the estimates of several studies, was responsible for one-third of East Asia's economic growth during the period.

Similarly, in the decade following the legalization of contraceptives in 1979, Ireland saw a sharp fall in the crude birth rate, leading to decreased youth dependency and a rise in the working-age share of the total population. By the mid-1990s, the dependency burden in Ireland had dropped to a level below that in the United Kingdom. Economists thus trace the role of demographic change in Ireland and relate it to its impressive growth performance (see Bloom and Canning, 2003c). The decline in youth cohort sizes and increased female participation in the labour market are perceived to be the outcome of fertility reduction, and contribute to economic prosperity. The decline in youth cohort sizes and rapid economic growth of the 1990s also led to a reversal of the out-migration of young workers, resulting in net in-migration of workers; made up partly of return migrants and also, for the first time, substantial numbers of foreign immigrants.

---

since forgone income is such a large portion of the cost of childrearing (see, e.g. Bloom and Canning, 2003a, 2003b).
3.2.6 Use of Value of Statistical Life and the Integrated Economic-Demographic Model

How big an overall contribution does better health make to economic growth? In answering this question, economists face difficulties to use the traditional national income and product accounts. There is a widespread recognition that GDP is an imperfect measure of social welfare because it fails to incorporate the value of health. While the national account include the values of inputs into health care, it does not take into account the welfare implication of health improvements manifested in increased life expectancy, reduced mortality, or increased quality of life. Since health is an important component of properly defined social welfare, measuring the economic cost of ill health only in terms of foregone GDP leaves out a potentially major part of its ‘full income’ impact, defined as its impact on social welfare. Judging countries’ economic performance by GDP per capita, therefore, fails to differentiate between situations where health differs.

What is the implicit value that people attribute to health? While this value is probably high, it is not infinite, since in the day-to-day context we are not willing to give up everything in exchange for better health. In some cases we are even prepared to pay to use harmful substances. The challenge is to make the high value attributed to health more explicitly visible by measuring the extent to which we are willing to trade-off health with specific market goods for which a price exists. This is undertaken in willingness-to-pay (WTP) studies (Surche et al., 2005, p.70). Given that individual’s day-to-day life economic decisions reflect how they value health and mortality risks, economists refer to individual willingness to forgo income to work in safer environments and social willingness to pay for health-enhancing safety and environmental regulation as approximate measure of the value of differences in mortality rates. Using evidence on market choices that involve implicit trade-offs between risk and money, economists have developed estimates of the value of a statistical life (VSL). Based on this VSL comes the concept of full income which is being used instead of traditional GDP measure in valuing the welfare implications of health improvements.

Earlier attempt to introduce the value of mortality reductions into national income accounting can be found in Usher (1973, 1980). He did this by generating estimates of the growth in ‘full income’ for six countries and territories (Canada, Chile, France, Japan, Sri Lanka, and Taiwan) during the middle decades of the 20th century. For the higher income countries in this group, about 30% of the growth of full income resulted from declines in mortality. Estimates of changes in full income are typically generated by adding the value of changes in annual mortality rates (calculated using VSL figures) to changes in annual GDP per capita.

---

17 Bloom, Canning and Jamison (2004) presents a simple illustration of the VSL calculation: If a worker requires (and is paid), ceteris paribus, 500 Euro a year of additional pay to accept a riskier but otherwise similar job, where the increase in the mortality rate is 1 in 10,000 a year, the value placed on reducing the risk is simply 500 Euro. The VSL is defined as the observed amount required to accept a risk divided by the level of risk – in this example the VSL would be 500/(1/10,000) = 5,000,000 Euro. Willingness to pay to avoid risk rises with income. A reasonable range of values for a country’s VSL appears to be 100-200 times GDP per capita, with values estimated in richer countries more like to occur toward the high end of the range (Bloom, Canning and Jamison, 2004).
Even these estimates of full income are conservative in that they incorporate only the value of changes in mortality and do not include the total value of changes in the health status (Surche et al., 2005, p.70). For the USA, Nordhaus (2003) found that the economic value of increases in longevity in the last 100 years is about as large as the value of measured growth in non-health goods and services. In a rediscovery of Usher’s pioneering work, Nordhaus (2003) tested the hypothesis that improvements in health status have made a major contribution to economic wealth (defined as full income) over the 20th century. A more detailed assessment reveals that ‘health income’ probably contributed to changes in full income somewhat more than non-health goods and services in the first half of the 20th century and marginally less than non-health goods and services since 1950. As emphasized by Suhrcke et al. (2005, p.71), if the results of this and other related studies (e.g. Costa and Kahn, 2003; Crafts, 2003; Cutler and Richardson, 1997; Miller 2000; Viscusi and Aldy, 2003) are confirmed, then the role of health (and of the activities that promote health, among others the healthcare system) should be reconsidered. It raises the possibility that the social productivity of spending on health (via the health system and via other sectors) might be many times greater than that of other forms of investment.

