1

EDITOR'S INTRODUCTION: THE NEW ECONOMIC HISTORY AND THE INDUSTRIAL REVOLUTION[1]

Joel Mokyr

The Industrial Revolution -- a Useful Abstraction

In the past years, there have been more and more voices that claim, to rephrase Coleman (1983), that the Industrial Revolution is "a concept too many."[2] The feeling is that the term is either too vague to be of any use at all or that it produces false connotations of abrupt change comparable in its suddenness to the French Revolution. The main intellectual motive for this revision has been the growing (though not universally shared) consensus that economic growth in the early stages of the British Industrial Revolution was slower than had hitherto been supposed. The idea of the Industrial Revolution, however, predates its identification with economic growth by many decades. The revision of national income statistics should therefore not, in itself, be enough to abandon the concept. Yet revisionist social historians have found in those revisions the support to state categorically that "English society before 1832 did not experience an industrial revolution let alone an Industrial Revolution. . . . [Its] causes have been so difficult to agree on because there was no 'Industrial Revolution,' historians have been chasing a shadow" (Jonathan Clark, 1986, pp. 39, 66). Wallerstein (1989, p. 30) suggests amazingly that "technological revolutions occurred in the period 1550-1750, and after 1850, but precisely *not* in the period 1750-1850." Cameron (1990, p. 563) phrases it even more vituperatively: "Was there an industrial revolution? The absurdity of the question is not that it is taken seriously but that the term is taken seriously . . . by scholars who should know better."

The important point to keep in mind is, of course, that from a purely ontological point of view, the British Industrial Revolution did not "happen." What took place was a series of events, in a certain span of time, in known localities, which subsequent historians found convenient to bless with a name. The argument whether the Industrial Revolution is a useful concept is therefore merely one about the efficiency of discourse: Does the term communicate? Do most people with whom we want to converse (colleagues, students, book purchasers) know by approximation what we mean when we use the term? And can we suggest a better term to replace it in our conversations? T. S. Ashton wrote in 1948 that the term was so widely used that it would be pedantic to offer a substitute (1948, p. 4; see also Crafts, 1985a, p. 68). Nothing has been learned since then to warrant changing that conclusion. Continuity or discontinuity, as McCloskey (1987) notes, are rhetorical devices. There is no "test" that we can apply: National income and aggregate consumption grew gradually; patents and cotton output grew much faster. Which one "measures" the Industrial Revolution?

Given this background, the sometimes strident voices calling for the banning of the word from our textbooks and journals seem off the mark and, to judge from the writings of scholars in the 1990s, have little influence. Economic historians, like all scholars, need certain terms and concepts with which they can conduct their discourse, even if arguments about the *precise* definitions of these concepts continue. But it is hard to argue that the concept should be abandoned for the simple reason that scholars feel that it communicates and insist

---

[1] This essay is a completely revised and largely rewritten version of my introduction to an earlier collection (Mokyr, 1985a). I am indebted to Gregory Clark, Stanley Engerman, C. Knick Harley, David Landes and Rick Szostak for comments on an earlier version.

[2] Among those, see especially E.L. Jones (1988, pp. 13-27); Clive Lee (1986, pp. 21-22).

on using it. In the years since the first edition of this book appeared, a number of important books and articles whose titles include the term *Industrial Revolution* have appeared, which demonstrates that their authors believe that the Industrial Revolution means something to their readers.[3]

To be sure, arguments about what exactly changed, when it started, when it ended, and where to place the emphasis keep raging. Such scholarly debate about the exact content of a central concept is common -- think of the arguments among biologists about the concept of species. Yet this is insufficient cause to abandon the term altogether: One might as well abandon such concepts as the Reformation or Imperialism.

How revolutionary was the Industrial Revolution? Compared to political revolutions, like the American and French revolutions that were contemporaneous with it, it was rather drawn-out, its dates usually set between 1760 and 1830 following Ashton (1948). To be sure, it was punctuated by some periods of feverish activity such as the year 1769, the *annus mirabilis* as Donald Cardwell (1972) called it, in which both James Watt's separate condenser and Richard Arkwright's water frame were patented. But, on the whole, economic changes, even economic revolutions, do not have their Bastille Days or their Lenins. Economic change is rarely dramatic, sudden, or heroic. Consequently, some scholars have found the revolutionary aspects difficult to stomach. John Clapham and Herbert Heaton, the doyens of economic history in the 1930s and 1940s, shunned the term *Industrial Revolution* altogether. In contrast, historians in the 1960s wrote of "Great Discontinuities" (Hartwell, 1971b) and "take-offs" (Rostow, 1960). Yet gradualism remained strong. Hughes (1970, p. 45) said it well when he wrote that anything that lasts so long is hard to think of as abrupt and added that "we cannot think of the events of the past seventy years as sudden. Seventy British years [in the period 1760-1830] passed no more rapidly."

There is merit to this argument, but not enough to abandon the terminology. Revolutions do suppose an acceleration of the rate of change, but how much does the rate have to change in order for it to qualify? Seventy years is a long period, but the changes that occurred in Britain between 1760 and 1830 dwarfed in virtually every respect the changes that had occurred in the previous seventy years.[4] The annual rate of change of practically any economic variable one chooses is far higher between 1760 and 1830 than in any period since the Black Death. The key concept is an increase in the rate of change, not the occurrence of change itself. The cartoon story of a preindustrial static society before 1750 with fixed technology, no capital accumulation, little or no labor mobility, and a population hemmed in by Malthusian boundaries is no longer taken seriously. Jones (1988) has stressed this point more than anyone else. At the same time Jones points out that before 1750 periods of growth were followed by retrenchment and stagnation. The Industrial Revolution was "revolutionary" because the technological progress it witnessed and the subsequent transformation of the economy were not ephemeral events and moved society to a *permanent* different economic trajectory. Moreover, it seems too much to demand that an event qualify as a revolution only if it follows a period of total stasis -- most political revolutions cannot meet this standard either. Furthermore, revolutions are measured by the profundity and longevity of their effects. In this regard, what happened in Britain after 1760 is beyond serious doubt. The effects of the Industrial Revolution were so profound that, as Paul Mantoux (1928, p. 25) notes, few political revolutions had such far-reaching consequences.

One of the more perplexing phenomena is that contemporaries seemingly were unaware of the Industrial Revolution. A number of scholars have commented on the notable absence of references to anything as dramatic in the writing of political economists and novelists writing in the years before 1830 (Cameron, 1994; cf. North, 1981, p. 160, Adams, 1996, p. 106, and McCloskey, 1994, p. 243). From this it is inferred, somewhat rashly, that contemporaries were unaware that they were living during an Industrial Revolution and from this it is further

---

[3] For instance Allen (1994); Crafts (1994; 1995a, 1995b, 1995c); Crafts and Mills (1994); Easterlin (1995); Engerman (1994b); Hawke (1993); Horrell (1995a); Fisher (1992); Huck (1995); Jackson (1994); Solar (1995); Goldstone (1996); Meignen (1996); Neal (1994); Nicholas and Oxley (1993, 1994); Snooks (1994); Teich and Porter (1996); Temin (1997);

[4] As Ashton (1948, p. 41) writes, "In the period 1700-1760 Britain experienced no revolution, either in the techniques of production, the structure of industry, or the economic and social life of the people."

inferred, even more rashly, that hence the term is useless. The latter inference is absurd: how many people in the Roman Empire referred to themselves as living during "classical antiquity?"[5] Yet the premise that contemporaries were unaware of the Industrial Revolution is simply and patently false. To be sure, they did not pay to it nearly the attention that subsequent historians have, but why should they have, not knowing where all this was leading? By confining oneself to reading Adam Smith (who published his *Wealth of Nations* in the very early stages of the Industrial Revolution), T.R. Malthus (who was above all interested in population and agriculture), or Jane Austen (who lived mostly in the South of England), one can easily misrepresent the perceptions of contemporaries. The Scottish merchant and statistician Patrick Colquhoun (1814, pp. 68-69) in a famous quote declared that "It is impossible to contemplate the progress of manufactures in Great Britain within the last thirty years without wonder and astonishment. Its rapidity ... exceeds all credibility. The improvement of the steam engines, but above all the facilities afforded to the great branches of the woolen and cotton manufactories by ingenious machinery, invigorated by capital and skill, are beyond all calculation..." At about the same time, Robert Owen (1815, pp. 120, 121) added that "The general diffusion of manufactures throughout a country generates a new character in its inhabitants... This change has been owing chiefly to the mechanical inventions which introduced the cotton trade into this country... the immediate effects of this manufacturing phenomenon were a rapid increase in the wealth, industry, population, and political influence of the British Empire." David Ricardo, despite being mainly interested in theoretical questions inserted a chapter on Machinery into the third edition of his *Principles of Political Economy* in which he is concerned with its impact on employment, an issue known as "the Machinery Question" and which only makes sense in the context of the Industrial Revolution (Berg, 1980).[6] Other writers and essayists, each from his or her own perspective, made similar comments. Similarly, literary references to the Industrial Revolution are not altogether absent, and Wordsworth, Blake, Charlotte Brontë, and Elizabeth Gaskell contain unambiguous references to the Industrial Revolution (see Mokyr, 1994, pp. 194-95 for details). Such references are relatively rare, but given the locational concentration of the Industrial Revolution in its earlier stages, this is not surprising.[7]

Nevertheless, there is a kernel of truth to the notion that the Industrial Revolution looms larger to us than it did to contemporaries. History is inevitably written with a certain amount of "presentism." Hindsight provides us with a tool to choose which details matter and which do not. In some instances, of course, this tendency should not be exaggerated. Some dead-ends and failures "mattered" as much as success stories and can be instructive for many reasons. The knowledge, however, that the Industrial Revolution set into motion a historical process of momentous global consequences is available to us and was not to contemporaries. It is a matter of taste and judgment to what extent that kind of knowledge should influence our work. Yet the

---

[5] Clearly awareness by contemporaries of the nature of the period in which they lived is not an absolute rule in Professor Cameron's book. He uses the term "Middle Ages" without qualm (chapter 3 of his textbook is called "Economic Development in Medieval Europe"). He may find it interesting to learn that the term was first used by one Christopher Keller or Cellarius in a book that appeared first in 1688. Although there, too, have been "countless reflections on the appropriateness of its label" the term has survived in conventional usage. See Fuhrmann, (1986), p. 16. I am indebted to my colleague Robert E. Lerner for bringing this reference to my attention.

