How beneficent is the market? A look at the modern history of mortality

RICHARD A. EASTERLIN
University of Southern California, Department of Economics, Los Angeles, California 90089–0253, USA

A short answer is, not very. In contrast to economic growth, where the free market is commonly viewed as a prime mover, mortality in the last two centuries has not been greatly helped by market forces. The classic sources of market failure – information failures, externalities, public goods, principal-agent, and free rider problems – have been pervasive. Nor has economic growth itself been behind the worldwide improvement in life expectancy. In the primitive state of nineteenth century health knowledge, the immense rise in urbanisation engendered by economic growth largely vitiated any positive level-of-living effects by increasing exposure to disease. Instead, public policy initiatives, based on new knowledge of disease and new institutions, have been essential to the improvement of life expectancy, both in urban areas and nationwide.

Sweepings from butchers’ stalls, dung, guts, and blood,
Drowned puppies, stinking sprats, all drenched in mud,
Dead cats, and turnip tops, come tumbling down the flood.
– Jonathan Swift, A Description of a City Shower, 1710

1. Introduction

The concept of human development has recently emerged to rival economic development as a worldwide objective of public policy (Crafts 1997, Steckel and Floud 1997, United Nations Development Program 1990). In assessing human welfare, advocates of this concept would, at a minimum, place indicators of social conditions, notably life expectancy and educational achievement, on an equal footing with traditional economic measures like GDP per capita and a poverty index; some would go further and include indicators of political and civil liberties. In this article, I focus on only one of the proposed new measures, life expectancy at birth (referred to subsequently simply as life expectancy).

I think it is fair to say that, so far as life expectancy is concerned, the common reaction among many economists and economic historians is scepticism of the broader measure of human development. This is because improved life expectancy is typically viewed as a natural by-product of
economic development or of the institutional conditions that foster economic development. Thus, elevating life expectancy to the status of a social goal commensurate with economic growth raises no issues that are not already being dealt with in the study of economic growth and its determinants.

It is this view to which the present article is addressed. Specifically, the questions of concern here are: (1) is life expectancy largely or wholly a function of economic growth? (2) If not, are the conditions commonly taken to foster economic growth, namely, free markets, private property, and freedom of contract, also responsible for promoting the advance of life expectancy? As a basis for forming tentative answers to these questions, this article examines the nature and causes of the historical improvement in life expectancy over the period chiefly from 1800 to 1950. The experience of the developed countries since the 1950s is not included, because of the shift in their disease environment from primarily infectious to non-infectious diseases.

Economic development is taken here in Kuznets' (1966) sense of modern economic growth, and is proxied, as usual, by real GDP per capita. Evidence on income distribution, although desirable, is not available in the historical and geographical detail needed, and hence distributional issues are not included. Life expectancy is measured by life expectancy at birth, the magnitude used in the United Nation's human development index. Because this measure is less sensitive to improvements in mortality at older than younger ages, one might prefer a measure such as the age-standardised death rate, but again the requisite data are not readily available.

The analysis starts with a brief conceptual section that helps highlight the main issues and most influential recent studies of historical experience. It then turns in Section 3 to the historical evidence on life expectancy and economic development in several developed and developing countries with relatively good data to assess in preliminary fashion the time series association between the two. Section 4 turns to the nature of technological change in infectious disease control. Section 5 considers the role of market forces in promoting the adoption and development of the new techniques of disease control. Finally, Section 6 describes the institutional innovations required to implement the new technology of disease control.

2. Conceptual background

A useful point of departure is a widely-cited article by demographer Samuel Preston (1975). Preston suggests that the improvement in life expectancy can be viewed as due to either of two components: (1) that arising from a movement along what economists would call an aggregate 'health production function', relating life expectancy to real GDP per capita, and (2) that due to an upward shift in the function caused by 'technological change', the ability to use given resources more productively to control disease and lengthen life (Figure 1). Preston's empirical analysis based on data for
Figure 1. Sources of increased life expectancy.

Source: Based on Preston (1975).

1900–1960 concludes that a shift in the function is the overriding source of improvement.¹

In explicitly separating economic development (as measured by GDP per capita) from other sources of life expectancy improvement, Preston’s analytical scheme seemingly highlights nicely the issue of economic growth as the source of the advance of life expectancy. To those who see development as the prime mover, it is a movement along the aggregate health production function that is chiefly responsible for improved life expectancy.

It is implicitly such a movement that is stressed in what is widely recognized as the most influential book on historical mortality in recent decades, Thomas McKeown’s The Modern Rise of Population (1976). Reasoning chiefly from cause-of-death data for England and Wales from the mid-nineteenth century onward, McKeown argues that medical advances cannot explain the observed mortality decline. He concludes that economic growth, and, in particular, improved nutrition, must be responsible for the observed reduction in mortality, although, as he admits and critics have repeated, no direct evidence on the nutrition of individuals is presented (cf. McKeown 1976, p. 130, and critiques by Szreter 1988, Curtin 1989, pp. 41–2, and Mercer 1990).

A recent article by S. Ryan Johansson sees McKeown’s work as import-

¹ Some will note the parallel with Nobel Laureate Robert Solow’s (1957) analysis of the sources of economic growth.
antly influencing the views of today’s neoclassical economists by apparently ‘providing scientific evidence that market forces solved health and mortality problems in the past’ (Johansson 1994, p. 108). Johansson charges economic historians generally, and especially Nobel laureate Robert W. Fogel in his project on human stature, with fostering this view. In fairness to Fogel, his discussion of the secular mortality decline specifically mentions factors that would shift the health production function. But his portmanteau use of the terms nutrition and nutritional status, plus the fact that his substantive discussion is almost entirely about diet, calorie consumption, and food supply, might lead some readers to suppose that he is only talking about movements along the health production function.

The separability of movements along the function from shifts in the function is not as simple as it seems (cf. Nelson 1973). I shall argue in Sections 3 and 4 below that not only the level, but the slope, of the aggregate worldwide function relating life expectancy to economic growth depends on the technology of disease control. Throughout much of the nineteenth century the slope was not much different from zero, because the mortality-raising agglomeration effects of economic growth largely offset any beneficial effects arising from improved living levels. New techniques of controlling major infectious disease, that came gradually into use in the last half of the nineteenth century, increased both the slope and the level of the function.

In the case of economic growth the underlying advances in technology have been in large part the result of decisions by individual firms. In the case of infectious disease control the decision-making units are much more diffuse. Firms – in the sense of private medical practitioners of all types – may employ better methods of treating disease. As will be shown, however, most of the improvement in life expectancy has been due to preventive rather than therapeutic measures, and the contribution of firms to preventive measures has been small (the chief exception being medical practitioners dispensing hygienic advice). The most important decision-making units have been households and governments – households, because such a wide array of decisions relating to household operation and the household environment are crucial to preventing disease – and governments, because the new methods of disease control typically necessitated government action. Of the two, governments have been more fundamental than house-

2 Fogel 1986, pp. 443–7; Goldin 1995, p. 205; Lindert 1986, pp. 531–3; and Perenoud 1991, pp. 20–21. The conceptual problem – of which Fogel himself is fully aware (1986, pp. 446–7) – is that nutritional status is affected not only by nutritional intake, but also claims against it, especially those due to disease. The reduction or elimination of disease can improve nutritional status and increase stature with no change in intake. Hence, trends in stature cannot be taken as a proxy for trends in per capita income or real wages, as Fogel and his collaborators initially suggested (Fogel et al. 1983, pp. 247–8).
holds, because the adoption of new household methods required health education programmes that were largely promoted by governmental agencies. Thus, thinking of the aggregate health production function in terms of its micro-level counterparts, the subsequent analysis suggests that the advance of knowledge shifted the production function of governments, households, medical practitioners, and non-medical firms, but that the pivotal player bringing about the shifts was the government.

