Journal of Economic Literature  
Vol. XLIV (March 2006), pp. 106–114

The Great Escape: A Review of Robert Fogel’s *The Escape from Hunger and Premature Death, 1700–2100*

Angus Deaton*

*Deaton: Research Program in Development Studies and Center for Health and Wellbeing, Woodrow Wilson School, Princeton University. I am grateful to Anne Case and to Adriana Lleras-Muney for helpful discussions. I gratefully acknowledge financial support from the National Institutes of Health through the NBER, Grant Number F01 AG05942-14.*

1. Introduction

In recent years, economists have changed the way that they think about health. Amartya K. Sen (1999) has successfully pressed the importance of recognizing aspects of well-being beyond real income and argued that health has the primary claim on our attention. Jeffrey Sachs, and the Commission on Macroeconomics and Health (2001) which he led, argue that good health is necessary (and perhaps even sufficient) for economic growth in the poorest countries of the world. And economists have begun to rise to the challenges posed by social epidemiologists who argue that it is socioeconomic status, including income, which is the primary determinant of health, not health care. Robert W. Fogel has been thinking about these issues for perhaps longer than anyone. His accumulated wisdom and the historical experience have much to bring to the current debates. In *The Escape from Hunger and Premature Death*, he restates his grand historical account of the synergistic improvement of living standards, morbidity, and mortality since the eighteenth century. He also develops the implications for understanding health in poor countries now, and for thinking about the future of health and living standards for those of us who are fortunate enough to live in countries that are already rich, whose inhabitants are tall and strong, whose life expectancy at birth is ten years or more in excess of three score and ten, and half of whose recently born children may well still be alive a century from now.
2. Ascending Mount Waaler

The analytical device that underpins much of Fogel’s book is the “Waaler” curve. Hans T. Waaler is a Norwegian economist turned epidemiologist who, in 1984, published a monograph in which he used a sample of 1.8 million people to investigate the links between height, weight, and mortality. What I call Mount Waaler is the three-dimensional plot of life expectancy (the height of Mt. Waaler) against height and weight (longitude and latitude). One useful cross-sectional cut through the mountain is to make a two-way plot between life expectancy and the body-mass index (BMI), defined as weight in kilos divided by the square of height in meters. A six foot man who weights 200 pounds has a BMI of 27.1; at 225 pounds, his BMI would be 30.5, just over the threshold of obesity, which by convention begins when BMI crosses 30. Among those who are underweight, with a BMI below about 20, mortality risk decreases with BMI; it is then more or less flat until BMI reaches 30, at which point it rises again. But BMI does not give us a complete picture of Mt. Waaler because, at any given level of the BMI, taller people live longer, at least up to a point. You reach the top of Mt. Waaler when you are 6’ 3” tall and weigh about 190 pounds, at least if you are a Norwegian man.

Provided that Mt. Waaler does not change its shape or height over time—and Fogel provides some evidence for constancy by comparing Waaler’s results with his own data from Union Army Veterans from the American Civil War—then the history of human physiology and, to an extent, the history of economic growth, can be usefully thought of as a(n uneven) ascent of Mt. Waaler. The escape from hunger and premature death is the movement up Mt. Waaler from the southwest, where the inhabitants are short, thin, and weak, toward the northeast, where the inhabitants are tall and powerful. We can also mark the position of the typical Indian or Chinese on Mt. Waaler and compare it to the typical American now or the typical American in 1800. And we can think about current trends in weights and heights, and where they are leading to on the mountain.

How did our ancestors move up the mountain over the last two centuries? By eating more, which required more food production, which was possible in turn because they were bigger and stronger, which was because they (or their parents) ate more, and so on. The escape from hunger and premature death was the escape from a nutritional trap where we could not work to produce food because we were too weak, and we were too weak because we could not work to produce food to make us strong. This synergistic improvement of health and living standards is referred to by Fogel and his collaborators (see particularly Fogel and Dora L. Costa 1997) by the inelegant term *techno-physio evolution*, “a synergism between technological and physiological improvements that is biological, but not genetic, rapid, culturally transmitted, and not necessarily stable” (p. 20).