[6] E.A. Wrigley (1994, pp. 30-31) makes essentially the same point when he notes that classical economists and their contemporaries were perfectly aware of the technological developments of their age and that it is impossible to doubt that Smith, Ricardo, and Malthus were as knowledgeable as anyone on these matters. Most political economists, however, rejected sustained economic growth as an equilibrium condition, largely on a priori grounds.

[7] As the area and the number of people affected by the Industrial Revolution increased, fiction, too, started to take note. In 1832 Elizabeth Gaskell moved to Manchester where she studied the same conditions that Friedrich Engels witnessed a decade later, resulting in her *Mary Barton* (1848). Both saw the same thing. Barton did not call it an Industrial Revolution (Engels did) but what they saw clearly disturbed them. Factory conditions are described in novels of the 1840s, obscure ones such as Frances Trollope's *Michael Armstrong, the Factory Boy* (1840) and Charlotte Elizabeth (Tonna)'s *Helen Fleetwood* (1840) and well-known ones such as Dickens's *The Old Curiosity Shop* (1841) and Disraeli's *Sybil* (1846). It is inconceivable that these authors were observing conditions that were brand-new.

thousands of scholars concerned with some aspect of economic growth, technological change, industrialization, and the emergence of the modern economies after 1750 are all employing this kind of judgment and for good reason. In 1815 it was impossible to discern whether the "wonderful progress of manufactures" was a temporary affair or the beginning of a sustained cumulative process of social and technological change, and some political economists believed, largely on a priori grounds, that progress would be temporary. Yet it is ludicrous for an economic historian at the end of the twentieth century to pretend to be equally ignorant.

In sum, in considering whether there "was an Industrial Revolution" I cannot do better than cite Max Hartwell, summarizing a career of study and reflection on the topic "Was there an Industrial Revolution?" succinctly: "There was an Industrial Revolution and it was British" (Hartwell, 1990, p. 575). Despite the announcements of opponents of the concept that modern research has demonstrated its vacuity, much recent work that looks beyond the aggregate statistics into the regional and microeconomic aspects of the Industrial - Revolution emphasizes the acceleration and irreversibility of economic change in the regions associated with the Revolution.[8]

The origin of the term *Industrial Revolution* was long attributed to two French-speaking observers writing in the 1830s, the Frenchman Jerome-Adolphe Blanqui and the Belgian Natalis de Briavoinne.[9] As David Landes shows elsewhere in this book, its origins can be traced back even further. All the same, there is little dispute that the term became popular following the publication of Arnold Toynbee's famous *Lectures on the Industrial Revolution* in 1884. The term is taken to mean a set of changes that occurred in Britain between about 1760 and 1830 that irreversibly altered Britain's economy and society. Of the many attempts to sum up what the Industrial Revolution really meant, the most eloquent remains Harold Perkin's: "A revolution in men's access to the means of life, in control of their ecological environment, in their capacity to escape from the tyranny and niggardliness of nature . . . it opened the road for men to complete mastery of their physical environment, without the inescapable need to exploit each other" (Perkin, 1969, pp. 3-5).

Although economic historians tend naturally to emphasize its economic aspects, the Industrial Revolution illustrates the limitations of the compartmentalization of historical sciences. More changed in Britain in those years than just the way goods and services were produced. The role of the family and the household, the nature of work, the status of women and children, the social role of the church, the ways in which people chose their rulers and supported their poor, what people wanted to know and what they knew about the world -- all these were altered more radically and faster than ever before. It is an ongoing project to disentangle how economic, technological, and social elements affected each other. The event itself transcended any definable part of British society or economic life; it was, in Perkin's phrase, a "more than Industrial Revolution."

What, then, was it that changed in the years that we refer to as the Industrial Revolution? We shall have to leave out of the discussion many of the aspects that made it a "more than Industrial Revolution" -- attitudes, class consciousness, family life, demographic behavior, political power, though all of these were transformed during the same period -- and concentrate on economic variables. Four different schools of thought about "what really mattered" during the Industrial Revolution can be distinguished.[10] The four schools differ in matters of emphasis and weight, yet they overlap to such an extent that many writers cannot be readily classified.

---

[8] For example, Marie Rowlands (1989, p. 124), who tries hard to find continuity in the economic changes in the West Midlands, is still describing it in dramatic terms: "There can be no question of the revolutionary impact of the introduction of the coal-fired blast furnace into the area from 1766. Within a single generation the furnaces . . . revolutionised not only the south Staffordshire economy but also its settlement pattern and landscape. . . . Agriculture became progressively more difficult, the night sky was illumined with flames and the day darkened with smoke, and the district began to be called the Black Country." Similarly, John Walton, writing of Lancashire, has no doubt that "there is something cumulatively impressive to explain. Nothing like it had been seen before. . . . The chain of events began in the 1770s and gathered . . . overwhelming momentum in the nineteenth century" (Walton, 1989, p. 64).

[9] Blanqui (1837, p. 389); Briavoinne (1839, vol. 1, pp. 185ff.).

[10] What follows is inspired by Hartwell (1971b, pp. 143-154), although the classification here differs to some extent.

1. **The Social Change School**. The Industrial Revolution is regarded by the Social Change School to have been first and foremost a change in the way economic transactions between people took place. The emergence of formal, competitive, and impersonal markets in goods and factors of production is the basis of this view. Toynbee ([1884] 1969, p. 58) writes that "the essence of the Industrial Revolution is the substitution of competition for the medieval regulations which had previously controlled the production and distribution of wealth." Karl Polanyi ([1944] 1985, p. 40) judges the emergence of the market economy as the truly fundamental event, to which everything else was incidental. A more recent contribution in this spirit, which emphasizes the emergence of competitive markets in manufacturing is Wijnberg (1992). Most modern social historians probably would view the central social changes as having to do with labor and the relation of workers with their work environment, other laborers, employers, and capitalists. An enormously influential work in this regard is E. P. Thompson (1963). Some recent contributions influenced by this work are Berg and Hudson (1992) and Randall (1991).

2. **The Industrial Organization School**. Here the emphasis is on the structure and scale of the firm -- in other words, on the rise of capitalist employment and eventually the factory system. The focal point is the emergence of large firms, such as industrial mills, mines, railroads, and even large retail stores, in which production was managed and supervised and where workers were usually concentrated under one roof, subject to discipline and quality control. The work of Mantoux (1928) is a classic example of this school, but Karl Marx's interpretation of the rise of "Machinofactures" also belongs here as do some modern writers in the radical tradition (Marglin, 1974-1975). A classic work discussing the Industrial Revolution from this point of view is Pollard (1965). In the same tradition is Berg (1994). More recently, Szostak (1991) has argued that changes in the organization of the firm were the causal factor in technological change and thus primary to it. Goldstone (1996) explicitly equates the Industrial Revolution to the emergence of the Factory system and argues that because China was unable for social reasons to adopt factories, the Industrial Revolution came late to it.

A somewhat different microeconomic approach to the Industrial Revolution emphasizes the distinction between circulating capital and fixed capital, a distinction that goes back to the classical political economy of David Ricardo and Marx. Some modern economists have defined the Industrial Revolution as a shift from an economy in which capital was primarily of the circulating kind (e.g., seed in agriculture and raw materials in domestic industry) to one in which the main form which capital took was fixed capital (e.g. machines, mines, and structures) (Hicks, 1969, pp. 142-43; Ranis and Fei, 1969).

3. **The Macroeconomic School**. The Macroeconomic School is heavily influenced by the writings of Walther Hoffmann and Simon Kuznets. Here the emphasis is on aggregate variables, such as the growth of national income, the rate of capital formation or the aggregate investment ratio, or the growth and composition of the labor force. Rostow (1960) and Deane and Cole (1969) are important proponents of this school, and their influence has extended to noneconomists (e.g., Perkin, 1969, pp. 1-2). Recent statements by E. A. Wrigley and Gary Hawke that baldly define the Industrial Revolution in terms of economic growth (Wrigley, 1987, p. 3; Hawke, 1993, p. 58) show that this approach still enjoys some support despite growing evidence that economic growth *during* the Industrial Revolution was unremarkable. Some writers, such as Gerschenkron (1962), prefer to aggregate on a sectoral level, dealing with the rate of growth of the manufacturing sector rather than the growth of the entire economy. Most practitioners of the New Economic History tend to belong to this school, because by its very nature it tends to ask questions about large collections of individuals rather than about single persons (Fogel, 1983, p. 29) and because of its natural interest in quantitative analysis.

4. **The Technological School**. The Technological School considers changes in technology to be primary to all other changes and thus focuses on invention and the diffusion of new technical knowledge. Technology is more than just "gadgets," of course: It encompasses techniques used for the organization of labor, consumer manipulation, marketing and distribution techniques, and so forth. The most influential book in this school is Landes (1969).