3. Life expectancy and economic growth

This Section addresses the presumption that the historical improvement in life expectancy is due to the favourable effect of economic growth on living levels. It first presents evidence that the long term trends in life expectancy and economic growth are not closely related, drawing on the record of six countries with relatively good data. It then considers why the relationship between the two has been quite weak in the historical past (see also Kunitz and Engerman 1992).

The first point of note is that life expectancy is marked by a sharp increase in the rate of improvement, starting in the late nineteenth century. Borrowing from Rostow (1959), I shall call this a ‘take-off’ to signify a substantial shift from a lower to higher rate of change. As in the case of economic growth, the precise dating of this take-off is somewhat arbitrary and varies from one country to another.

In England and Wales – for whom the historical record of life expectancy is longest, dating from 1541 – the take-off occurs around 1871. Prior to this there is a slow improvement in life expectancy that starts in the first half of the eighteenth century, but this does little more than raise life expectancy by 1871 to a level only slightly higher than around 1600 (Figure 2). In Sweden, another country for which the historical record is well-researched, the take-off occurs around 1875 (Figure 3). As in England, there is a preceding period of mild improvement dating from the late eighteenth century, which may also have been preceded by a period of worsening mortality.

The pattern of mild, followed by rapid, improvement occurs in four other countries included here (Figures 4–7). In France a take-off occurs around 1893; in Japan, 1923; and in the Third World countries of Brazil and India, around 1940 and 1945, respectively. Taking all six countries together, the rate of improvement in life expectancy in the half century after take-off is

---

3 Negative as well as positive living level effects of economic growth – e.g., via tobacco and alcohol consumption – are usually recognised, but the positive effects are typically assumed to predominate until fairly high income levels are reached (cf. Auster et al. 1969, pp. 134–6).
Figure 2. Life expectancy in England and Wales since the sixteenth century.


Figure 3. Life expectancy in Sweden since 1778–82.

The market and the modern history of mortality

Figure 4. Female life expectancy in France since 1816–20.


Figure 5. Life expectancy in Japan since 1875.

Figure 6. Life expectancy in Brazil since 1872.


Figure 7. Life expectancy in India since 1896.

Table 1. Life expectancy improvement in the half century before and after take-off in six countries.

<table>
<thead>
<tr>
<th>Country</th>
<th>Take-off date</th>
<th>Life expectancy at take-off (years)</th>
<th>Change in half century</th>
<th>Ratio (4)/(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>England and Wales</td>
<td>1871</td>
<td>41.0</td>
<td>3.0</td>
<td>12.0</td>
</tr>
<tr>
<td>Sweden</td>
<td>1875</td>
<td>45.4</td>
<td>4.6</td>
<td>17.2</td>
</tr>
<tr>
<td>France*</td>
<td>1893</td>
<td>44.9</td>
<td>3.4</td>
<td>20.3</td>
</tr>
<tr>
<td>Japan</td>
<td>1923</td>
<td>42.6</td>
<td>5.8</td>
<td>30.8</td>
</tr>
<tr>
<td>Brazil</td>
<td>1940</td>
<td>36.7</td>
<td>8.0</td>
<td>28.9</td>
</tr>
<tr>
<td>India</td>
<td>1945</td>
<td>32.1</td>
<td>8.3</td>
<td>28.3</td>
</tr>
</tbody>
</table>

Note: *Data are for females.
Sources: See notes to Figures 2–7.

Table 2. Take-off dates for economic growth and life expectancy in six countries.

<table>
<thead>
<tr>
<th>Country</th>
<th>Economic growth</th>
<th>Life expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>England and Wales</td>
<td>1783–1830</td>
<td>1871</td>
</tr>
<tr>
<td>Sweden</td>
<td>1868–1890</td>
<td>1875</td>
</tr>
<tr>
<td>France</td>
<td>1830–1870</td>
<td>1893</td>
</tr>
<tr>
<td>Japan</td>
<td>1885–1905</td>
<td>1923</td>
</tr>
<tr>
<td>Brazil</td>
<td>1933–1950</td>
<td>1940</td>
</tr>
<tr>
<td>India</td>
<td>1952–1963</td>
<td>1945</td>
</tr>
</tbody>
</table>

Sources: Figures 2–7 and Rostow (1978).

from three to six times greater than in the half century before (Table 1, cols. 3 and 4).

If the McKeown hypothesis is correct, then the take-off dates for life expectancy should conform closely to those for economic growth. But the two do not fit closely at all (Table 2). The contrast between Sweden and England is striking. Their take-offs in life expectancy are almost identical, but those in economic growth differ by about three-quarters of a century.

Taken as a whole, the evidence for the six countries suggests that the rapid improvement in life expectancy started later than modern economic growth, but spread more rapidly. In none of the six countries did rapid improvement in life expectancy start before the 1870s, whereas economic growth was underway in two, perhaps three, of the six countries by that time. But the time span of the take-off dates for life expectancy is much shorter than that for the take-offs in modern economic growth – about seven
decades compared with seventeen. Also, in the four countries in which life expectancy takes off before 1940 there is little evidence of any impact on life expectancy of the substantial retardation in economic growth between World Wars I and II.

My interest here is in the factors responsible for the high rate of improvement in life expectancy post take-off, because it is this improvement that has lifted life expectancy to wholly unprecedented levels. But a brief word is needed about the mild improvement observed in all six countries before take-off, because of the temptation to assume that this is the product of economic growth. It should be noted, first, that neither India nor Sweden provide much support for this assumption – there is little or no increase in *per capita* income in India in the first half of the twentieth century or in Sweden prior to 1850 (Maddison 1995, pp. 196, 204; Bengtsson 1997). Second, other explanations are possible. In all six of the countries included here the phase of mild improvement in life expectancy is associated with a reduction in smallpox mortality as vaccination was introduced (Fenner *et al.* 1988). Not only was smallpox itself a major cause of death and thus lowered mortality directly, but its reduction may also have lowered mortality from other diseases by reducing the proportion in the population of persons whose immune systems had been seriously damaged by smallpox (Mercer 1990, Sundin 1995, Skold 1996). Other factors, such as a gradual improvement in hygienic knowledge may also have been at work (Riley 1987). Thus, economic growth is by no means the most obvious explanation of the phase of mild improvement.

Indeed, there is reason to doubt whether the acceleration in economic growth that started in western Europe after 1800 was having any sizeable positive effect on life expectancy. Some analysts have pointed out that in the middle of the nineteenth century there is a flattening of the trend in life expectancy both in England and France, despite rising income levels (cf. Figures 2 and 4, and Fogel 1994; Schofield, *et al.*, 1991, ch. 1; Szreter 1988, 1997; Vallin 1991). They suggest that the positive effect of economic growth on life expectancy due to better living conditions was countered by another effect of economic growth – the redistribution of population to high-mortality urban centres. Before its take-off, life expectancy in urban areas was about ten years less than in rural, reflecting the more rapid spread of disease where population density is high and also under the crowded conditions of factory production as modern economic growth took hold (note the estimates for early dates in Figure 8). As the population became more concentrated in low-life-expectancy urban areas, there was a negative effect of economic growth on life expectancy. This effect can be illustrated for England and Wales. Between 1831 and 1861 the proportion of population in urban areas rose from about one-third to one-half, and this redistribution

---

4 The specific mechanisms linking economic growth and urbanisation are detailed in Easterlin (forthcoming).
The market and the modern history of mortality

A. Deficit of London versus rural England and Wales, 1811–1911.

B. Deficit of Paris versus France, 1816–20 to 1901–05.


Figure 8. Shortfall of urban life expectancy, specified country and period.

Table 3. Illustrative calculation of effect of urbanisation on life expectancy in England and Wales between 1831 and 1861.

(1) Analytical relationship:

\[ e_0' = e_0^u \left( \frac{t}{u} \right) + e_0^r \left( \frac{t - u}{r} \right) \]

where \( e_0 \) = life expectancy at birth, and \( t, u, r \) refer to total, urban, and rural population respectively.