The story has many interesting and not obvious implications. Inventing a Wellsian time machine to take us all back to eighteenth-century England would be as good for our health as transporting us to the moon without spacesuits. Our bodies are simply too large to survive on the average food supply then available. When the food supply is low, people cannot be large because large bodies use up too many calories simply in resting and maintenance, leaving nothing for work. Our ancestors could manage to survive and procreate because they were much smaller than we are. But, because they were small and subject to the laws of Mt. Waaler, they lived shorter lives than we do. Both they and their lives, if not nasty and brutish, were certainly short. If we were so transported, of course, most of us would adapt quite quickly to meager rations by becoming thinner, but only the short would be able to survive, so that the average height of the population would shrink, though
more slowly than average weight. The reverse process is how our ancestors made their great escape from hunger and death, and turned themselves into the large, long-lived animals that we are today. According to Fogel, the fact that we are bigger and stronger can account for around half of the growth in national income in Britain since 1790 (p. 34). It also accounts for not only the increase in longevity since 1790—the movement up Mt. Waaler—but also much of the reduction in morbidity in the United States of the twentieth century, including that from noninfectious diseases, including musculoskeletal disease, digestive disease, heart disease, and cancer.

The process of getting bigger and stronger is still running. A recent compilation of evidence from ten western European countries (A. E. Cavelaars et al. 2000) showed that, in each decade, men became taller by 0.6 inches and women by 0.3 inches. This is true in the tallest countries, Norway, Sweden, Denmark, and The Netherlands, as well as in the shortest, France, Italy, and Spain, so that the gap between the top group and the bottom group, 3.5 inches for men and 2.7 inches for women, has not changed over time. The same is not true for the United States, long the tallest country in the world, and still so at the end of World War II. But height has been stagnating in the United States for a decade, and Americans are now shorter on average than many Europeans, including not only the very tall Dutch and Scandinavians, but even the citizens of the former East Germany (see John Komlos and Marieluise Baur 2004). While Americans are not expanding upwards, they continue to expand outwards, and the average American, like the average Briton, is now heavier than the weight that would minimize mortality risk given average height. While we are still some way from the top of Mt. Waaler, we are not at the highest point given our heights. Consistent with the laws of the mountain, the inhabitants of the United States also have shorter life spans than those of many other taller countries, again including Sweden, Norway, The Netherlands, and Germany. Of course, Mt. Waaler is certainly not the whole story; life expectancy in the United States is also less than in Japan, Singapore, Hong Kong, Israel, Greece, Costa Rica, Guam, and Puerto Rico.

Fogel's argument that nutrition was a key element of the historical great escape is surely right, as is the proposition that poor nutrition remains an impediment to health in much of the world today. Much less obvious is the idea that nutritional deficits are an important part of the health story in the rich world today. Yet there is a good deal of evidence, even and in some cases particularly—in populations whose most obvious nutrition-related problem is obesity and over nutrition. One key line of work here is associated with David J. P. Barker and his group at the University of Southampton in England. The “Barker hypothesis,” sometimes referred to as the “womb with a view” hypothesis, is that events in the womb have long-lasting effects on health throughout life, and perhaps particularly for health outcomes that express themselves in late life. Nutritional insults in utero, which prevent the fetus developing its full potential, cause the selective abandonment of function, with evolution disfavoring those features whose primary function is to prevent disease in late life beyond the normal reproductive span. Although the Barker hypothesis remains deeply controversial, there is fascinating evidence that supports it. Gabriele Doblhammer and James W. Vaupel (2001) and Doblhammer (2002) have shown that life expectancy at 50 varies seasonally with month of birth. In the northern hemisphere, 50 year olds who were born in October and November, so that their mothers had access to cheap and plentiful fresh fruits, vegetables, and eggs through most of their pregnancy, can expect to live about three-quarters of a year longer than those who were born in the spring. In the southern hemisphere, the same seasonal pattern
occurs, although shifted by six months, except that those who were born in the Northern hemisphere but died in the South (European immigrants to Australia) display the Northern pattern. There is other evidence, for example from the Dutch famine of 1943, that nutritional deficits in pregnancy have long-term effects on obesity, with deficits in the first trimester predicting later adiposity, and deficits in the third trimester inhibiting it. And indeed, explosions of obesity and associated diseases (adult-onset diabetes, heart disease, and so on) often come close on the heels of a loosening of nutritional constraints, when those whose parents were undernourished, who were themselves undernourished in utero, move into an environment in which food is plentiful and heavy manual work is no longer required. For example, in the black township of Khayelitsha near Cape Town in South Africa, more than a half of the adult women have BMIs above 30 (Anne Case and Angus Deaton 2005).