The attitudes of many writers regarding the revolutionary nature of the period is to some extent determined by the school to which they adhere. The most confirmed advocates of discontinuity have typically been *technological* historians. Quantitative analysis of patent statistics reveals a sharp kink upward in the late 1750s (Sullivan, 1989). Insofar as the level of technical innovation can be approximated by patenting, this finding lends support to the discontinuity hypothesis. Nonquantitative economic historians with a strong interest in technology have had little difficulty with the discontinuity implied by the use of the concept of the Industrial Revolution. David Landes's chapter in this book represents a summary of this view, which goes back at least to the writings of A. P. Usher and before.[11] Another leading technological historian, D.S.L. Cardwell (1972, p. 139), uses the term *revolutionary epoch* (which he reserves for the years 1790-1825), whereas Arnold Pacey (1975, p. 216) prefers to apply the term *revolutionary* to the last third of the eighteenth century. In a more recent work, however, he has no qualms about using the term *Industrial Revolution* (Pacey, 1990, chap. 7). H. I. Dutton (1984), Richard Hills (1979, p. 126), and Bertrand Gille (1978, p. 677) stress the technological discontinuities of this period. Maurice Daumas, despite reservations, accepts the concept for the case of Great Britain between 1775 and 1825 (1979, p. 8). Akos Paulinyi expresses the sentiments of many when he writes that "the perception [that denies the revolutionary character of the innovations during the Industrial Revolution] bewildered me because in no book on the history or philosophy of technology is it doubted that the technological changes which took place between 1760 and 1860 introduced a new era" (1986, p. 261). In his recent book on science and technology, Ian Inkster supports this view and adds that "removing the Industrial Revolution may simply lead to boredom" (1991, p. 61). Without necessarily accepting this view, it seems fair to object to a de-dramatization of the events purely because of some preconception that "nature does not make leaps."

On the other hand, historians interested in macroeconomics and emphasizing economic growth have in recent years found little support for discontinuities. In this they differ from earlier aggregative approaches such as Rostow (1960) and Deane and Cole (1969), which seemed to find sudden leaps in the macroeconomy. As Harley's essay in this book makes clear in more detail, modern research has established that economic growth before 1830 was slower than was previously thought. This could lead to the conclusion that the acceleration, if there was one at all, does not merit the adjective *revolutionary*. Table 1.1 presents average annual compound rates of growth of the economy before and during the Industrial Revolution, contrasting earlier and more recent efforts.

---

[11] Usher (1920, p. 247), in a chapter entitled "The Industrial Revolution," cites with approval J. A. Blanqui for stressing the profound changes occurring in his own lifetime (the 1830s) and adds that the two revolutions, the industrial in England and the political in France, each in their own way contributed to a break with the past "so complete that it is difficult for us to reconstruct the social life of the old régime."

TABLE 1.1 Estimated Annual Rates of Growth, 1700-1871
(in percentages)

Period	National Income per cap. (Deane & Cole)	National Income per cap. (Crafts)	Indust. Product (Hoff-mann)	Indust. Product (Deane & Cole)	Indust. Product (Harley)	Indust. Product (Crafts)	Indust. Product (Cuenca)
1700-1760	0.44	0.3	0.67	0.74	n.a.	0.62	--
1760-1800	0.52	0.17	2.45	1.24	1.6[a]	1.96	2.61[c]
1800-1830	1.61	0.52	2.70	4.4	3.2[b]	3.0	3.18
1830-1870	1.98	1.98	3.1	2.9	n.a.	n.a.	

a - 1770-1815
b - 1815-1841
c - 1770-1801

SOURCE: Computed from Harley (below); Hoffmann (1965); Cuenca (1994).

Compared to Deane and Cole's national income statistics, Crafts' figures reveal an aggregate growth that was much slower during the Industrial Revolution. Industrial production is more ambiguous: Hoffmann's data, computed in the 1930s, clearly show a rapid acceleration during the period of the Industrial Revolution, but Deane and Cole's series is much more erratic and, like the revisionist data of Harley and Crafts, show that most of the quantitative expansion occurred *after* 1800. All the same, Crafts and Harley explicitly deny adhering to a school that would negate the profound changes that occurred in Britain during the Industrial Revolution (1992) and restate that "industrial innovations . . . did create a genuine Industrial Revolution reflected in changes in Britain's economic and social structure," even if their impact on economic growth was more modest than previously believed (p. 3). The point stressed by Crafts and Harley, as well as by students of other episodes of rapid technological change, is worth repeating: There is typically a long lag between the occurrence of changes in technology, even those of fundamental importance, and the time they start affecting aggregate statistics such as industrial production and national income per capita.

The revisionist view of the Industrial Revolution proposed by Harley and Crafts has led to lively exchanges with scholars critical of their methodology and views. Landes (below) still feels that during the Industrial Revolution growth of per capita income accelerated to the extent that we are justified in considering the Industrial Revolution a breaking point. In a different mode, a number of scholars have attacked the quantitative methodology underlying the revisionism and pointed out that rather than based on new research, the new series proposed were a reshuffling of the same raw materials used by Deane and Cole and questioned one detail or another in the technical procedures (Hoppitt, 1990; Jackson, 1992, 1994; Cuenca, 1995). In particular, as table 1.1 indicates, Javier Cuenca has questioned the estimates of industrial output growth produced by Crafts and Harley. Given the significant role of the lower industrial output growth estimates in GDP (Jackson, 1994, p. 91) these scholars can be seen to have taken issue with the fundamental revisionism which contends that *during* the Industrial Revolution aggregate growth rates were far lower than Deane and Cole had originally postulated.[12] All the same it remains a matter of consensus that we do not observe, and indeed *should* not observe a sharp break in aggregate long-term growth rates.

On a different front, the Crafts-Harley has been criticized by Berg (1994) and Temin (1997). Part of the economic logic of the Crafts-Harley view of slow growth was that productivity growth and technological progress were confined to a few relatively small sectors such as cotton, wool, iron, and machinery whereas much of the rest of manufacturing remained more or less stagnant till after 1830. Temin maintains that this argument is inconsistent with the patterns of British foreign trade, which clearly shows that Britain maintained a comparative advantage not just in the rapidly expanding "new industries" but in a host of small, older industries such as linen, glass, brewing, pottery, buttons, soap, candles, paper, and so on. Temin relies on export figures to make a point about comparative advantage and to infer from it indirectly that technological progress occurred on a variety of fronts. Anecdotal evidence and examples of progress in industries other than the paradigmatic high-flying industries can be culled together from specialized sources.[13]

Nonquantitative analysts also disagree on the issue. The Social Change School tends to be divided: Toynbee and his contemporary H. Gibbins (1895) thought that the changes that mattered most were rapid. Modern social historians such as Jonathan Clark would clearly disagree. More recent work (e.g., Berg and Hudson, 1992) asserts that the pendulum has swung too far in the direction of gradualism and points to a number of radical and discontinuous social changes. The same holds for the Industrial Organization School; whereas Mantoux clearly believed in sudden and rapid change, modern scholars in this tradition are more gradualist in their views and stress the dynamic elements in the pre-1760 economy. Maxine Berg (1994) has resisted the new quantitative orthodoxy of Harley and Crafts while insisting at the same time (p. 281) that "industrial growth took place over the whole eighteenth century and not just in the last quarter of it." In any event, there is no justification for extreme statements such as that of Musson (1978, p. 149), who flatly declares that by 1850 Britain was not a very different economy than it had been in 1750. After all, the population of Britain had tripled by that period, and at least in some regions everything, from the landscape to the occupational structure, had been turned upside down. The statement is, perhaps, closer to the truth for southern and eastern England and the Scottish Highlands, but even there it is debatable.

Debates on gradualism vs. sudden change are not specific to the literature on the Industrial Revolution or even economic history. There has always been an intellectual current that believed with Charles Darwin and Alfred Marshall that Nature makes no leaps. Within evolutionary biology, a debate between gradualists and saltationists has been conducted with equal intensity and perhaps similarly inconclusive results (Mokyr, 1990b, 1991a). After many years of undisputed reign by gradualists, a new compromise is emerging that allows for

---

[12]The most effective criticism was made by Cuenca (1994) who has questioned the procedures used by Crafts and Harley (1992) to estimate the growth of the cotton industry during the Industrial Revolution. Cotton output was the fastest growing component of industrial production, and its relative share in industrial output is thus a crucial variable in the estimation of industrial output. Cuenca argues that cotton prices fell rapidly after 1770 and hence output was growing faster than is generally believed. His revisions in the prices of cotton raise the rate of output growth of industrial production from the 1.27 percent per year estimated by Crafts and Harley to a much higher level of 2.61 percent, higher even than Deane and Cole's estimate. In their "reply", Crafts and Harley (1995) dispute the price series used by Cuenca and point out that his figures imply that in 1770 the relative share of cotton in the industrial sector was far larger than was hitherto assumed which explains the large increase in aggregate industrial output claimed by Cuenca. In any case, even the radical revisions in *industrial* growth proposed by Cuenca do not change GDP growth rates by all that much, from the 1 percent per year (1760-1801) estimated by Crafts to about 1.4 percent (ibid., p. 142). Still, such seemingly small differences in growth rates compounded over 40 years would mean that GDP would be 75 percent higher in 1801 than in 1760, as opposed to 49 percent by the lower growth rates. Since population grew at around 0.8 percent per annum over the same period, meaning that population in 1801 was about 40 percent higher than in 1760, these differences imply rather dramatic differences in income per capita growth.

[13]On the hardware industry, see Berg (1994), ch. 12. On many of the other industries classic industry studies carried out decades ago have not yet been supplanted such as Coleman (1958) on the paper industry, Mathias [1953,(1979)] on brewing, Haber (1958) on the chemical industries, Church (1970) on the shoe and boot industry, McKendrick (1961, and 1982b) on potteries, and Barker (1960) on glass.

sudden outbursts of accelerated change although not insisting that *all* historical change is necessarily of that kind. It seems that economic historians and evolutionary biologists have been walking on parallel paths.