(2) England and Wales, 1831 actual:

\[ 40 = 33 \times (0.34) + 43 \times (0.66) \]

(3) England and Wales, 1861, if only rural-urban shift in population distribution:

\[ 38 = 33 \times (0.49) + 43 \times (0.51) \]

(4) England and Wales, 1861 actual:

\[ 41 = 38 \times (0.49) + 45 \times (0.51) \]

Source: Data from Woods (1985).

would, ceteris paribus, have reduced life expectancy by two years, from 40 to 38 years (Table 3, panels 2 and 3). In fact, life expectancy rose by one year between 1831 and 1861, as improvements within the rural and urban sectors slightly outweighed the effect of population redistribution (panels 2 and 4). These within-sector increases could be due to economic growth, the reduction of smallpox, or other factors.

The implication of the agglomeration effects of economic growth for the mid-nineteenth century slope of Preston’s aggregate health production function relating life expectancy to economic growth is noteworthy. At the nationwide level the positive slope commonly taken to characterise the relation between life expectancy and per capita income is sharply reduced, because higher income brings with it higher urbanisation, and thus greater exposure to infectious disease.5 Indeed, some scholars argue that the slope of the relationship in a pre-modern mortality regime may have been negative (Mosk and Johansson 1986). This inference regarding the pre-modern slope of the function implies that the great historical improvement in life expectancy had to be due to a shift in the function, because movements

---

5 The flatness of the nationwide slope of the function is consistent with positive slopes within both the rural and urban sectors. In the Table 3 illustration, for example, the increases of \( e_0 \) of 5 and 2 years within each sector might be assumed (contrary to the argument above) to be due entirely to income growth within each sector. These increases are greater than the nationwide increase in \( e_0 \) of 1 year associated with this income growth, which reflects in addition, the adverse effect of urbanisation generated by economic growth.
along the function would have had, at best, a mild positive effect. But if the function shifted, then there ought to be evidence of ‘technological change’, that is, of new methods of preventing or curing disease. The next Section turns, therefore, to the historical evidence of technological change in disease control.

4. Technological change and life expectancy

The starting point for any history of disease control technology must be recognition of the appallingly low state of knowledge of disease in the first part of the nineteenth century. At that time there was no correct knowledge of the causes of disease, very little of the mode of transmission, and almost none of how to treat disease. This is not to say that there were no beliefs on these matters – quite the contrary, there was a firmly established body of doctrine on the nature, causes, and treatment of disease. But these beliefs were as likely to be counter-productive as productive, centering, as they did, on treatment by means of emetics, cathartics, diuretics, and bleeding (Rosenberg 1979, p. 13). Note the contrast with the state of knowledge regarding methods of production. When modern economic growth got started, people already knew how to grow food and manufacture goods – the technology of economic growth increased the ability to do better what people were already doing successfully. However, with regard to controlling disease, the fact is that there was very little useful knowledge before the mid-nineteenth century. Consider the example of a Philadelphia tallow chandler in the fall of 1826 who ‘complained of chills, pains in the head and back, weakness in the joints and nausea . . .’:

[B]efore seeing a regular physician, he was bled till symptoms of fainting came on. Took an emetic, which operated well. For several days after, kept his bowels moved with Sulph. Soda, Senna tea etc. He then employed a Physician who prescribed another Emetic, which operated violently and whose action was kept up by drinking bitter tea (ibid.; cf. also Starr 1982, pp. 32–7; Warner 1986).

Elsewhere, the author who quotes this passage, in a masterpiece of under-statement, observes: ‘[I]t is difficult to recapture the medical world of 1800 . . . a world of thought structured about assumptions alien to a twentieth-century medical understanding’ (Rosenberg 1987, p. 71).

This is not to say that before the nineteenth century there had been no improvement whatsoever in knowledge relevant to the control of disease, but given the long history of humanity, the advances that had occurred were surprisingly recent and had had only limited effect (Hall 1967, Ackernecht 1968, Dixon 1978, ch. 2). The most important practical advances had been the use of quarantine and cordons sanitaires starting in the fourteenth century to prevent the spread of plague, and the development in the latter part of
the eighteenth century of inoculation and then vaccination against smallpox.

The major breakthroughs of the nineteenth and twentieth centuries that were eventually to bring fatal infectious diseases generally under control can be briefly outlined. They fall under three principal headings:

1. New methods of preventing the transmission of disease, including education of the public, starting in the mid-nineteenth century,
2. New vaccines to prevent certain diseases, starting in the 1890s, and
3. New drugs to cure infectious disease (antimicrobials), starting in the late 1930s.

The first major step in preventing the transmission of infectious disease came with what has come to be called the ‘sanitation revolution’. Starting in the 1840s in England, this was a movement aimed at cleaning up the cities through purer water supplies, better sewage disposal, paving streets, education in personal hygiene, and the like. Though based on a misguided theory of disease transmission – the miasmatic theory, which linked disease to bad smells and vapours – its emphasis on cleaning up public places and homes led to a gradual reduction in the transmission of waterborne and airborne diseases. The sanitation revolution is usually dated from Edwin Chadwick’s landmark 1842 Report on the Sanitary Condition of the Labouring Population of Great Britain. This report, and similar studies elsewhere (Griscom [1845]1970, Shattuck et al. [1850] 1948, Citizens’ Association of New York, 1866) assembled demographic data and the testimony of medical experts to document the association between filth and high mortality. The domestic household counterpart of the sanitation revolution was a new emphasis on cleanliness. As Wohl (1983, p. 66) states: ‘If sanitary engineering associated with Chadwick represents the public face of the public health movement, the less well-known private aspect is represented in the efforts of the voluntary health visitors and sanitary workers who, entering the homes of the poor, tried to scour the inhabitants as well as their flats’.

Next came a series of discoveries establishing how certain diseases were specifically transmitted (Table 4, panel A). Two critical breakthroughs were the mid-nineteenth century discoveries of Snow and Budd that identified impure water as a vehicle for the transmission of two highly-feared killers, cholera and typhoid. The specific evidence of impure water as a carrier of disease helped strengthen the case for the reforms being urged by the sanitationists. Also, in 1867 Joseph Lister, influenced by Pasteur’s research on the bacteriological origins of disease, introduced antiseptic surgery, starting a trend toward sharply diminished mortality in surgical procedures by reducing the transmission of infection during surgery (Biraben 1991, Gariepy 1994).
By the last quarter of the nineteenth century the discoveries of Pasteur, Koch and others, and the laboratory techniques and methodology that had been developed, had laid the foundation for the new science of bacteriology, and essentially validated the germ theory of disease. For the first time the causal agents in a number of major diseases were identified (Table 4, panel B). Further breakthroughs also occurred in identifying the mode of transmission of specific diseases, most notably, the vectors carrying the germs of two major killers, malaria and yellow fever (Panel A). A basis was laid for the systematic development of immunology, and a new systematic approach opened for the prevention of disease by the development of vaccines (Parish 1965, Plotkin and Mortimer 1988). The conquest of diphtheria by von Behring in 1892 was the first in a series of developments that brought a number of major infectious diseases under control via immunisation (Table

**Table 4. Discoveries in the control of major fatal infectious diseases since 1800: mode of transmission and causal agent.**