Today, movements in life expectancy in rich countries are driven by trends in chronic disease among those over 50—infant and maternal mortality rates are no longer important—so that, if the “womb with a view” stories are even partly correct, it is not nutrition today that is important for population health today, but nutrition before 1950. Indeed, the consequences of the nutritional improvements of the first half of the last century, not only in quantity of food but in the increased availability of fresh produce throughout the year and in a better understanding of nutritional requirements, have still to deliver their full payoff in reductions in morbidity and mortality. Middle-aged and elderly Chinese who are currently alive are the survivors of a period in which nutritional and other insults killed 20 percent of their contemporaries in their first year of life. Although they survived, many of them experienced the same insults as those who died, albeit in less severe form, and so will suffer from a particularly high burden of chronic disease in later life (p. 91).

Moving from poor to rich, the rising life expectancy and falling morbidity of the middle-aged and elderly in the United States now owes much to the fact that, compared with 1910, infant mortality in 1930 had been cut by a half, and by 1940, by two-thirds. The children who survived their first year in the 1930s, and who are seventy now, have a lower burden of chronic disease than did their parents, whose early health environment was much more dangerous. This pattern of declining morbidity from chronic disease is a challenge to the standard account of the epidemiological transition, according to which the era of infectious disease, largely affecting children, gives way to an era of chronic disease, largely affecting the elderly. According to Fogel, much of whose evidence comes from comparisons of Union army veterans with their modern counterparts, chronic diseases have declined alongside infectious disease, although the demise of the latter has made the former more important in relative, although not absolute, terms.

3. Disease, the Germ Theory, Income, and Public Health

Is Fogel’s account of history right? His research program, as well as much current thinking about the social and medical determinants of health, traces back to the work of Thomas McKeown (1976, 1979), who argued that economic growth and better nutrition were the fundamental causes of the remarkable improvement in population health since the eighteenth century. McKeown’s main target was the idea that population health had improved through better medical treatments, particularly through drugs and vaccines, and he drew a series of famous graphs showing the reduction in mortality from various diseases over time (tuberculosis, typhus, scarlet fever, dysentery, cholera, smallpox, etc.), and pinpointing the dates of important prophylactic innovations, such as sulfonamide and penicillin against infectious diseases, or
streptomycin for tuberculosis. These graphs provide clear evidence that, with the possible exception of vaccination against smallpox, innovations in treatment accounted for little, if any, of the escape from disease; declines were well underway before the relevant medical innovations were discovered and continued at much the same pace afterwards. McKeown then turned his guns toward public health measures and argued that, they too, had little effect. By elimination, McKeown concluded that the reduction in mortality came from the general improvement in living standards, particularly the associated increase in nutrition.

Fogel's research program begins where McKeown's ended, replacing indirect, negative evidence by direct, positive evidence for the importance of nutrition, documenting levels of calorie availability and their effects on health, with Mt. Waaler the single most important bridge between the two.

Subsequent research has successfully challenged McKeown's conclusions, at least in part. While the arguments about the limited role of new medical treatments are still accepted—for the past if not for the present—the arguments about public health have been debunked, see the important paper by Simon Szereter (1988), as well as Samuel H. Preston's (1996) succinct overview. Today, it is clear that public health measures, particularly the provision of clean water and better sanitation (for example, to stop sewage being discharged into drinking supplies), were the fundamental forces for mortality reduction during the century from 1850 to 1950, see for example David M. Cutler and Grant Miller (2005) on cities in the United States. Although progress began under the sanitarians, John Duffy (1990), whose (incorrect) understanding of disease (if it smelled bad, it caused disease) was essentially unchanged from medieval times, the process was reinforced and speeded up once the germ theory of disease had supplied a more useful scientific basis for policy. Even once the scientific understanding was there, starting about 1870, it took many years for the appropriate public health measures and behavioral changes to diffuse into the population, and to make our health what it was by mid-century, see Nancy Tomes (1998), Richard A. Easterlin (2004, chapter 7), and Joel Mokyr (2002, chapter 5). But if there is one single factor that was primarily responsible for the great escape, it was the discovery of the germ theory of disease and its implementation through public health measures, first in northwestern Europe and North America, later (and more quickly) in southern and southwestern Europe, and most recently (and most rapidly of all) in the third world since World War II (Davidson R. Gwatkin 1980, Preston 1975, 1980).