A moment of reflection and a few simple computations indicate that for a country that undergoes structural change while it grows, very sudden accelerations in the growth rate of the kind that Rostow envisaged are simply impossible. Thus the finding that the aggregate effects of the Industrial Revolution are not overwhelming before 1820 is not surprising. It is useful for this purpose to regard Britain during the period of the Industrial Revolution as a dual economy in which two economies coexisted although the argument would be no different if we considered a continuum of many sectors. One was the traditional economy, which, although not stagnant, developed gradually along conventional lines, with slow productivity and slowly rising capital-labor ratios. This sector contained agriculture, construction, domestic industry, and many traditional "trades" that we would now classify as industrial but which in the eighteenth century and before were partially commercial: bakers, millers, tailors, shoemakers, hatters, blacksmiths, tanners, and other craftsmen. The modern sector consisted of cotton, iron smelting and refining, engineering, heavy chemicals, mining, some parts of transportation, and some consumer goods such as pottery and paper. At first, however, only segments of these industries underwent modernization, so that dualism existed *within* as well as *between* various products, which makes calculations about the performance of the modern sector rather tricky.[14] According to McCloskey's (1985) computations, the traditional economy was large, if relatively shrinking. The average size of agriculture and "all others" between 1780 and 1860 was 79 percent of the British economy, meaning that in 1760 it was likely to have composed close to 90 percent of the British economy. Productivity growth in this sector is estimated by McCloskey at about 0.6 percent per annum. During the same period productivity in the modern economy grew at a rate of 1.8 percent per annum.

Two-sector growth models imply that abrupt changes in the economy *as a whole* are a mathematical impossibility because the aggregate rate of growth of any composite is a weighted average of the growth rates of its components, the weights being the respective shares in output. Even if changes in the modern sector itself were discontinuous and its growth rate very high, its small initial size would limit its impact on the economy-wide growth rate, and its share in the economy would increase gradually. In the long run the force of compound growth rates was such that the modern sector swallowed the entire economy. How long was the long run? A numerical example is illuminating here. Assume two sectors in a hypothetical economy, one of which (the modern sector) is growing at the rate of 4 percent per annum while the other (the traditional sector) is growing at the rate of 1 percent per annum.[15] Suppose that initially the modern sector produces 10 percent of total output. Then the aggregate growth rate is at first 1.3 (=.9x1 + .1x4) percent. After ten years the aggregate rate of growth will have increased to 1.39 percent per year. After thirty years of "dual growth" the share of the modern sector will have increased to 21 percent of the economy and after fifty years to one-third. Only after

---

[14] Some approximate idea of the differences between the two sectors can be obtained from comparing pre-1760 rates of output growth to those between 1760 and 1800. Real output in cotton, for example, grew at 1.37 percent per annum in 1700-1760 and 7.57 percent in 1760-1800. In iron output, the growth rates were, respectively, 0.60 percent and 4.10 percent. In two traditional industries the acceleration is less marked: In linen the growth rates were 1.25 percent and 1.44 percent, and in leather 0.25 percent and 0.57 percent, respectively (all data from Crafts, 1985a, p. 23).

[15] Note that these rates differ from the ones McCloskey presents, since what is relevant here is *total* output growth, not productivity growth. The average rate of growth of "manufactures, mining, and building" in 1801/11-1851/61 was 3.57 percent, whereas that of "agriculture, forestry, and fishing" was 1.5 percent per annum (Deane and Cole, 1969, p. 170). For the closing decades of the eighteenth century, industrial output grew according to Crafts's calculations at a rate of 2.11 percent per annum and agricultural output at 0.75 percent. Crafts has also revised Deane and Cole's figures for the nineteenth century, but the differences are not large enough to affect the point made here. As was noted above, the rate of growth of the "modern sector" must have been faster than that of "industry." For instance, the consumption of cotton--the raw material of the modern industry par excellence--increased at the annual rate of 10.8 percent between 1780 and 1800 and at the rate of 5.4 percent between 1800 and 1840. In his essay below, Clark radically revises the growth of agriculture and claims that there was practically no growth of agricultural output in the eighteenth century. Yet the traditional sector was more than agriculture, and some sectors clearly were benefitting from improvements elsewhere in the economy.

seventy-four years will the two sectors be of equal size (at which point aggregate growth equals 2.5 percent per year), and a full century after the starting point the traditional sector will have shrunk to about 31 percent of the economy. The British economy as a whole was changing much more slowly than its most dynamic parts, because growth was diluted by slow-growing sectors (Pollard, 1981, p. 39). These hypothetical numbers fit the actual record rather well, and they indicate that it is hardly surprising that it took until 1830 or 1840 for the economy-wide effects of the Industrial Revolution to be felt.

In reality the "modernity" of industries and enterprises was a continuum rather than a dichotomy, and the example is thus highly simplified. The distinction between the modern and traditional sectors leaves an inevitable gray area, and it has been criticized effectively in recent work as a simplification (Berg and Hudson, 1992). Not all industries that mechanized were growing quickly (e.g., paper), and not all industries in which output was growing rapidly were subject to rapid technological change.[16] In some industries, such as instrument and clock making, important technological changes were occurring in a traditionally organized industry. The distinction also abstracts from what actually happened in that it does not take into account that the modern and the traditional sectors affected each other. Although technological change in the traditional sector was slow by comparison, its productivity was affected by what happened in the modern sector. For instance, construction technology may have changed slowly, but improvement in transportation technology allowed the shipment of bricks throughout Britain, which made cheaper and better buildings possible. Agriculture benefited in some ways from technological developments in manufacturing, including the production of clay and, later, metal drainage pipes and various agricultural machines and implements. The development of coke ovens allowed the extraction of tar from coal. Gaslighting, one of the most neglected of the "great inventions," allowed many artisans and craftsmen in the traditional sector to work longer hours and reduced the cost of night work (Falkus, 1982). These intersectoral spillover effects imply that the distinction between the traditional and modern sectors is to some extent arbitrary. The coexistence of the old and the new is important, and the interaction of the two sectors greatly affected the growth of the aggregate. These interactions do not, however, change the principle of gradual change of the aggregate economy.

Despite the abstraction involved in distinguishing between a modern and a traditional sector, many economic historians still think that two-sector models are useful (Crafts, 1985a; McCloskey, 1985). The modern sector was more than industry but not all of industry. Its production was carried out in workshops or factories where workers were concentrated in workplaces away from their homes, many of which were located in urban or suburban areas. The traditional sector, roughly speaking, covered industries and services that remained little affected by the new technology. Much of the production was still carried out in the household or small workshop (though some larger establishments employed nonmechanized techniques), where the worker had few personal interactions with other workers or supervisory personnel. The interaction of the two sectors was, of course, reciprocal. From the point of view of the modern sector, the traditional sector was important because it determined the sociopolitical environment in which the new industries operated. And, although the modern sector was largely self-sufficient in capital and partially so in raw materials, it depended on the traditional sector for its labor supply and skills.

Utilizing the distinction between a modern and a traditional sector allows us to summarize what happened to the British economy during the Industrial Revolution as a three-pronged economic change. First, a small sector of the economy underwent quite rapid and dramatic technological change. Second, as a consequence, this sector grew at a rate much faster than the traditional sector so that its share in the overall economy continued to increase. Third, the technological changes in the modern sector gradually penetrated the membrane of the traditional sector so that parts of the *traditional* sector eventually became modernized. The economy grew, but because its sectoral composition changed, it did more than just increase in size, it was "growing-up" (Mokyr, 1976b).

---

[16] There is no a priori economic reason that suggests that industries in which technological change was rapid would also necessarily experience rapid output growth. If technological progress was especially important in industries for which demand was inelastic, these industries would possibly grow slower than industries for which demand was highly price and income elastic.

The idea that the Industrial Revolution was primarily a story of rapid economic growth has thus been discredited. One obvious reason is the composition effects just described. But there are other arguments raised by scholars in recent years that have cast some doubt on this view. One is that the assumption that the pre-1750 economy (despite some obvious fluctuations in population and income) was essentially trendless is difficult to sustain. Although answers to the questions about what happened to long-term income before 1800 are even more limited by data problems, the circumstantial evidence seems to indicate that on the eve of the Industrial Revolution Britain was already a wealthy and sophisticated market economy. This means it must have been growing during *some* stages of its medieval and early modern past.

Moreover, in addition to the stormy developments in production technology, the British economy in the eighteenth century was subject to other, more gradual forces that affected the long-term growth of income. The most prominent of these forces were the growth of trade and the division of labor it brought with it. For Adam Smith, not surprisingly, the gains from trade and specialization were the main sources of economic growth. As Table 1.1 indicates, economic growth preceded the Industrial Revolution and thus can hardly have depended on it. Jones (1988) emphasizes that the technological changes of the last four decades were superimposed on an economy that was already growing. Had there been no Industrial Revolution, growth would have continued in the long run, though at a much slower (and decelerating rate). The Smithian and the technological elements of economic change, though interrelated at many nodes, could operate independently of each other. The Industrial Revolution was neither a necessary nor a sufficient condition of economic growth. In the very long run, however, without continuous technological change, growth would slowly grind to a halt. The gains from trade and specialization, which in Smith's vision were the key to wealth, would have run into diminishing returns, as further declines in transportation or transactions costs would have yielded smaller and smaller marginal gains. Similarly, gains from improvements in the allocation of resources due to more effective economic institutions and the development of markets in factors and resources, eventually start yielding less and less as most of the easy gains are made early on. Changes in technology, that is, changes in human knowledge and ability to understand and utilize the laws of nature, is the the only dynamic element that seems thus far to be exempt from diminishing returns.