Bloom, Canning and Jamison (2004) presents an application of VSL in calculating the devastating economic impact of AIDS in sub-Saharan Africa, and shows that the intrinsic value of this health deterioration is much more than otherwise derived using GDP as a measure. Their estimate consists of two components: the change in GDP per capita and the value of changes in mortality rates as estimated in the VSL literature. The first step in obtaining the VSL estimate entails calculation of the impact of AIDS on mortality rates. By 2000, the epidemic had, on average, progressed to the point that mortality rates (in middle-ages) were beginning to increase substantially. In 1990, a 15-year-old male had a 51 percent chance of dying before his 60th birthday, and this had increased to 57 percent by 2000. For females, the increase was from 45 to 53 percent. Taking the average of the change in annual mortality probabilities gives 0.35 percent a year from 1990 to 2000. The next step entails calculation of the economic cost of these mortality increases. They use a conservative estimate of 100 times GDP per capita as the VSL, and find that Africa’s mortality changes imply an economic cost of the epidemic approximately equal to 15 percent of Africa’s GDP in 2000 (assuming that about 50 percent of the population is aged 15–60 and that 90 percent of AIDS deaths are in this age group). This corresponds to a decline in income of 1.7 percent a year from 1990 to 2000, far higher than existing estimates of the effect of AIDS on GDP. (Bloom, Canning and Jamison, 2004, p.14)\textsuperscript{18}

Murphy and Topel (2006) present an inclusive application of the VSL concept in evaluating the economic impact of health improvements in the USA during the last century. In their

---

\textsuperscript{18} Rosen (1988) came up with similar analysis where valuation formulas for age-specific mortality risks are derived from life-cycle allocation theory under uncertainty and related to empirical estimates of the value of life. A change in an age-specific mortality risk affects all subsequent survivor functions and reallocates consumption and labor supply over the entire life cycle. The value of eliminating a risk to life at a specific age is the expected present value of consumer surplus from that age forward.
paper, health improvements are perceived to be of two distinct types – reduction in mortality that leads to extended life (i.e. life expectancy), and rise in the quality of life. On the one hand, with increased life the utility from goods and services accrue over a longer period; on the other hand, improvements in the quality of life raise utility from given amounts of goods and leisure. In turn, both contribute to the increase of aggregate social value attributable to the increase in individuals’ maximization of lifetime expected utility. After deriving the value of a life (lives) for representative individual(s), the social value (i.e. increase of society’s ‘output’) of improvements in medical advances and public health infrastructure is derived by aggregating the benefits over the current and expected future population that benefit. They estimate economic gains from declining mortality in the USA over the 20th century, and to value the prospective gains that could be obtained from further progress against major diseases. Gains in life expectancy over the 20th century are found to be worth over 1.2 million USD per person to the current population. From 1970 to 2000 gains in life expectancy added about 3.2 trillion USD per year to national wealth, with half of these gains due to progress against heart diseases alone. Envisaging the prospect, the paper estimates that even modest progress against major diseases would be extremely valuable. For example, a permanent 1 percent reduction in mortality from cancer has a present value to current and future generations of Americans of nearly 500 billion USD, while a cure would be worth about 50 trillion USD.