<table>
<thead>
<tr>
<th>Date</th>
<th>Disease</th>
<th>Investigator</th>
<th>Date</th>
<th>Disease</th>
<th>Investigator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1847</td>
<td>Measles</td>
<td>Panum</td>
<td>1880</td>
<td>Typhoid (bacillus found in tissues)</td>
<td>Eberth</td>
</tr>
<tr>
<td></td>
<td>Puerperal fever</td>
<td>Semmelweis, Holmes</td>
<td></td>
<td>Leprosy</td>
<td>Hansen</td>
</tr>
<tr>
<td>1854</td>
<td>Cholera</td>
<td>Snow</td>
<td>1882</td>
<td>Malaria</td>
<td>Laveran</td>
</tr>
<tr>
<td>1859</td>
<td>Typhoid fever</td>
<td>Budd</td>
<td></td>
<td>Tuberculosis</td>
<td>Koch</td>
</tr>
<tr>
<td>1867</td>
<td>Sepsis (surgical)</td>
<td>Lister</td>
<td></td>
<td>Glanders</td>
<td>Loeffler and Schutz</td>
</tr>
<tr>
<td>1898</td>
<td>Malaria</td>
<td>Ross, Grassi Looss</td>
<td>1883</td>
<td>Cholera Streptococcus (erysipelas)</td>
<td>Koch Fehleisen</td>
</tr>
<tr>
<td>1900</td>
<td>Yellow fever</td>
<td>Reed</td>
<td>1884</td>
<td>Diphtheria</td>
<td>Klebs and Loeffler</td>
</tr>
<tr>
<td>1906</td>
<td>Dengue</td>
<td>Bancroft</td>
<td>1885</td>
<td>Typhoid (bacillus isolated)</td>
<td>Gaffky</td>
</tr>
<tr>
<td></td>
<td>Rocky Mountain spotted fever</td>
<td>Ricketts, King</td>
<td>1886</td>
<td>Pneumococcus</td>
<td>A. Fraenkel</td>
</tr>
<tr>
<td></td>
<td>Typhus</td>
<td>Nicolle</td>
<td>1887</td>
<td>Malaria</td>
<td>Bruce</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Soft chancre</td>
<td>Ducrey</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gas gangrene</td>
<td>Welch and Nuttall</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1892</td>
<td>Plague</td>
<td>Yersin, Kitasato</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Botulism</td>
<td>van Ermengem</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1894</td>
<td>Dysentery bacillus</td>
<td>Shiga</td>
</tr>
</tbody>
</table>

Table 5. Discoveries in the control of major fatal infectious diseases since around 1800: vaccines and drugs.

<table>
<thead>
<tr>
<th>Date</th>
<th>Disease</th>
<th>Developer</th>
<th>Date</th>
<th>Drug</th>
<th>Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1798</td>
<td>Smallpox</td>
<td>Jenner</td>
<td>1908</td>
<td>Salvarsan</td>
<td>Ehrlich</td>
</tr>
<tr>
<td>1881</td>
<td>Anthrax</td>
<td>Pasteur</td>
<td>1935</td>
<td>Sulfanomides</td>
<td>Domagk</td>
</tr>
<tr>
<td>1885</td>
<td>Rabies</td>
<td>Pasteur</td>
<td>1941</td>
<td>Penicillin</td>
<td>Fleming, Florey, Chain</td>
</tr>
<tr>
<td>1892</td>
<td>Diphtheria</td>
<td>von Behring</td>
<td>1944</td>
<td>Streptomycin</td>
<td>Waksman</td>
</tr>
<tr>
<td>1896</td>
<td>Cholera</td>
<td>Kolle</td>
<td>1947</td>
<td>Broad spectrum</td>
<td>antibiotics$^a$</td>
</tr>
<tr>
<td>1906</td>
<td>Pertussis</td>
<td>Bordet-Gengou</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1921</td>
<td>Tuberculosis</td>
<td>Calmette, Guerin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1927</td>
<td>Tetanus</td>
<td>Ramon, Zoeller</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td>Yellow fever</td>
<td>Theiler</td>
<td>1948</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Typhoid fever</td>
<td>Weigl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950</td>
<td>Polio</td>
<td>Salk</td>
<td>1954</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1954</td>
<td>Measles</td>
<td>Enders, Peebles</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Panel B: Baldry (1976).


The developments summarised so far were techniques that reduced mortality through the prevention of disease, but the ability to cure disease remained elusive, and, as in the past, physicians could do little to help those who were seriously ill (Thomas 1983, chs 3–5). The successful development of antimicrobials that could attack the newly identified causes of disease without harmful side effects was the next major step in bringing infectious disease under control, but it did not come until almost a half century after the causes of a number of diseases had been found. The discovery of sulfanomides in 1935 was a start, but the most important breakthrough in terms of general effectiveness was the development of penicillin in 1941 and the long list of other antibiotics to which it subsequently gave rise (Table 5, panel B; Bottcher 1964, Baldry 1976).$^6$

$^6$ For a striking demonstration of the advance in medical therapy after the 1930s, compare the recommended treatments of major infectious diseases in Winslow (1931) with those in Beeson (1980).
Thus, in little more than a century the ability to control infectious disease was totally transformed – first by techniques that prevented the spread of certain major infectious diseases, then by vaccines that protected people from contracting some of these diseases, and, finally, by the development of cures. The evidence of major advances in the knowledge and technology of disease control from the mid-nineteenth century onward, such as those listed in Tables 4 and 5, seems indisputable. But since any such enumeration of discoveries and technological breakthroughs is somewhat subjective, it may be helpful to quote a summary assessment of the progress of knowledge given in 1983 by a medical historian: ‘In a single century the understanding of disease increased more than in the previous forty centuries combined. The two crucial developments in this regard were the rise of technology and the application of the basic biological sciences to medicine, using new rules of experimentation and new criteria of proof’ (Hudson 1983, p. 121).

As these new techniques of disease control were introduced, mortality rates plunged, life expectancy took off, and non-infectious gradually replaced infectious diseases as the leading causes of death. This development, known as the ‘epidemiologic transition’ (Omran 1971, Bobadilla et al. 1993, Gribble and Preston 1993) is illustrated by cause-of-death data for England and Wales (Table 6). The fact that a large share of the mortality decline there took place before 1940 makes clear that much of the control of infectious disease was accomplished by preventive measures, before the introduction of antimicrobials.

One would expect that the new techniques of disease control would improve life expectancy more rapidly in urban than rural areas, and that the gap between the two areas would consequently narrow. This is because the sanitation revolution was first and foremost a drive to clean up the cities.

Table 6. Death rate and percentage distribution of deaths by cause, England and Wales, 1871–1951 (age-standardised).

<table>
<thead>
<tr>
<th></th>
<th>1871</th>
<th>1940</th>
<th>1951</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death rate (per thousand)</td>
<td>22.4</td>
<td>9.3</td>
<td>6.1</td>
</tr>
<tr>
<td>All causes</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Infectious diseases</td>
<td>31</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Bronchitis, pneumonia and influenza</td>
<td>14</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>Diseases of the circulatory system</td>
<td>9</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td>Diarrhea and enteritis</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Accidents</td>
<td>4</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Neoplasms</td>
<td>2</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Other causes</td>
<td>36</td>
<td>29</td>
<td>24</td>
</tr>
</tbody>
</table>

Source: Caselli (1991, pp. 89–90). Data are averages of males and females. Detail may not add to total because of rounding.
Moreover, efforts to educate the public on the importance of personal hygiene were directed especially at, and more easily reached, the highly concentrated urban than the widely dispersed rural population. And, in fact, the historical shortfall of urban compared with rural life expectancy was steadily eliminated (Figure 8 above). The initial differential and subsequent trend in rural versus urban life expectancy is the opposite of what one would expect based on per capita income. Although per capita income was initially lower in rural areas (Williamson 1981, 1982), life expectancy was higher. And while rural income grew more rapidly, converging toward urban levels, life expectancy grew more slowly (Preston and van de Walle 1978, p. 279, Preston et al. 1981, Sawyer 1981).

The point was previously made that under the mortality regime prevailing in the first half of the nineteenth century – before the onset of sustained advance in the knowledge and technology of disease control – the positive relation between life expectancy and per capita income that would be expected on the basis of improved living levels was undercut by the positive association between urbanisation and per capita income. The differential trend between rural and urban mortality brought about by the new technology of disease control significantly altered this relationship. As unfavourable urban health conditions were gradually removed by the new techniques, and excess urban mortality eliminated, the adverse effect of urbanisation on life expectancy evaporated, leaving only the positive effect of per capita income via higher living levels. Thus the new technology of disease control had the effect of increasing the slope of Preston’s functional relationship between life expectancy and per capita income, as well as shifting that relationship upward (cf. Mosk and Johansson 1986, p. 420).