Despite the disagreements in interpreting the Industrial Revolution, it is appropriate to note that there are many areas of broad agreement. The consensus is that within the relatively narrow confines of production technology in a number of industries, more numerous and more radical inventions occurred during the Industrial Revolution than ever before in so short a period. It is equally uncontroversial that these changes had a far-reaching effect on the lives of only a minority of Britons throughout our period. The Industrial Revolution was, above all, a regional affair, affecting Lancashire and parts of the adjoining counties and the Scottish Lowlands but leaving most of the rest of the country without visible marks. As late as 1851, only about 27 percent of the British labor force worked in the industries that were *directly* affected by the Industrial Revolution, although almost everyone had been touched by it indirectly as consumer, user, or spectator.

Furthermore, it is beyond dispute that one of the problems with assessing the macroeconomic and social impact of the Industrial Revolution in its early stages is that it occurred simultaneously with other events whose effects are impossible to disentangle from those of the Industrial Revolution proper. Unlike a chemical experiment, history does not provide us with the circumstances to test the effects of one element by holding the others unchanged. First, for most of the period under discussion here, Britain was at war. Wars disrupted commerce and finance, increased taxation, and siphoned off labor to unproductive uses. Second, the Industrial Revolution coincided with the resumption of population growth in Britain, which until the middle of the eighteenth century had slackened off. There were ever more people who needed to be fed and clothed, threatening to materialize the dire predictions of the Reverend Malthus. The economic impact of population change was further complicated by the fact that it was in large part due to an increase in the birth rate. Like many underdeveloped countries today, this left Britain with an ever-younger population in which the proportion of small children who did not yet work was increasing.[17] Third, the Industrial Revolution happened to occur during a period of worsening weather conditions, leading to a string of poor harvests, high food prices, and scarcity. Some of the worst harvests, as fate would have it, coincided with the war years, as they did in 1800/01 and 1812/13, compounding the misery.

These three extraneous factors -- wars, population growth, and poor harvests -- were not caused by the Industrial Revolution nor did they affect it directly. From the point of view of the economic historian looking for causes and effects, they are contaminations in an economic experiment that could be carried out only once. Economic history does not lend itself to neat and clean analysis: Even if we had far more data than we do, contaminating events and feedback loops make it exceedingly difficult to reach definite conclusions about causality. Yet the importance of the Industrial Revolution in British and indeed world history is such that we cannot afford not to try.

2. What Was the Industrial Revolution?

Technological determinism does not enjoy a great reputation among scholars, and in many accounts it is usually preceded by the telling adjective "crude."[18] In the metaphor coined by a famous if anonymous schoolboy cited by T. S. Ashton, the Industrial Revolution is defined as "a wave of gadgets that swept Britain." In this view, invention becomes an exogenous variable that then affects the endogenous variables: factories, urbanization, social change, and, with a long lag, economic growth. This is an unsatisfactory cartoon of history. Inventions do not rain down upon an economy like manna from heaven. They are stimulated by economic and social pre-existing conditions. They emerge in the minds of some people for some reason which may or may not be identified, are communicated, adapted, refined, implemented, and imitated. An innovation may succeed or it may be resisted so fiercely that it never has a chance to compete. Some societies exhibit a quality that, lacking a better term, we will call "technological creativity." Technological creativity is not the same as inventiveness; it also includes the willingness and ability to recognize and then adopt inventions made elsewhere. We have barely begun to understand why some societies are technologically creative and others are not, and why some societies that are technologically creative at some time cease to be so later on. I will argue below that Britain, indeed, was a technologically creative society, and that we can make some reasonable hypotheses as to why and how she became so. Regardless of its source, the Industrial Revolution was above all an age of rapidly changing production technology propelled by technological creativity.[19] This view attributes to technology an important historical role, and the challenge is to somehow disentangle those cases in which technological change "may indeed have had some independently initiating role from others in which it is better understood as secondary, dependent, or derivative" (Adams, 1996, p. 107).

---

[17] The dependency ratio (defined as those aged 0-14 and those aged 60 and over divided by those aged 15-59) thus increased from 815 in 1751 to 942 in 1801 (1826 = 1000) (Wrigley and Schofield, 1981, p. 447).

[18] For a recent summary of this literature, see Smith and Marx, eds. (1994). This collection highlights two kinds of technological determinism: one that views technology as an autonomous force which acts on other variables but is not explained itself, and another that regards technology as one of the central forces determining economic performance. Economic historians have rarely felt particularly guilty at assigning a major role to technology in history because of their preoccupation with *material* conditions. Moreover, technological historians such as David Landes and Lynn White have done much to improve our understanding of the cultural and economic sources of technological progress. In so doing, they have identified technological innovation as one channel through which existing social conditions and changes in human knowledge lead to economic change, and they can hardly be accused of "crude" determinism.

[19] To some students, the definition of the Industrial Revolution in technological terms may seem commonplace, even banal. Yet in some corners there are serious reservations about this view. Braudel (1984, p. 566) states categorically that "if there is one factor which has lost ground as a key explanation of the Industrial Revolution, it is technology."

The story of the most important innovations of the Industrial Revolution has been told elsewhere many times.[20] Without repeating all the details here, it may be useful to make a few distinctions that help to make sense of the story. Technological change consists of the creation of new knowledge and its diffusion and implementation, sometimes referred to as innovation. As always there is a gray area between the two, and here it is rather large. On many occasions when a known technology is introduced in a new place, it has to be modified and adapted to suit a different environment and sometimes a different product, and thus it acquires some of the characteristics of invention. Inventions and innovations are very much complementary processes, and asking whether technological change proceeds more by one or the other is like asking whether a pianist makes music with the left or the right hand. An invention that is not adopted remains a dead letter and at best ends up in a footnote in a text on the history of technology. On the other hand, without new inventions the process of innovation will lose steam and eventually reach a dead end.

We can envisage the relation by using the economist's terms of average- and best-practice technique. At any given point of time an industry uses a variety of techniques. Some producers use the most recent and most up-to-date (best-practice) technique, but because of a variety of diffusion lags not all firms use state-of-the-art technology all the time. As best-practice techniques are diffused, the average-practice technique pursues and eventually catches up with the state-of-the-art technique. If, however, the technical frontier advances continually through invention, average-practice never catches up with best-practice. Invention keeps throwing new fuel on the fires of innovation and progress. The rate of progress of an industry is thus a function of both the rate of advance of the best-practice techniques and the mean diffusion lag.

Many of the inventions that made the British Industrial Revolution were, in fact, adaptations of inventions made overseas. Thus the Fourdrinier paper machine, introduced by Bryan Donkin in London in 1807, was originally invented by the Frenchman N. L. Robert in 1798. Gaslighting, the Leblanc soda process, chlorine bleaching, and the wet-spinning process for flax were Continental inventions imported into and perfected in Britain. By being receptive to these foreign technologies, as much as through their own inventions, Britain's industries displayed an unprecedented technological creativity that lay at the foundation of the British Industrial Revolution.

Inventions, too, come in different sizes and packages. If we counted successful inventions mechanically as if they represented one unit each, we would find that the vast bulk of inventions made during the Industrial Revolution -- or in our own time -- were small, incremental improvements to known technologies. Such "gap-filling" inventions are often the result of on-the-job learning-by-doing or of a development by a firm's engineers who see opportunities to produce a good cheaper or better. Over time, a long sequence of such *micro*inventions may lead to major gains in productivity, impressive advances in quality, fuel and material saving, longevity, and so on. At times the accumulated effect of incremental inventions changed the nature of the product. Consider one example: the sailing ship. Since the emergence of the fully rigged, three-masted ship in the fifteenth century, the art of shipbuilding had not been stagnant: Ships were cheaper to build and to maintain, more seaworthy, and more durable in 1800 than in 1450. Yet there had been no radical changes in either planking or rigging, no discontinuous leaps in ship design (Gilfillan, 1935) since 1500. The same is true for technologies as diverse as the cultivation of grains, the smelting of iron ore, the printing of books, and the making of guns.

Far rarer, but equally important, were dramatic new departures that opened entirely new technological avenues by hitting on something that was entirely novel and represented a discontinuous leap with the past. Such *macro*inventions created what Dosi (1988) has called technological paradigms, entirely new ways of thinking about and carrying out production. Within the new paradigm, once it is created, incremental *micro*inventive activity takes over: radically novel techniques need to be adjusted, extended, refined, and debugged.[21] It is rare that a totally

[20] See, for instance, Ashton (1948); Cardwell (1972); Landes (1969); Mantoux (1928); Mokyr (1990a, chap. 5); Mokyr (1992a).

[21] As one of the great engineers of the Industrial Revolution, John Farey, told a Parliamentary committee in 1828, "The inventions which ultimately come to be of great public value were scarcely worth anything in the crude state, but by the subsequent application of skill, capital and the well-directed exertions of the labour of a number of inferior artizans . . . brought to bear to the benefit of the

new invention is fully ready to go into production from the start. But it is equally clear that without occasional leaps of this kind, the process of continual incremental improvement within an existing technological paradigm would run into diminishing returns and eventually give out.

An exact criterion to distinguish macro- from microinventions is not easy to define. On the whole, a successful macroinvention meets three criteria: novelty, workability, and potential for further improvement. It involves a new technique to carry out production or consumption in a way that was radically different from anything before. Yet a radical idea, even a blueprint, that could not actually be materialized in practice was useless. Without the workmanship, the materials, and the supporting maintenance technology, the new idea would not survive. Macroinventions typically open new avenues to further improvements in production, reducing cost and enhancing product quality, finding new applications and new permutations, so that eventually it also acquired economic significance. However, it need not be a single event. Many macroinventions consisted of a number of steps that together were necessary for the new paradigm to emerge. The number of steps has to be small enough, however, to preserve some sense of discontinuous change.