4. CAUSALITY – A CRITICAL FACET IN THE HEALTH - INCOME RELATIONSHIP

The endogeneity problem, resulting from the feedback effects or measurement errors, along with unavailability of appropriate data to address this issue, stands in the way of providing an informative interpretation of causality. The presupposition of the existence of causality, nevertheless, is valid and very important to study. This rendered continuous engagement of researchers in theorizing about and developing models of the causal process in the realm of the health-income relationship, and evaluating those models using standards of evidences. There exist various techniques with which this problem may be tackled. The counterfactual approach, which defines a causal effect as the difference between an outcome that is observed when a specified treatment or event (or intervention) occurs and the outcome that would have ensued had the treatment or event (or intervention) not occurred, provides the core conceptual and notational framework for analyzing problems of causality in both statistics and econometrics (Bachrach and McNicoll, 2003, p.444). However, given that the experimental designs are generally infeasible in most social science research, the researchers have usually relied on observational data e.g. surveys, censuses, or administrative records etc. that are beset with formidable problems in estimating a causal effect. Particularly, two distinct sources of possible bias arise in the counterfactual analysis with observational data: (a) outcomes for the treatment and control groups may differ even in the absence of treatment, and (b) the potential effect of the treatment may differ for the treatment and control groups (Winship and Morgan, 19).

19 An important recommendation with regards to the analysis of causation is to begin with studying the effects of causes rather than trying to define what the causes of a given effect (Holland, 1986).
1999). In a similar vein Smith (2003) points to the various problems with the counterfactual account of causation, including conflation of direct and indirect effects of treatments, heterogeneity in people’s preferences for treatments and in the effects that treatments have on them, non-independence of individual-level units of observation, and assumptions about manipulability of causes. In the way forward to tackle various problems, the literature has come up with various methodological alternatives evolving around natural and quasi-natural experiments, and various forms of regression techniques.

Also, the potential pitfalls of applying the techniques based on individual-level analysis of causality to infer causal explanation at the aggregate level is emphasized by, for instance, Smith (2003), and Bhrolcháin and Dyson (2007). Particularly in the demographic studies, Bhrolcháin and Dyson (2007, p.1) assert that “aggregate phenomena and demographic change through time should be at the heart of demography, and therefore represent a central object of causal investigation in the discipline”. Consequently, as an alternative to the counterfactual or intervention methods, Bhrolcháin and Dyson (2007) propose an interesting approach by assigning several criteria for causation based on selected examples of aggregate demographic change. The criteria are: (i) **Time order** (i.e. that the cause should precede the effect); (ii) **Contiguity** (i.e. the shorter the time between the cause and the effect, the stronger the basis for causal inference); (iii) **Duration** (i.e. the causal inference is strengthened where the effect continues during the entire period in which the cause is operating); (iv) **Distinctiveness** (i.e. the causal inference is more straightforward where both cause and effect are clearly differentiated and identifiable in a temporal context); (v) **Direction** (i.e. that the effect should be in the expected direction); (vi) **Proportionality** (i.e. the causal linkage is better grounded when the scale of the effect can be considered proportional to the scale of the cause); (vii) **Recurrence** (i.e. the causal inference is strengthened if the linkage occurs in a variety of settings); (viii) **No cause, no effect** (i.e. where the putative cause is absent, the effect is absent too); (ix) **Mechanism** (i.e. to establish a causal link, a plausible set of intermediate links is required showing how the cause brings about the effect); and (x) **No alternative**, (i.e. all reasonable alternative explanations, including confounding, must be considered and ruled out) (Bhrolcháin and Dyson, 2007, p.25). They present several case studies related to various historical intervention events in the field of demography and link the outcomes with the aforementioned criteria. For example, there was an abrupt drop in births in Japan due to the folk belief that the year of the “Fire Horse” (i.e. 1966) is an ill-fated year for a girl to be born, when the crude rate fell by about 25 percent. Again, a maternity pay provision policy during the early 1980s in Sweden entitled the women who had a child within 24 months of having a

---

20 Winship and Morgan (1999) provides a survey of methodological alternatives on estimating causality with observational data. The list includes, inter alia, the use of control variables to eliminate potential differences between the treatment and control groups that are related to the outcome; standard regression and matching approaches that include the regression discontinuity design, propensity score techniques, and dummy endogenous variable models; use of instrumental variables (exclusion restrictions) to render feasible estimation of causal effects. While dealing with longitudinal data interrupted time-series design, differential linear growth rate models, the analysis of covariance etc. have been proposed. Hartwig (2008) applies dynamic panel Granger causality framework to test the Health care expenditure – growth relationship.
previous birth to continue to receive maternity pay in respect of the earlier birth. Such speed premium in Sweden caused a short-term rise in fertility and a longer-term change in the timing of births. Other cases examined by Bhrolcháin and Dyson (2007) include: the sharp rise of Romania’s fertility rate just after the abolition of the legal abortion in 1966; the consequence of famine on demography with reference to the Bangladesh famine during 1975-75; the abrupt changes in mortality and fertility due to the collapse of communism in eastern Europe; the impact of the family planning programs; and the cross-country nature of the fertility transitions. All of the above events, while examined along the ten criteria of causalities, provide useful insights into the direction of causation.