There is good reason to doubt the ability of a nutritional trap, by itself, to hold its victims for long. In currently poor countries, for which nutritional wage stories were first proposed and fully worked out in papers by James Mirrlees and Joseph Stiglitz in the 1970s, subsequent empirical work has consistently failed to provide support. In modern economies, even very poor ones, the trap cannot be binding; the 2,000 or so calories that can provide the means key to escape can be bought with only a fraction of the daily wage (Shankar Subramanian and Deaton 1996). Anyone who has tried to lose weight can ruefully attest to the number of miles of running or hours on an exercise machine that are required to use up a few hundred calories, and the converse also holds, a single kilo of bread will fuel all but the most energetic day's work. People whose life chances were crippled by the lack of a few hundred calories, and who understood the nature of their problem, would have devoted every hour of every day to the search for calories, cultivating their own crops or accepting even the lowest paying work. Were real wages really so low, the prices of staples so high, and the opportunities so limited, that people were trapped in a way that is not true today in economies whose per capita incomes are as low or lower than those of Europe at the outset of the industrial revolution?
Nutritional traps are much easier to understand once disease is given its proper place in the story. Disease interacts with nutrition, and each reinforces the other. Malnutrition compromises the immune system, so that people who do not have enough to eat are more likely to succumb to infectious disease. At the same time, disease prevents the absorption of nutrients so that, even when food is obtainable—through own cultivation or in exchange for work—it cannot be turned into nutrition. A child who is suffering from acute diarrhea cannot be cured by giving her more food.

What difference does the primacy of the germ theory and of public health make to Fogel's story? Fogel fully understands the importance of cleaning up the water supply and of the associated reductions of disease, as well as the interaction of diseases and nutrition, and he makes these points repeatedly, not only in *The Escape* but also in his magisterial essay on mortality in the *Handbook of Family Economics*, which was written around the same time as the lectures that became *The Escape* and should be read alongside it or, for that matter, by anyone who would like to read one of the great papers in the mortality literature. Yet Fogel, while making all of the right points, always eventually turns away from disease and from public health toward nutrition. In one sense, this is just fine. Given the power of the interaction between disease and nutrition, it makes little sense to fight over which was the most important or to try to parse the total into shares that add up to 100 percent. And even if we choose to accept the primacy of the germ theory and of its systematic implementation, as I do, then it was the removal of human waste from the drinking water (for example) that permitted nutrition to do its work on the human body, making us all bigger and stronger, and enhancing the efficiency of labor, particularly in manual occupations. So the improvement in nutrition and its effects are real enough, and important, precisely as Fogel argues.

But there is another version of the story that is harder to defend. The synergism between economic growth and the growth of the size and the durability of the human body can turn into an overemphasis on links between economic growth and health, and an underemphasis on the role of disease and its prevention. Fogel is rarely guilty of such overemphasis, though he is at some pains to emphasize the close tracking of health and income, provided the latter is appropriately measured. While the synergism is surely there, it is far from automatic. In Britain, the United States, and much of Europe, there were periods in the nineteenth century when urbanization ran ahead of the rate of public health provision and population health deteriorated during periods of rapid economic growth. In a country by country examination of mortality decline, whether in Europe in the nineteenth century (Easterlin 1996; 2004, chapter 7) or in the third world since World War II (William Easterly 1999), we sometimes find that economic growth is correlated with improvements in longevity, but just as often not. While it is hard to imagine the absence of a correlation between health and income in the longest of long runs, the relationship can vanish for substantial periods of time. This is important, not just for getting the history straight, but for health policy throughout the world now. If economic growth reliably improved nutrition in poor countries now, and if nutrition is the primary barrier to health, then we should worry about growth, and let health look after itself. But if causation runs the other way, as Jeff Sachs argues, or if growth by itself is no guarantee of health improvement, then some sort of public action, whether through public health or provision of health systems, is required to turn economic growth into improvements in health. Economic growth frequently needs help to guarantee an improvement in population health. There are other, and sometimes faster, ways of climbing Mt. Waaler than to wait for income to make people taller and stronger.
4. Back to the Future

Fogel devotes two chapters and an appendix of The Escape to his thoughts about the future, in many cases referring back to the material in his last book The Fourth Great Awakening. He is not afraid to make bold predictions about the lifespan, the share of GDP devoted to healthcare, and leisure, all of which will be longer or bigger (and better) than they are today. Like several other commentators, notably Jim Oeppen and Vaupel (2002), he endorses the view that life expectancy will continue to increase as it has done in the past, by at least two to three years every decade. Oeppen and Vaupel calculate that over the last 160 years (which is the span over which we have data) “record life expectancy,” the highest life expectancy in the world, has increased by about 0.24 years per year for women and 0.22 years per year for men. So that female life expectancy in the United States will be around a hundred years by 2070. Since this is a period estimate, and since mortality is projected to fall further, the first cohorts who can expect to live to be centenarians will be born much earlier, and among the more privileged groups, may already have been born (Vaupel and A. E. Gowan 1986, Fogel 2005). Fogel notes that these forecasts are much more optimistic than those currently being used by the Social Security Administration, and therefore more pessimistic for the solvency of social security and Medicare, at least if the increased number of years at risk are not offset by later onset and decreased severity of chronic disease.