The steam engine is a case in point.[22] It was conceptually one of the most radical inventions ever made. Energy, as used by people, comes in two forms: kinetic energy (work or motion) and thermal energy (heat). The equivalence of the two forms was not suspected by people in the eighteenth century; the notion that a horse pulling a treadmill and a coal fire heating a lime kiln were in some sense doing the same thing would have appeared absurd to them. Yet converting heat into work must be regarded as one of the most crucial advances ever made; energy was exploited through controlled fire, the domestication of animals, and the use of watermills and windmills. However, heat and work were not yet convertible into each other, so that wood and fossil fuels could not be used to produce motion and watermills could not produce heat.[23]

By breaking through this separation, then, the steam engine was truly radical. Its invention stemmed from the realization that the earth was surrounded by an atmosphere and that differences in atmospheric pressure could be utilized to harness energy. Suggestions of this kind had been made throughout the second half of the seventeenth century, but the half-baked sketches and flights of the imagination did not add up to much until 1690 when Denis Papin produced a prototype of a piston that moved up and down in a cylinder due to alternative heating and cooling. Thomas Savery's vacuum pump notwithstanding, the first truly successful steam engine was not produced until 1712 when an English engineer named Thomas Newcomen produced the first working steam engine. Large, cumbersome, noisy, and voracious in its appetite for fuel, the Newcomen engine must have appeared fierce and somewhat awesome to contemporaries. It was a prime example of what some have called "a hopeful monstrosity."[24] Newcomen engines were, however, viable and were used widely as pumps in mines where fuel was plentiful and flooding a threat. It was not until 1765, however, that the steam engine could be turned into an economic revolution, when James Watt introduced the separate condenser, as well as number of other important microinventions.

A second macroinvention of enormous economic importance was the invention of mechanical spinning. Since time immemorial, spinning had been carried out by a distaff-and-spindle method in which the spindle was dropped while the worker twisted the rovings of raw material and turned it into yarn. The index finger and thumb of the spinner, or (usually) spinster, were essential to this process, because it was their motion that drew out the

community . . . such improvements are made progressively, and are brought into use one after another, almost imperceptibly" (cited by Inkster, 1991, pp. 84-86).

[22] For a similar argument, see Cipolla (1965).

[23] There was one exception to the rule. Gunpowder as used in the West was a method to convert heat into kinetic energy. But it was an uncontrolled conversion, and the uses of gunpowder for civilian purposes prior to the invention of dynamite were limited. It is telling that Christiaan Huygens, a Dutch scientist, proposed in 1673 to build a combustion engine prototype using gunpowder.

[24] The term was actually coined by biologist Richard Goldschmidt to denote mutations that create new species.

fibers and carried out the true "spinning." The addition of the spinning wheel in the Middle Ages did not change that principle; the wheel just helped wind the finished yarn on a rapidly turning spindle. Replacing the human finger by a machine turned out to be a difficult problem, and it took until the last third of the eighteenth century to finally find a solution. When it happened, not one but two inventions emerged, which together changed spinning forever. One was the throstle, or water frame, invented by Richard Arkwright in 1769, which used two pairs of rapidly turning rollers to mimic the human fingers. The other was the Hargreaves spinning jenny (1765), based on the insight that it was possible to impart the twist by the correct turning of the wheel itself, with metal bars guiding the spun yarn onto the spindle. These two were then combined in 1779 by a third inventor, Samuel Crompton, into a hybrid of the two, appropriately called the mule. For more than a century, the mule remained the backbone of the British cotton industry.

The inventions in spinning led to a technology that was radically different from what came before. Economically, its importance was that it delivered a yarn that cost a small fraction compared to the previous technique and yet was of far higher quality than anything that could have been produced in Britain before. The new spinning technology practically created an industry *de novo* (prior to 1770 cotton had been a small industry, in the shadow of its cousins, the woolen and linen industries). Above all, the spinning machines were truly a novel concept, one that could subsequently be further improved. The novelty was in the substitution of a machine for the fine movements of the human fingers, one of the most delicate and flexible mechanisms designed by nature. Although cotton spinning was concentrated in a small part of Britain (Lancashire), its ramifications were truly global. It led to the destruction of the Indian cotton-spinning industry, which previously had supplied the high-quality yarns needed to make calicoes. Across the Atlantic, the growth of the British cotton industry led to the emergence of the cotton economy and the survival of slavery in the United States.

The *economically* most important inventions were not necessarily the most spectacular macroinventions, though that was the case with the steam engine and cotton-spinning machinery.[25] Consider, for instance, the invention of the puddling-and-rolling technique by Henry Cort in 1784, which solved the problem of efficiently converting the output of blast furnaces, pig iron, into what industry needed, wrought iron. Arguably, it was the most indispensable invention of the era because unlike steam power and cotton there was no substitute for iron. Yet Cort's invention was hardly a radical departure; rolling had been practiced for centuries, and the conceptual novelty of the process was modest. On the other hand, consider the Jacquard loom, invented in France in 1804. This loom wove complicated patterns into fabric using instructions that were embedded in an endless chain of cards, which had holes that were prodded by special rods. What these cards contained was a revolutionary new insight: the binary coding of information, a system that was conceptually novel and a harbinger of things to come. The importance of the insight was fully recognized by Charles Babbage, the inventor of the "analytical engine," which was the precursor of the modern computer. Yet the Jacquard loom produced largely an up-market, expensive product (silk and high quality worsteds) and did not produce a very different product than the old draw loom. Its economic significance, compared with Cort's invention, was relatively small.

The most radical of macroinventions of the time had even less of an economic impact: hot air ballooning (invented in France in 1783). It never had much commercial use, and even its military use, though attempted, was less than decisive. Yet it was one of the most radical technological events of all time: the first manned flight, defeating the tyranny of gravity. It was typical of the period, the last third of the eighteenth century, in which traditions, conventions, and old boundaries were recklessly cast aside and new ideas tried everywhere. In 1798, Edward Jenner discovered the smallpox vaccination process, in which for the first time non-human substances were introduced into the human body to confer immunity -- an unprecedented step in the history of medical technology. Other examples abound: the use of gas for lighting, the bleaching of fabrics with chlorine, new designs in waterwheels, the preservation of food through canning, and the idea of interchangeable parts in clocks and firearms all date from this period.

---

[25] The "social savings" of an invention is defined as the addition to total consumer surplus generated by it. It thus depends on the difference in costs between using the technique in question and the next best alternative.

A technological definition of the Industrial Revolution is a clustering of macroinventions leading to an acceleration in microinventions. The macroinventions not only increased productivity at the time but opened enough new technological vistas to assure that further change was forthcoming. Such a definition does not pretend to exhaust what happened in Britain in those years. The macroinventions were significant in large part because they created the germs of what came later: a gradual diffusion, adaptation, improvement, and extension of the techniques developed during the Industrial Revolution. The high-pressure steam engine led to the railroad and steamship. Improvements in cotton-spinning and weaving were reinforced by innovations in the preparatory stages in yarn-making, such as carding and slubbing and the finishing processes such as bleaching and printing. The inventions in cotton manufacturing spread to wool and linen. The cheap wrought iron found many new uses for iron, including construction, water mills, ships, machines, and specialty tools. The Leblanc soda-making process (1787) and bleaching powder (1798) laid the foundation for a chemical industry. In the absence of subsequent microinventions, some macroinventions remained little more than curiosa. Thus Faraday's invention of the electrical motor in 1821 remained of largely academic interest until the principle of self-excitation was developed in the late 1860s. Ballooning, too, could not be exploited commercially until small, lightweight engines could be mounted on the balloons for steering.

Despite the obvious importance of changes in technology in the British economy, their analysis and measurement have been slippery, and economists have found it exceedingly difficult to quantify them. Innovations and inventions are difficult to count and they do not follow the laws of arithmetic. An invention may supersede a previous invention, it may be independent of it, or it may in fact supplement it and improve it. The combined effects of two inventions could thus be equivalent to one, two, or a larger number of improvements. Yet economic historians have felt intuitively that if technological change is to be analyzed, it has to be quantified in some way. There are two alternative ways of measuring the level of technological change. One is the counting of patents or related statistics, which is a microeconomic approach. The other is estimating total factor productivity, which is mostly a macroeconomic approach.

Patent statistics have always tempted economists. Jacob Schmookler (1966), whose work is often cited in this respect, was preceded in his interest in patents by economic historians such as Ashton (1948, p. 63) who pointed to the sharp rise in patents as a symptom of the Industrial Revolution. Recently, the patent statistics have been subjected to quantitative analysis (Sullivan, 1989; for the United States, see Sokoloff, 1988, and also Griliches, 1990). Yet the counting of patents has always been subject to sharp criticism. First, it is a measure of *invention*, not of *innovation*. The statistics reveal nothing about the subsequent usefulness of the invention: Arkwright's and Watt's patents would be counted together with that of the inventor who took out a patent on nightcaps specially designed for sufferers from gout and rheumatism. Weighting the patents by their "importance" is of course far from easy. Second, not all important inventions were patented. The reasons for this range from the inability of the inventor to pay the required fee (£100 for England, £350 for Great Britain as a whole) to the inventor's preference for secrecy. This objection would perhaps not be so damaging if the inventions that were patented were in some sense a representative sample of the larger population of inventive activity. But recent research strongly suggests that that was not the case (Griffiths, Hunt, and O'Brien, 1992; MacLeod, 1988). Patenting statistics thus measure the propensity of inventors to patent as well as the distribution of inventive activity over high- and low-patent industries. As such, its usefulness as an index for the level of inventive activity is limited.