Associated with the epidemiological transition there was also a take-off in stature much like that in life expectancy. Recent work by economic historians has increasingly recognised that stature is a function of disease as well as diet, because disease seriously affects the capacity of the body to retain nutrients (Steckel 1995, Engerman 1997, Steckel and Floud 1997). Because the epidemiological transition especially reduced the mortality and illness of infants and the young, among whom the incidence of infectious disease is highest, one would expect the transition to have beneficial effects on stature. That this is so is indicated by the time series assembled by Fogel and his collaborators. In the six European countries for which historical estimates are available, the average improvement in male stature in the century prior to the third quarter of the nineteenth century was 1.1 centimetres. In the subsequent century – the period of the epidemiological transition – it was

---

7.7 centimetres (Fogel 1993, p. 20; Easterlin 1996, p. 82). In every one of the six countries the rate of improvement in stature was considerably higher in the most recent century than in the earlier. As was noted above regarding the mild pretake-off improvement in life expectancy, even the early improvement in stature may partly reflect a reduced incidence of infectious disease.

5. Life expectancy and the market

To recapitulate the argument to this point, in the state of knowledge regarding health and mortality prevailing in the early nineteenth century, economic growth had at best only a small positive effect on life expectancy. This is because the positive effect of economic growth via improved living levels was substantially offset by a growing exposure to disease as the population became more urbanised. Only as advances in knowledge led to the development and use of new methods of controlling infectious disease both in urban areas and nationwide did rapid improvement in life expectancy occur.

But even if economic growth via its effect on living levels was not directly responsible for the great improvement in life expectancy, isn’t it possible that the institutions that fostered economic growth were responsible for the development and use of the new methods of controlling infectious disease? Rodrik (1996) has noted the broad consensus among economists that economic growth is fostered by free markets, private property, and enforcement of contracts, exemplified in economic history by the influential work of Douglass North (1990). Weren’t these same institutions at work on the problem of infectious disease? Incomes were rising and infectious disease was an important concern; wasn’t it profitable for firms to attack the problem?

It is this question – the role of the market in the great improvement of life expectancy – to which this Section is addressed. By the ‘market’ I mean supply and demand conditions operating within the institutions of private property and free contract to allocate resources via the incentive of private profit to the satisfaction of human wants – in this case wants with regard to the elimination of disease and reduction of mortality.

My approach to answering this question is to look at the specific techniques that reduced infectious disease so dramatically – control of the mode of transmission, immunisation, and antimicrobials – and consider, first, the extent to which the market allocated resources to each.

5.1. Control of the mode of transmission of disease

For the present purpose, it is helpful to think of the techniques under this head as falling into two groups – those requiring a change in the contami-
nating behaviour of individuals, firms, and other agents, and those calling for correction of environmental conditions.

The contaminating behaviours of individuals encompass such things as coughing, sneezing, spitting, and nose-blowing; toilet habits; behaviour in regard to personal washing and bathing; practices regarding the sources and handling of drinking water and milk; methods of food handling and preparation; customs regarding the care and feeding of infants and children; practices relating to care of the sick; sexual behaviours; and attitudes toward rodents and insects. In the nineteenth century everyday behaviour in all of these respects generated significant negative externalities with regard to infectious disease. A few examples: spitting by men on the floor at home and in public places was often an accepted behaviour; the fly, rather than being regarded as a carrier of disease, was thought of affectionately as the ‘friendly fly’ (Rogers 1989, Tomes 1998, p. 98). Writing of the habits of the poor in Wakefield, England in 1869, Sir John Simon reports, ‘people are seen easing their bowels into the beck [stream] which afterwards supplied them with drinking water’ (Wohl 1983, p. 94).

At the firm level, worker and management practices fostering the spread of disease (often unwittingly) were also common. The following is an example from a statement by Stephen Smith, a physician, on the results of a sanitation survey in New York City in 1865:

> I hold in my hand a list of cases of smallpox found existing under circumstances which show how widespread is this disease. Bedding of a fatal case of smallpox was sold to a rag-man; case in a room where candy and daily papers were sold; case on a ferry-boat; woman was attending bar and acting as nurse to her husband who had smallpox; girl was making cigars while scabs were falling from her skin; seamstress was making shirts for a Broadway store, one of which was thrown over the cradle of a child sick of smallpox; tailors making soldiers’ clothing, have their children, from whom the scabs were falling, wrapped in the garments; a woman selling vegetables had the scabs falling from her face, among the vegetables, etc. etc.(Smith 1911, pp. 108–9).

A description of mid-nineteenth century London’s ‘town dairies’ – ‘half-underground dens and cellars in which the cows were kept for the greater part of the year, standing knee-deep in filth’ – states that ‘it was difficult to find a sample of London milk which would fail to show the presence of blood or pus when examined under the microscope’ (Drummond and Wilbraham, 1939, pp. 299–300).

Such behaviours and practices, of central significance for the transmission of disease, are not a simple function of income and prices. They are rooted in the established norms of society, and its customs and beliefs. Each generation as it is raised internalises various health beliefs and learns what is socially acceptable behaviour. Historically, the market, by crowding people more closely together in towns, cities, and factories, magnified the
negative externalities of disease-transmitting behaviours and practices. Writing in 1842 on *The Condition of the Working Classes in England*, Friedrich Engels observed: ‘Dirty habits ... do no great harm in the countryside where the population is scattered. On the other hand, the dangerous situation which develops when such habits are practiced among the crowded population of big cities, must arouse feelings of apprehension and disgust’ (as quoted in Wohl 1983, p. 4).

Nor can this set of behaviours be corrected by the simple assignment of property rights. In the absence of knowledge of the mechanisms of disease causation and transmission, such assignment is not even conceivable. But given such knowledge, enforcement is not possible because of the overwhelming magnitude of transaction costs. ‘If you had to sue everybody who sneezed in your vicinity, you would have no time remaining for any other activity ... [S]ocial customs and “manners” create society’s best control mechanism’ (Phelps 1992, p. 418). But appropriate customs and manners do not arise spontaneously. They result chiefly from increasing awareness among the public of the consequences of one’s actions for the spread of disease – awareness that depends on the dissemination of new knowledge of disease.

The second principal source for preventing the transmission of infectious disease has been correction of environmental conditions that expose population to disease. Here too the contribution of market forces has been dubious. Some environmental techniques for controlling contagious disease, such as insect or rodent control, are quite clearly public goods. The individual may take defensive measures – the use of screens, mosquito netting, rat traps, etc. – but in situations of dense habitation these are likely to be ineffective in the absence of community action. What is needed are measures that go beyond the individual’s resources, such as the spraying of insecticides on the breeding grounds of insects (Musgrove 1996, p. 11).

Some environmental conditions important for the control of infectious disease do involve goods that are or have been provided by the market to some extent. This is notably true with regard to those conditions that were the initial target of the sanitation revolution – improved water supply and waste disposal. Didn’t rising income generate a growing demand for these goods, and their resulting supply work to remove this source of infectious disease?