Fogel reminds us that optimism is the right word. That longer life gives us opportunities that we used not to have, more time to get to know our grandchildren and great-grandchildren, more time for second careers, for voluntary activities, and all the other components of a good life. But he is also entirely optimistic about the prospect that we will spend a much larger share of GDP on healthcare. He argues that as we grow richer we can afford and will want to spend more money on improving both the quality and length of our lives. “The increasing share of global income spent on healthcare expenditures is not a calamity; it is a sign of the remarkable economic and social progress of our age” (p. 107) and “Public policy should not be aimed at suppressing the demand for health care. Expenditures on health care are driven by demand, which is spurred by income and by advances in biotech that make health interventions increasingly effective” (p. 95). Fogel also argues that large expenditures on health are a boon to the economy and that the healthcare industry will play the role of leading industry, just as in the past did railroads, automobiles, or computers. If there are difficulties with the funding of healthcare, or with social security and pension schemes that were predicated on a lower level of life expectancy, the problems lie only with “clumsy system of financing” (p. 77), in particular the absence of private accounts for social security, and of health saving accounts for health.

While Fogel is honest enough to insist that contributions to saving accounts would have to be compulsory, it is unclear that the third of income that would be set aside would be as easily and as willingly met as he thinks, or that a third would be sufficient if expensive new innovations in health care continue to appear at their current rate. Nor is Fogel sensitive to the equity consequences of a system of private accounts, particularly health savings accounts. It is hard to think of a more effective way of increasing the positive correlation between health and wealth than health saving accounts that guarantee that people who suffer from chronic illness will enter retirement, not only with a poorer health histories and lower expectations of remaining life, but also with retirement saving accounts that are depleted (potentially even below zero) compared with those whose health histories have been more fortunate.
Fogel’s enthusiasm for increased spending on health care is doubly curious because, like McKeown before him, he credits medicine with little of the reduction in mortality, either in the past, for which his argument is plausible, or in the present, for which it is not. He writes “it is likely that past public health reform, improvements in nutrition and other living standards, and the democratization of education, have done much more to increase longevity than has clinical medicine” (p. 102), and goes on a page later to write “The main thing that physicians do is to make life more bearable: reduce morbidity and tell people how to take care of themselves.” It is strange to be so sanguine at the prospect of spending a third or more of GDP on something that brings so limited a benefit, and if we need a leading industry, perhaps we could select one whose output contributes more to our wellbeing?

My own view is that, at least since 1970, medicine has been much more effective and important in prolonging life than Fogel gives it credit for. At the same time, and perhaps paradoxically, much of medical expenditure is driven by suppliers, not by patient demand, and carries little or no benefit, “flat of the curve medicine,” or worse, iatrogenic medicine, medicine that actually hurts people. Cutler (2004) provides evidence for both the good and the bad of modern medicine. The rapid declines in mortality after 1970, particularly mortality from cardiovascular disease, are pretty much what would be expected from new procedures and new drugs, given the randomized controlled trials prior to their introduction. After successful trials for diuretics in the late 1960s, physicians began to prescribe them for reducing hypertension, which quickly started a downward trend in mortality, which was maintained by a whole succession of subsequent new drugs and new treatments. Cutler also argues that if we use standard numbers for the value of additional years of life, the total cost of the health care system is more than justified by its results, a proposition about averages that is entirely consistent with the existence of enormous waste, flat of the curve medicine and unnecessary supplier-induced demand. Some of the most compelling evidence for waste comes from the Dartmouth Atlas and the work of John Wennberg, see for example the October 7th, 2004 issue of Health Affairs, which documents the extraordinary geographical variation in treatments and expenditures that can neither be explained by patient needs nor be shown to correspond to patient outcomes.

Fogel’s position, in contrast to these views, appears to be an amalgam of McKeown (living standards and nutrition are important, physician-supplied medicine not) and of Chicago economics (if the market supplies all this health care, then that must be a good thing, even if it does very little). I do not believe that either of these positions is useful for thinking about medicine or healthcare expenditures in any of the rich countries of the world. Yet this (important) area of disagreement does little to diminish my admiration for Fogel’s work, nor the eagerness with which I turn to his work. There is no other writer on these matters whose scholarship is more comprehensive, nor from whom so much can be learned. So that, even if it is impossible to agree with all of the conclusions that he draws from the immense body of his research, traveling through the literature with Fogel as guide is an experience that is to be recommended with the greatest enthusiasm.

REFERENCES