Total factor productivity measurements take a completely different road: they are, if anything, measures of innovation, not of invention. The economic logic behind total factor productivity estimates is that output grows due to either increases in inputs or shifts *of* the production function (such as technological change). If the weighted contributions of the inputs are subtracted from the growth rates, the "residual" measures the rate of productivity growth, which is associated with innovation.[26] The two best-known attempts to compute total factor productivity for Britain during the Industrial Revolution were made by McCloskey and Crafts, and they are

[26] The actual estimation (e.g., McCloskey, 1985) often uses the "dual" approach, which consists of looking at input and output prices. This approach is formally equivalent to the production function approach but utilizes different information.

discussed in detail in the chapter by Harley below. Between 1760 and 1800, Crafts and Harley estimate, total factor productivity "explained" about 10 percent of total output growth; in the period 1801-1831 this went up to about 18 percent. This seems rather unimpressive, but it should be kept in mind that growth is concerned with output per worker (or per capita). If we look at output per worker, we observe that for the period 1760-1830 practically the entire growth of per capita income is explained by technological change.[27] Economic growth was slow, as Harley and Crafts have shown, but what little there was is explained by the residual. Differences in the exact procedure are still not entirely resolved.[28] Precisely because growth per capita was so slow and there is little to explain, small differences in procedures and estimation will produce different residuals. For instance, Voth (1998) radically revises labor inputs and claims that because labor input per capita increased in the fifty years before 1800, the residual is extremely small and possibly negative. Coming from a different direction, Clark (below) has revised the growth of per capita income between 1760 and 1800 and finds it to be essentially zero. The apparent dominance of invention over abstention suggested by total factor productivity analysis, one of the most striking findings of the New Economic History, seems less secure now than it did a decade ago. Clearly it is unwarranted to expect that major technological breakthroughs will lead to more or less simultaneous increases in productivity. Most of the payoff to such breakthroughs occurs in the more remote future and is spread over a long period.

Identifying the residual with technological change, in any event, is far from warranted. The residual is a measure of our ignorance rather than of our knowledge. Any errors, omissions, mismeasurements, and aggregation biases that occur on the output and the input sides would, by construction, be contained in the residual. For instance, we simply do not know much about the flow of capital services. If horses or machines worked longer hours or factory buildings occupied for more than one shift, it is unlikely to be registered in our estimates as an increase in capital inputs. Moreover, changes in the *quality* of inputs would also be captured in the residual. If labor becomes more productive because workers are healthier or better disciplined, total factor productivity will increase though technology has remained unchanged. Furthermore, the residual is affected by market imperfections and external economies, economies or diseconomies of scale, changes in factor mobility, and so on. The residual is more than productivity change, and productivity change is more than technological change. At the same time, not all technological progress necessarily shows up in the residual.

A related literature has emerged in recent years concerning the question of “exogenous” vs. “endogenous” growth. Modern new growth theory, pioneered by economists such as Paul Romer, has tended to attribute a much larger portion of economic growth to endogenous factors, that is, factors that are part of the economist’s models. The sense of economists is that they prefer models that do not rely on unexplained and exogenous events, replacing them, as one recent paper has it, “with explanations of historical experience” (Greasley and Oxley, 1997, p. 935). Endogenous growth theory argues that technological progress is really

---

[27] The contribution of total productivity toward per capita output can be computed from data provided by Crafts (1985a, p. 81) and Crafts and Harley (1992, table 5).

	Per Capita Growth	Contrib. of Capital/ Labor Ratio	Contrib. of Resources per Capita Ratio	Total Contrib. of Nonlabor Inputs	Total Factor Producti-vity Growth	Producti-vity as a % of Total per Capita Growth
1760-1800	0.2	0.2*0.35 = 0.07	-0.065*0.15 = -0.01	0.06	0.14	70
1800-1830	0.5	0.3*0.35 = 0.105	-0.1*0.15 = -0.015	0.09	0.41	82

[28] Crafts and Harley themselves find somewhat larger contributions of capital and correspondingly lower contributions of productivity, which results from their procedures lumping capital together with land and thus overstating total input growth somewhat.

produced by the system, either by people getting better skills and education or by some capital good that brings it about. This view implies that the time series properties of industrial output will be quite different than the old exogenous growth models in which economic growth triggered by exogenous technological shocks eventually reverted back to a steady-state rate. In these models the output series does not exhibit persistence to shocks that is, it does not possess a unit root. An interesting debate has developed in the pages of the periodical literature on whether the time series of industrial output in Britain between 1780 and 1914 exhibits a unit root, the argument being that trend vs. difference stationarity presents a strong test of the kind of process that generates economic growth (Crafts, 1995a; Greasley and Oxley, 1997; Crafts and Mills, 1997). The idea is that if the series can be shown to be Difference Stationary, economic growth will be "endogenous" because a production function of the Romer type exhibits persistence and does not revert back to its trend. Trend stationarity, on the other hand, means that the growth process did not exhibit persistence: productivity shocks would, if not followed by others, peter out and the system required a constant infusion of new productivity-increasing technological advances if a technology-driven process of economic growth is to be sustained.

The econometric argument, thus far, is inconclusive. Nicholas Crafts has argued that at least some part of the growth was exogenous and that trend stationarity cannot be rejected. Others have re-examined their data and concluded the reverse. One problem is that too much ink is spilled on devising the right test for persistence and too little, some would say none at all, attention to the underlying data. For a wide range of goods the quantity indices used by all participants in the debate consistently understate the rate of growth and so tend to be biased. It is not clear whether such a bias would increase or decrease the likelihood of rejecting the unit root hypothesis, but it does mean that many of the tests have been run on flawed data. While a few products such as cotton and coal are thought to be of more or less uniform quality, improvement in the *quality* and nature of capital goods, from steam engines to cattle to streetlights, makes the series employed by Crafts and others a source of concern. Performing a conclusive Unit Root test on consistently measured output data is difficult enough, as Christiano and Eichenbaum (1990) pointed out years ago -- doing so with output data which are not and could not be measured in a consistent way strikes me as demanding too much credulity-suspension. To be sure, one can do such analyses on disaggregated series, and here too there is some evidence of persistence.

Moreover, exact economic meaning of such persistence is still rather unclear. It means that a sudden technological "shock" due to an invention of sorts will disturb the rate of growth of output for ever, which is what one would expect if the aggregate production function exhibited increasing returns. But what if technology is itself a Markov process in which present values depend on the past? In that case what looks like output responding forever to a sudden technology shock is nothing but output responding to new knowledge building on itself. Beneath the changes in technology there are changes in human knowledge not readily observed in production time series. That knowledge does not have to be scientific, as I argued it was not. But there was an accumulation of experience, of tricks, of practical engineering knowledge "what works" "what material is suitable" and "what tool is appropriate here." Unless the econometrician observes the underlying knowledge directly, she will mistakenly infer that it is the output that follows the "persistent" trajectory. We know something about how this knowledge was transmitted, diffused, improved upon, and eventually discarded. Very little of it had anything to do with formal education and other readily observable accumulation of human capital, least of all in Britain. The role of physical capital, as we shall see below, was equally ambiguous. Thus far, it remains very much an open question if the insights of the "new growth theory" can be applied to the Industrial Revolution (Crafts, 1996).

Technological change was only one event in the series of events that transformed Britain in this period. To what extent it caused the other changes or were caused by them remains a matter of interpretation. Whatever its exact role, it is impossible to provide any definition of the Industrial Revolution without it. Thus, if one insists on economic growth, capital accumulation, or changes in the organization of production as integral parts of the Industrial Revolution, it is difficult to separate them from the changes in technology. Even the most convinced detractors of the concept of the Industrial Revolution will concede two things. One is that although income per capita did not rise much between 1760 and 1830, it is hard to see how Britain could have sustained a more than doubling of its population while fighting a number of major wars had not its economic potential increased.[29] Moreover, the undeniable sustained growth that occurred in the British economy *after* 1830 would not and could not have occurred without the changes in technology in the previous seventy years.[30]

Secondly, most scholars agree that simple causal mechanisms will not explain something like the Industrial Revolution and that positive feedback and interactive path-dependent models will be needed if the phenomenon is to make sense. One example will suffice to convey this point: Many scholars emphasize commercial changes in this period and regard the rise of a national market and improvement in transport as causes of the changes in technology (Szostak, 1991). Adam Smith, writing before the Industrial Revolution or in its very early stages, had a view of economic development in which specialization and "the gains from trade" were at center stage. Yet improvements in technology subsequently fed back into improved transport, allowing even greater specialization and internal trade. Due to the inventions of John Loudon McAdam and Thomas Telford, improved roads and canals were constructed. Ships were built with planks cut by steam- or water-driven mills. Eventually, the high-pressure steam engine and the precision-tool industry, developed during the Industrial Revolution, were applied to land and sea transport, leading to changes in commerce that would have been unimaginable even to that inveterate optimist, Adam Smith. Thus gains from trade and specialization interacted with gains from technological progress, and such interactions led to a long and sustained path of economic development. Monocausal, linear models based on concepts of equilibrium or steady states will have difficulty doing justice to the historical reality.

To understand the phenomenon of the British Industrial Revolution, we have to ask two related questions: What were the causes of technological progress in Britain? What other elements permitted its society to adapt and transform itself to absorb the innovations and become the "workshop of the world."

3. The Causes of the Industrial Revolution

Why was there an Industrial Revolution? In this crude form the question is unanswerable. In more focused versions of the question some answers have been provided, and although full agreement is still remote, the discussion is one of the more lively in the historical literature. Examples of more focused formulations are: Why did the Industrial Revolution occur in Britain and not in France (or in the Netherlands, Germany, Spain)? Why did it start in the last third of the eighteenth century rather than, say, a century earlier? Can we find factors that should be regarded as "necessary preconditions" for the Industrial Revolution to have taken place? Historical causality in the analysis of an event as momentous as the Industrial Revolution is not likely to generate much consensus, since a multitude of different models can be devised to explain it.