To answer this, it is helpful to start by recalling that in the mid-nineteenth century the flush toilet was a rarity – in cities the most common facility was a vault privy, modelled on its country cousin, the out-house, but in poor neighbourhoods, even these were rare.8 The result was burgeoning

8 ‘[I]n mid-century Darlington: “In 1 yard 66 persons are obligated to use 1 privy; in another 65, and in a third 63, in a fourth 54, in a fifth 45, in a sixth 41, in a seventh 35 and so on”’ (Wohl 1983, p. 87; cf. also Winslow 1943, pp. 244–5).
accumulations of human excrement as city size rocketed. In some areas, these accumulations, because of their potential use as manure, had value as an economic good – there was a saying that the ‘chamber pot is a penny savings bank’ (Drummond and Wilbraham 1939). But the resulting market only aggravated the problem of infectious disease, because of accompanying negative externalities. It is worth repeating an oft-quoted and apt passage from Chadwick describing conditions in British towns around 1840:

In one part of the street there is a dunghill, – yet it is too large to be called a dunghill. I do not mistate its size when I say it contains a hundred cubic yards of impure filth, collected from all parts of the town. It is never removed; it is the stock-in-trade of a person who deals in dung; he retails it by cartfuls. To please his customers, he always keeps a nucleus, as the older the filth is the higher is the price. The proprietor has an extensive privy attached to the concern. This collection is fronting the public street; it is enclosed in front by a wall; the height of the wall is about 12 feet, and the dung overtops it; the malarious moisture oozes through the wall, and runs over the pavement. The effluvia all round about this place in summer is horrible. There is a land of houses adjoining, four stories in height, and in the summer each house swarms with myriads of flies; every article of food and drink must be covered, otherwise, if left exposed for a minute, the flies immediately attack it, and it is rendered unfit for use, from the strong taste of the dunghill left by the flies. (Chadwick 1842, p. 119).

The market was certainly at work here, but hardly in a way that reduced exposure to disease.

The market was at work too with regard to the provision of water, in the form of piped water supplied by private water companies, from containers sold by private vendors, or from sales at a local street pump or tap (Hohenberg and Lees 1985). Privately provided piped water was allocated almost wholly to meet the demands of the middle and upper income groups, while those in crowded urban slums might have to walk a quarter of a mile to the one water tap in the neighbourhood (Wohl 1983, pp. 61–3; Briggs 1985, pp. 134–5; Goubert 1989). In these circumstances, it is not surprising that working class families in mid-nineteenth century Burton-on-Trent ‘purchased an average of nine buckets of water a week for a family of five or more for all purposes’ (Wohl 1983, p. 63, emphasis in original). Note that this statistic relates to working class families – the poor would have fared worse.

9 ‘Dogs’-dung ... called “Pure” from its cleansing and purifying properties’ was also valued, and collected by specialised workers. See the description of ‘Pure-finders’ in Mayhew (1851, pp. 306ff).

10 Conditions in American cities were much like those reported in the Chadwick Report. See, for example, the reports by Griscom [1845] 1970, Shattuck [1850] 1948, Smith 1911.
But water supply involves much more than a problem of unequal distribution. As cities grew, and the cost of transportation of human waste to rural areas became prohibitive, carters turned for disposal to the closest stream, pond or river. Water courses in and around large cities were transformed into enormous cesspools. William Budd’s contemporary description of the ‘Great Stench’ arising from the accumulation of sewage in the Thames River in the summer of 1858 is indicative:

For the first time in the history of man, the sewage of nearly three millions of people had been brought to seethe and ferment under a burning sun, in one vast open cloaca [sewer] lying in their midst. . . . Stench so foul, we may well believe, had never before ascended to pollute this lower air. . . . The river steamers lost their accustomed traffic, and travellers, pressed for time, often made a circuit of many miles rather than cross one of the city bridges (as quoted in Winslow 1943, p. 288).

Sewage disposal thus led increasingly to contaminated water supply.\(^\text{11}\) The problem was aggravated by industrial wastes from factories (Cain 1977, pp. 375–6; Hassan 1985). Since pathogenic organisms can exist in water which to the naked eye is pure, not even the wealthy – despite their ability to pay – were assured of protection from this source of infectious disease.

All of this boils down to a simple point – under the conditions of agglomeration arising from nineteenth century economic growth, the market could not be counted on for the provision of pure water in adequate amounts or for the proper disposal of sewage. Rather, market forces were tending to increase exposure to infectious disease.

5.2. Immunisation

On the face of it one might suppose that a newly available vaccine would find a ready market. But this assumes a belief in the efficacy of modern medicine that may not exist. Despite the documented success of smallpox vaccination, a strong anti-vaccinationist movement existed in Great Britain well into the nineteenth century. In sub-Saharan Africa in the 1960s, mothers sometimes hid babies from the national or international teams dispensing smallpox vaccinations (Fenner et al. 1988, Hanlon et al. 1988, cf. also Cutts et al. 1989). In addition, those who have appropriate knowledge may be priced out of the market. There is also a free rider problem – the incen-

\(^{11}\) Attempts by some cities to go further upstream for water did not necessarily solve the problem of contamination. In the United States after impure water came to be recognised as a source of disease, it took about half a century before the belief that water purified itself after traveling six miles was replaced by the view that ‘no river is long enough to purify itself’ (Marcus 1979, p. 192). For an excellent analysis of the interdependent problem of urban sanitation and water supply in the United States, see Cain 1977.
tive for vaccination diminishes as others become immunised. And in the case of the immunisation of infants and children, there is a type of principal-agent issue. The child must rely on the parents’ decision to immunise. But parents may be negligent, or simply not have the time needed for a round, say, of three inoculations of DTP or polio vaccine in a year.

Immunisation also involves a problem like that which arises with hygiene education. One person may opt for the new practice or knowledge, but the failure of others to do so may leave that person at risk. Mention was made earlier of the synergistic relation between smallpox and other diseases. The vaccination of one person may protect her against smallpox, but if those who fail to get vaccinated suffer from damage to their immune systems caused by smallpox, they may expose the person who was vaccinated to greater risk from other diseases such as typhoid or tuberculosis.

These considerations add up to a questionable case for reliance on the market to foster the spread of immunisation. A recent publication of the World Bank puts it more strongly: ‘Had it been left to private markets during the last few decades, it is inconceivable that today some 80 per cent of the world’s children would be immunized against the six major-vaccine-preventable childhood diseases’ (Musgrove 1996, p. 14).

**5.3. Antimicrobials**

Here, at last, one might suppose is an area that can be conceded to the market. To be sure, regarding antimicrobials as well as vaccines, there are issues of quality control and of monitoring claims for effectiveness by private producers. But can’t one rely, generally speaking, on the market as a vehicle for distributing drugs?

The answer appears to be no. There are significant externalities associated with the private distribution of drugs; most importantly, the market fails to take adequate account of the fact that the excessive use of antibiotics fosters the growth of drug-resistant bacteria. This problem quickly came to the fore in developed countries shortly after antimicrobials were introduced, and seriously undercut the high hopes originally held for these drugs (Lappé 1982). The seriousness of the problem is nowhere better demonstrated than in today’s developing countries, where, unlike the developed nations, an uncontrolled free market is typically the primary vehicle of drug distribution. A quotation from a World Bank study is particularly informative, because it comes from a book explicitly devoted to a search for free market solutions:

> The proliferation of modern pharmaceuticals in developing countries can have harmful effects ... [I]n many developing countries medical practitioners do not exercise any control over the use of modern prescription drugs such as antibiotics, as do practitioners in the developed countries. Throughout Latin America, for example, prescription
medications, usually manufactured by multinational pharmaceutical firms, can often be purchased over the counter in pharmacies or shops or from medicine vendors. . . . [I]n some regions of India, indigenous practitioners supply modern medicines on a large scale. . . . In Mysore and the Punjab 80 per cent of the medicines are modern, and 50 per cent of the patients receive penicillin injections, generally from unqualified practitioners supplied by pharmacists . . . [T]he greatest source of hazard [is] the tendency of 'pseudo-indigenous practitioners' to use the most powerful drugs possible, such as chloramphenicol, to obtain quick results (Roth 1987, p. 137, citations in original deleted).

This problem of excessive use of antibiotics under free market conditions is exacerbated by the fact that consumers who are ill are likely to demand the most powerful drugs without regard to their longer term effects.

5.4. Sources of technological change in disease control

The foregoing has been concerned with how well the market allocates resources to the specific techniques responsible for the marked reduction in infectious disease – control of the mode of transmission, immunisation, and antimicrobials. Beyond this, however, there is the question whether market forces were behind the development of these techniques.