5. CONCLUSION

The political economy consequences of the nature of the relationship between health and income at the micro and aggregate level bear critical implications. Which policy options should the government pursue to achieve the health and other social targets? Some argue for a massive scaling up of public health and other social sector expenditure (e.g. Sachs, 2004; CMH, 2001; WHO 2002). A second school of thought favours a more multisectoral approach: emphasizing cross sector synergies, general investments in infrastructure, and improved governance for achieving better social-sector outcomes (see for example Leipziger et al., 2003). Others perceive social-sector development goals to be superfluous, and consider economic growth as the primary driver of social outcomes. Emphasis is placed on policies that target generalized macroeconomic growth more than anything else (e.g. Pritchett and Summers 1993; World Bank; 1993). Privatisation, deregulation, trade liberalization etc. have been the buzz words preached by the World Bank towards achieving economic stability and development, which often were enacted at the expense of social spending.

While social and economic gains contribute to the improvement in health, many of the gains are due to specific efforts to address major causes of disease and disability, such as providing quality and more accessible health services, introducing new medicines and health technologies, and promoting healthier behaviours (Levine et al., 2004). Examples of effective public health programs, not necessary hinging upon the national income level, exist to facilitate understanding the determinants of the changes in population health (see for example Levine et al., 2004; Chandra, 2006). Drawing on numerous successful cases of large scale public health interventions around the world Levine et al. (2004) identifies the elements of success: (i) predictable, adequate funding from both international and local sources; (ii) political leadership and champions; (iii) technological innovation within an effective delivery system, at a sustainable price; (iv) technical consensus about the appropriate biomedical or public health approach; (v) good management on the ground; and (vi) effective use of information. Reflecting on the above criteria, it has been found that large scale health interventions have worked successfully even in the world’s most underdeveloped and remote regions, in the face of extreme poverty and weak health systems.
Recently, Africa set a remarkable example of public health success in reducing measles cases and fatalities. The deaths from measles fell by 75% (from an estimated 506,000 to 126,000) during 2001-2006, which is due to the firm commitment and resources of national governments, and support from the Measles Initiative (a consortium of the American Red Cross, the United States Centers for Disease Control and Prevention (CDC), the United Nations Foundation, UNICEF and WHO) (UNICEF, 2007). Chandra (2006) asserts this incidence as striking in the context that the majority of these deaths have been preventable since 1960, when an economical and highly potent vaccine for measles was invented. It is apparent that the decline in child-mortality is not associated with improvements in African incomes, but the result of an effective international intervention that occurred 40 years after the discovery of the innovation of the measles vaccine. In the similar vein, public health campaigns to eradicate small-pox and cholera in 1960s and early 1970s in India may had been the consequence of political willpower, the presence of an extensive public health service, or generous funding from the WHO (Chandra, 2006). In a time span of 1985-1991, a regional polio elimination effort led by the Pan American Health Organization immunised every young child in the Americas resulting in the complete elimination of polio as a threat to public health in the Western Hemisphere. Reduction of guinea worm disease by 99% in 20 endemic African and Asian countries is the result of a multipartner eradication effort since 1986 fostering healthier behaviours (Levine et al., 2004). Researchers consider the economic return to investments in these public health improvements to be very high.

The crucial question then is whether the policy should follow the “Health-first” or a Growth-first” approach. In January 2000, the Commission on Macroeconomics and Health was established by World Health Organization to assess the place of health in global economic development in the realm of the health related Millennium Development Goals. The Commission made strong recommendations to promote health sector investments asserting that extending the coverage of crucial health services, including a relatively small number of specific interventions, to the world’s poor could save millions of lives each year, and would translate into hundreds of billions of dollars per year of increased income in the low-income countries. In this respect, quantification of health’s contribution will highlight the importance of investing in health in installing a virtuous cycle of economic development, which until now has been much less appreciated.
REFERENCES


Please note:

You are most sincerely encouraged to participate in the open assessment of this discussion paper. You can do so by posting your comments.

Please go to:


The Editor