Economic explanations of invention typically focus on demand conditions as the source of technological change. Demand may similarly be invoked as the main causal factor in the control of infectious disease. As has been seen, in mid-nineteenth century England the prior slow century-long advance in life expectancy had significantly slowed as a result of rapid urbanisation and industrialisation. Health conditions among the poor in urban centres were increasingly recognized as appalling, and epidemic outbreaks of cholera and typhoid aroused concerns generally. These problems contributed to a growing demand for solutions (Flinn 1965, Brown 1988, Mercer 1990, Cain and Rotella 1994).

But while demand increased in the nineteenth century, such demand was not new. Sickness has been the eternal bane of humanity, but effective solutions – though much in demand – have been hard to come by. As Nathan Rosenberg points out:

Many important categories of human wants have long gone either unsatisfied or very badly catered for in spite of a well-established demand. It is certainly true that the progress made in techniques of navigation in the sixteenth and seventeenth centuries owed much to the great demand for such techniques in those centuries, as many authors have pointed out. But it is also true that a great potential demand existed in the same period for improvements in the healing arts generally, but that no such improvements were forthcoming (Rosenberg 1976, pp. 267–8).
That study of the ‘healing arts’ was far from neglected in the sixteenth and seventeenth centuries is suggested by the fact that at Europe’s leading universities at that time there were more salaried chairs in medicine than in science (Ben-David 1984, p. 52).

Rather than demand, the actual sequence of the solutions that were found suggests that it was supply-side developments that governed advances in the control of infectious disease, specifically, the increased feasibility of invention due to the advance of knowledge (cf. also Easterlin 1995). The most intense demand comes from those who are sick and is for the cure of disease, but the development of cures came last, not first, in the actual sequence of technological developments in the control of infectious disease. The first major breakthrough was in preventing the transmission of disease, and reflects the lesser difficulty with which knowledge of transmission can be obtained vis-à-vis developing a cure. Typically, the mode of transmission of a disease is more amenable to observation than its causes, and the development of a cure must wait upon identification of the pathogen and physiological mechanisms responsible for a particular disease. This is evidenced today in experience with the newest major infectious disease, HIV, where the modes of transmission were quickly identified and led to measures directed toward control well before effective therapies started to appear. Before the nineteenth century, the only major advances in control of fatal infectious diseases (notably bubonic plague and leprosy) were also methods of preventing transmission.

The starting point for the cumulation of systematic knowledge of the transmission of disease was the development of vital statistics, dating from the work of Petty and Graunt in the late seventeenth century. Petty, Graunt, and their followers promoted the collection of data on mortality, cause of death, and associated conditions, prepared tabulations of these data, and performed rudimentary analyses with a view to identifying links between environmental conditions and mortality (Riley 1987). This approach, together with the slow growth of medical knowledge, laid the basis for the breakthrough epidemiological studies in the nineteenth century of the type done by Chadwick, Snow, Budd, Villermé, Shattuck, and other analysts of health conditions in the early industrialising countries. Although much of this work was guided by a mistaken miasmatic theory of disease, its policy conclusions – an attack on filth – had the important effect of reducing human contact with disease-causing pathogens in the environment.

Knowledge of the causes and mechanisms of disease, and their application to the development of systematic immunisation and chemotherapy had to wait upon the emergence of microbiology. This, in turn, depended on advances in instrumentation (especially the microscope), development of laboratory research techniques, and the growth of related disciplines such as chemistry, anatomy, and physiology. It is this sequence in the evolution of basic biomedical knowledge – from epidemiological studies to identification
of causes and mechanisms – that principally explains, I believe, the chronology of advances in the control of major infectious disease, not demand conditions.

Nor was the market responsible for these supply-side advances. The major breakthroughs rarely came from individuals or firms motivated by personal profit.12 Some of the innovations were due to dedicated public servants like Chadwick. Others were made by medical practitioners seeking to understand and ameliorate the everyday problems of disease that they encountered, such as Snow, Budd, and Koch. Still others were due to scientists like Pasteur, Fleming, and Florey, working in research laboratories of universities, research institutes, hospitals and clinics, or governments (cf. Roemer 1991, p. 48; 1993, p. 12). It was the work of individuals like these that was chiefly responsible for the growth of biomedical knowledge and its eventual payoff in the control of fatal infectious disease.

6. Institutional innovation in the control of infectious disease

In what is widely regarded as one of the early classics in health economics, Arrow (1963, p. 947) observed: ‘[W]hen the market fails to achieve an optimal state, society will, to some extent at least, recognise the gap, and non-market social institutions will arise attempting to bridge it.’ The non-market institutions required to bring infectious disease under control are testimony to the accuracy of this generalisation.

I take ‘institutions’ here in North’s (1990) sense of both formal and informal arrangements. And, indeed, both types have been required – informal arrangements in the form of a change in social norms relating to responsibility for disease, and, also, formal establishment of an apparatus for state intervention. These are taken up in succession below.

One of the effects of the sanitation revolution was a gradual transformation in attitudes toward responsibility for disease (Griscom [1845] 1970; Hanlon et al. 1960, p. 446; Flinn 1965, p. 59; Institute of Medicine 1988, ch. 3; Duffy 1992, p. 128. Previously, disease had been attributed to ‘acts of God’ or individual failings, such as sinfulness, lack of moral character, and the like. However, the growth of knowledge regarding modes of transmission of disease made it increasingly clear that the individual might be the victim of forces beyond his or her control, and that these forces were within the purview of social action. As awareness of this possibility grew, so too did support for state intervention in the interest of ‘public’ health (Briggs 1985, 2, p. 150; Kearns 1988).

12 Perhaps the most important exception is the discovery of sulfanomides by G. Domagk at I.G. Farben in the 1930s.
What was lacking, however, was an effective mechanism for intervention. It was in the solution of this problem that the sanitation revolution probably made its greatest contribution. The key institutional innovation was the establishment of a network of local boards of health under the supervision of a central authority (usually a national health agency, although in the United States this function was performed by state health boards), armed with the weapons of inspection, quarantining, closing facilities, and the like. The novelty of the institutional innovation of the public health system is recognised by both contemporaries and historians. In 1890, looking back on the evolution of the public health apparatus, Sir John Simon, ‘the greatest of the Victorian medical officers’ (Wohl 1983, p. 8), was to observe that ‘on the new foundations of Science, a new political superstructure has taken form’ (Simon 1890, p. 463). Writing in the mid-twentieth century, George Rosen, author of the classic history of public health, cites Edwin Chadwick’s chief contribution as his recognition that ‘what was needed was an administrative organ to undertake a preventive program by applying engineering knowledge and techniques in a consistent manner’ (Rosen 1968, p. 167).

Initially the focus of the public health network was on sanitation – establishing pure water supplies, sewage disposal, paving streets, and the like. But the functions changed over time as knowledge and technology advanced. With growing acceptance of the germ theory, a bacteriological view of public health tended to reinforce ‘sanitary science’, and expand the functions of health departments. Bacteriological laboratories became part of the new departments, and research and diagnosis of pathogens became significant functions. Regulation of food and milk supply developed as the role of food handling in the transmission of disease became recognised. Recognition grew of the need for housing standards, building regulations, and appropriate enforcement authorities. The production and distribution of vaccines became important too. And gradually some of the original activities of health departments were spun off to other municipal agencies, such as responsibility for water supply, waste removal, and ‘nuisances’, although oversight and regulation functions continued.13

For households, the domestic hygiene counterpart of the new sanitary science centred initially on ventilation, disinfection, plumbing, water purification, isolation of the sick, and general cleanliness. Since the new knowledge was not proprietary, the market could not be relied on to disseminate it. Nor were there competitive profit-making pressures on households analogous to those fostering the adoption of new production techniques by

---

13 The association between the growth of knowledge regarding disease and the expansion of government regulatory and educational activities from the latter part of the nineteenth century onward is apparent in a number of articles written from the perspective of the early 1930s in the first edition of the Encyclopaedia of the Social Sciences. See, e.g., the articles on food and drug regulations, building regulations, inspection, health education, sanitation, water supply, milk supply, housing, and slums.
The market and the modern history of mortality 285

profit-making firms. At first, the new knowledge was promoted especially by women reformers through voluntary organisations. But public health agencies gradually assumed an increasing role, and voluntary domestic hygiene was supplemented by compulsory quarantine and disinfection. As knowledge grew, education expanded to encompass food handling and infant and child care, and health programmes were introduced into the schools. Because women were principally responsible for household care and child-rearing, these educational efforts were especially directed towards women. Thus, in contrast to economic growth, female, rather than male, education has played a central role in the improvement of life expectancy.

One indication of the success of these educational efforts is the changing pattern of consumer demand. Mokyr and Stein (1997) point out that in England soap consumption rose sharply in the late nineteenth century, despite a rising price (cf. also Wohl 1983, p. 71). In regard to the late nineteenth century United States, Tomes (1990, p. 531) reasons that ‘the rush to develop and to patent sewer traps, toilet designs, window ventilators and water filtration systems ... suggests that entrepreneurs found a lucrative market among [middle and upper class] householders anxious to safeguard their families against infection’.

The public health system was not established without a struggle. From its inception in the sanitation revolution, the public health movement encountered serious opposition, because of the proposed expansion of the government’s role in the economy (Brown 1988, Szreter 1997). Specific proposals were fiercely debated in local and national political arenas. The backbone of the opposition was provided by those whose vested interests were threatened – landlords, water companies, proprietors of refuse heaps and dung hills, burial concerns, slaughterhouses, and the like (Briggs 1985, 2, ch. 7; for the United States, cf. Wells 1995; for Germany, Evans 1987, ch. 2). They appealed to the preservation of civil liberties, and sought to debunk the new knowledge cited by the public health advocates, a strategy reminiscent of today’s response of the tobacco industry to evidence of the adverse health effects of smoking. Public health reformers reacted strenuously, asserting that ‘laissez-faire is the paltriest of all philosophies, in sanitation – as in everything else’ (Tomes 1998, p. 129). Thus, the sanitation revolution and its sequela precipitated an extended clash of ideologies between advocates of laissez-faire and proponents of state intervention in the interest of ‘public’ health.

The institutional impact of advancing biomedical knowledge went beyond the official public health system. Voluntary associations arose, usually dedicated to a specific purpose, such as education in regard to infant care, or the diffusion of knowledge about tuberculosis. These voluntary organisations served a useful purpose in supplementing the governmental system and pointing to new needs for action, but they were, for the most part, relatively short-lived, and it was the governmental system that formed the backbone of the new institutional structure dedicated to the promotion
of public health. More importantly, as the germ theory became accepted, it revolutionised the training of doctors and nurses, and gave birth to what we know as the modern hospital (Abel-Smith 1960, 1964; Vogel 1980; Rosenberg 1987). Also with the development of ‘community’ medicine, new professional associations arose, such as the American Public Health Association, along with schools of public health and a specialised professional literature (Rosen 1958, pp. 516–25; Duffy 1992, p. 253). Thus, the new knowledge and technology of disease control spawned a growing institutional infrastructure needed for its implementation. This infrastructure is itself testimony to the failure of the market adequately to rise to the task of controlling infectious disease.

7. Conclusion

Let me summarise some of the impressions from this look at the modern history of mortality. The improvement of life expectancy, like economic growth, has been based on a new technology involving new institutional, capital, and labour requirements. But for life expectancy, the nature of the new technology and associated requirements is quite different from those for economic growth. The technology comprises new methods of controlling major infectious disease. At the core of the institutional requirements is a public health network. The capital requirements involve chiefly new public expenditures, and the labour requirements are for the bearers of the new technology – specialised personnel in the fields of public health and medicine, and homemakers educated in personal hygiene and household sanitation.

The point of departure for understanding the great improvement in life expectancy in the last century and a half must be the abysmal state of knowledge that prevailed at the start of this period, and still exists today in many places. The causes of the major infectious diseases were not known, and almost nothing was known about the way in which these diseases are transmitted. In the absence of valid knowledge of the ‘health production function’, resources allocated to the prevention or cure of disease were, with a few exceptions, ineffective.

As economic growth began to raise living levels in Western Europe after 1800, one might have expected resistance to disease to have grown and life expectancy to have been raised, even though health knowledge remained low and health practices of questionable value. But this reasoning regarding the effect of economic growth on life expectancy is seriously incomplete, for it fails to take account of the urbanisation requirements of the new technology on which modern economic growth is based. The rapidly rising concentration of population in urban centres substantially increased exposure to disease, and largely vitiated any effect of increased resistance.

Only with the growth, first, of epidemiological and, then, bacterial knowl-
knowledge did effective techniques emerge for controlling infectious disease. These techniques focused primarily on the prevention of the spread of disease – first via controlling the mode of transmission, and subsequently via immunisation. It is these methods of prevention that have been chiefly responsible for the great improvement in life expectancy. In the last half century the advance of knowledge has also added methods of curing disease to the arsenal available to fight infectious disease, particularly with the development of antibiotics, but the great bulk of the reduction in infectious disease has been accomplished largely by preventive methods.

The control of infectious disease involves serious issues of market failure – information failures, externalities, public goods, principal-agent problems, and so forth. The market cannot be counted on for such things as the provision of pure water and milk, the proper disposal of sewage, control of pests such as mosquitos and rats, the supply of uncontaminated food and other manufactured products, immunisation of children and adults against major infectious diseases, and the dissemination of new knowledge regarding personal hygiene, infant and child care, food handling and preparation, care of the sick, and the like. There is also a serious market failure problem with regard to the distribution of antimicrobials because of externalities associated with the development of disease-resistant bacteria.

The title of this article posed the question, how beneficent is the market? The ubiquity of market failure in the control of major infectious disease supplies the answer: if improvement of life expectancy is one’s concern, the market cannot do the job. Implementation of the new techniques of disease control has required the development of new institutions, centering on the public health system. The functions of this system have included in varying degrees health education, regulation, compulsion, and the financing or direct provision of services. Action by households, and particularly women responsible for home care, has also been essential in the control of infectious disease. However, such action is premised on the implementation of new knowledge, and the government has typically played a key role in the dissemination of this knowledge to the public.

The establishment of a public health system has also required acceptance of social responsibility for the control of major infectious disease. This shift in norms came about as the advance of biomedical knowledge increasingly pointed to factors beyond individual control as the primary source of disease, in much the same way that progress in economics in the twentieth century has led to increased acceptance of social responsibility for unemployment and inflation. In time, intervention in the interests of public health came to be seen as positive and necessary, not simply as a residual function, doing ‘what the market can’t or won’t do’ (Institute of Medicine, 1988, p. 46).

None of this is to say that the situation with regard to public sector intervention for the control of infectious disease has been, or is, optimal. But the
assumption that the market, in solving the problem of economic growth, will solve that of human development is belied by the lessons of experience. Rather than a story of the success of free market institutions, the history of mortality is testimony to the critical need for collective action.

Acknowledgements

I have benefited greatly from discussion with Moses Abramovitz, Nauro B. Campos, Dean T. Jamison, James Robinson, and participants in seminar presentations at Lund University, Northwestern University, the University of Illinois-Urbana, the University of North Carolina, the University of Southern California, and the World Bank, especially Lee J. Alston, Tommy Bengtsson, William Easterly, Robert E. Gallman, Joel Mokyr, Larry Neal, and Jeffrey B. Nugent. Nancy Easterlin supplied the epigraph. Helpful research assistance was provided by Nauro B. Campos, Christine M. Schaeffer, and Donna H. Ebata. I am grateful to the University of Southern California for financial support.

References


SZRETER, S. (1988). The importance of social intervention in Britain’s mortality