To start with the last question, the notion that certain changes were a sine qua non for the Industrial Revolution has become increasingly difficult to maintain (Gerschenkron, 1962, pp. 31-51). Some factors present in Britain facilitated the Industrial Revolution and in this sense can be said to be causal. Others impeded its progress, and the Industrial Revolution proceeded in spite of them. The term *facilitated* does not mean, however, that there were any elements that were indispensable. After all, factors that were neither necessary nor sufficient for the outcome can still be thought of as causal. For instance, heart attacks cause deaths, though not all deaths are caused by them and not all heart attacks are fatal. Moreover, insofar as heart attacks are themselves caused by other factors, it is debatable to what extent they are ultimate causes or just "transmission mechanisms." The causal explanation of the Industrial Revolution runs into similar quandaries. Economic historians have increasingly come to concede that the positive effect that factor X had on the Industrial Revolution does not

---

[29] The population of England in 1760 was 6.1 million; in 1830, 13.1 million (Wrigley and Schofield, 1981, p. 534). The populations of Wales and Scotland grew at comparable rates.

[30] Gross domestic product per person-hour, which grew at 0.5 percent in the United Kingdom in the period 1785-1820, accelerated to 1.4 percent in the period 1820-1890. Real income per capita between 1820 and 1870 is estimated to have grown at 1.5 percent per annum (Maddison, 1982, pp. 31, 44).

entitle factor X to the status of "necessary factor." Counterfactual analysis has to be resorted to, at least implicitly, to assess the indispensability of the various elements.[31]

It is not even certain that the question Why did the Industrial Revolution occur in Britain rather than in some other country? is necessarily the best way to approach the material. For one thing, as we have already indicated, the Industrial Revolution was not so much a national as a regional affair. This has been stressed again recently in a collection devoted to this issue (Hudson, 1989). The regional argument was presented most cogently by Sidney Pollard (1981, 1985). Instead of dividing the European continent into "economies," Pollard prefers to look at regions that transcended national boundaries and shared a common economic fate. Thus one ought to prefer a comparison of, say, a region consisting of Lancashire and the West Riding of Yorkshire with a region consisting of southern Belgium and the northern *départements* of France.

Pollard's criticism of the national economy as the unit of analysis is not likely to remain unchallenged itself. The best arguments for the choice of nation-state as the appropriate unit of analysis are still in Kuznets (1966, pp. 16-19), who pointed out that nations share common heritages and histories, and thus people tend to be more interested in their national history than in regional histories. Moreover, a nation-state has a common government that is the major legislative and policy-making body, and insofar as it affects economic development, the unit under its jurisdiction should be the unit of analysis. The state was also, in most instance, the agency that collected economic statistics. Consequently, for better or worse, most of our data (e.g., foreign trade statistics, fiscal returns, price and wage figures) come on the national level.

It might be debated whether Britain was a unified economy or not on the eve of the Industrial Revolution (compare Crafts, 1985a, p. 3, and Szostak, 1991, p. 79, with Berg and Hudson, 1992). Yet it was certainly becoming more of one after 1760, and with the possible exception of the United Provinces, it was the most unified economy in Europe. Above all, it is hazardous to disavow comparisons of national units on account of *intra*national variances because the regional differences were themselves a *consequence* of the process of national development. As Rick Szostak (1991) has recently emphasized, no nation can devote itself entirely to one industry. With the improvements in transportation, interregional specialization became an inevitable part of the phenomenon that we are trying to analyze, namely the concentration of some industries in the Northwest of the country. The rise of the Yorkshire woolen industry was the mirror image and the "cause" of the demise of its counterpart in the West Country. The south of England remained largely unaffected by the Industrial Revolution because it specialized in agriculture.

A second criticism of the question Why was England first? has been raised in a pioneering paper by N.F.R. Crafts (1985b; see also Rostow, 1985). Quite simply put, Crafts's argument is that much of the Industrial Revolution was self-sustaining. In the extreme view, there is no point in asking why some nations underwent economic development and became rich while others remained poor and backward; it was all a matter of pure luck, a roll of the dice and in the limiting case causal analysis is useless. Much of the persuasiveness of this view depends on the accuracy of its premises. If we think of the Industrial Revolution as a sequence of strongly interrelated phenomena, it becomes indeed something close to a single event whose explanation may be beyond us. In reality, however, the set of facts we are trying to explain are to some extent independent events; by 1830 Britain had become a leader in a variety of industries, from papermaking to engineering to chemicals. If a coin is tossed once and heads comes up, there may be nothing to explain. However, if the coin is tossed dozens of times and heads comes up in every one of them, a closer look at the fairness of the coin would be called for. Much depends on how independent the events were; if they were strongly correlated, the "chance" explanation may hold true. If the correlation is weak, the plausibility of the "random-event" explanation is weakened. An

[31] Counterfactual analysis involves constructing a hypothetical world that never was. It is helpful in testing the hypothesis that factor X was a necessary condition in bringing about outcome Y; i.e., that in the absence of X, Y would not have taken place. Although the New Economic History is often credited with, or blamed for, introducing this mode of analysis, it has always been a staple tool of traditional historians. Thus Craig (1980, p. 1) begins his magisterial survey of German modern history: "It is certainly unnecessary to apologize for introducing Bismarck's name at the outset. If he had never risen to the top of Prussian politics, the unification of Germany would probably have taken place anyway but . . . surely not in quite the same way."

analogy from genetics is instructive here: We know that mutations are chance events, copying-errors in the DNA. Yet the number of mutations can be affected by radiation or mutagenic chemicals, and a sharp rise in the number of mutations would itself not be a chance event because mutations are unlikely to lead to further mutations. Can we, in economic history, define something equivalent to mutagens, environmental agents stimulating invention and innovation?

Landes, in his Tawney lecture (1994, p. 653), insists that big processes call for big causes and that models in which small initial differences are reinforced over time to produce an ever-widening gulf are unrealistic. Yet models of positive feedback have actually those characteristics, and it is easy to think of some historical processes in which increasing returns, induced technological change, frequency dependent processes, and the co-evolution of institutions and technology led to a spiral in which similar societies landed on quite different locations. Positive feedback can occur, for instance, when there are learning effects or under increasing returns (Arthur, 1994). In those cases technological change leads to lower prices, which could lead to the realization of scale economies, complementarities with other industries, demonstration effects, self-fulfilling expectations, bringing about even lower prices. Once the process had started, it fed on itself. Just as we have vicious circles in which backwardness breeds poverty and poverty breeds more backwardness, we have virtuous circles in which the reverse is true. More recent work in the theory of economic development has formalized much of this thinking (e.g., Matsuyama, 1991, and Arthur, 1994). If so, the role of contingency and accident in economic history may be far larger than people have supposed. In this approach, economic theory has to be complemented by insights from chaos theory: Comparatively minor differences in initial conditions can lead to major differences in historical outcomes. The Industrial Revolution in this interpretation was a "bifurcation point." Thus, as historians are gradually learning from evolutionary biology and chaos theory, accidents and contingency are increasingly seen under the right circumstances to matter a great deal, and fairly small historical events can set an economy off into one direction or another.[32] The key qualification is “under the right circumstances” -- did the structure of the economy switch from one of predominantly negative feedbacks to one in which positive feedback loops dominated? If so, can the Industrial Revolution be represented by such a model? Some thinking in modern systems-analysis seems to be moving into that direction. For instance, Stuart Kauffman has suggested that technological change can occur when the dynamic system's parameters are such that it is either in a subcritical region in which innovation is isolated and peripheral or in a supercritical one in which “new goods and services create niches that call forth the innovation of further new goods and services...such avalanches create enormous arenas of increasing returns because of the massive early improvements climbing learning curves along the novel technological trajectories” (1995, p. 296). If, as he suggests, the boundaries between the “regions” can be quite thin, it might be fruitful to think of the Industrial Revolution as a grand traverse from an economic system dominated by negative feedback and diminishing returns to one of positive feedback and sustained, indeed, explosive growth and innovation. Such interpretations must remain speculative, but they tend to underscore the central finding of the historians of technology, namely that the Industrial Revolution was not the beginning of economic growth or of technological progress, but it was the beginning of sustained, divergent, self-reinforcing and accelerating economic change. Up to a point, Pollard's (1996, p. 373) recent summary captures these dynamics: “the discovery of discovery itself became a commonplace and a driving force...as soon as inventions became widespread rather than isolated they provided mutual support for each other... Technology had 'taken off'”. Yet positive feedback stories of this kind have to contend with the problem that offsetting negative feedback can equally be discerned and eventually in the long run “nothing failed like success.” For that reason, we need to examine the details of Britain’s economy and society to understand why she came to play the role she did.

a. Geography

[32] Paul Mantoux realized this long ago when he pointed out that "only a negligible quantity of ferment is needed to effect a radical change in a considerable volume of matter" (1928, p. 103).

Britain's geographical advantages over other economies have often seemed to be good explanations for its economic success after 1750. In one book, a social historian states it as self-evident that "England is built [*sic*] upon an underground mountain of coal. Its exploitation was the motor-force in the revolution in production that created modern industrial society" (Levine, 1987, p. 97).[33] The belief that "geography is destiny" is an old and venerable one. Some major objections can be raised againt the view that places too heavy an emphasis on accidents of nature as causal factors. In part, the impact of such accidents is ambiguous. Resource availability plays a somewhat bizarre role in the historiography of technological progress. On the one hand, resource abundance is considered a blessing because it cheapens production and encourages the development of complementary techniques. On the other hand, many authors maintain that the challenge imposed by resource *scarcities* stimulates invention. Thus the deforestation of Britain is alleged to have led to a rise in timber prices, thus triggering Britain into adopting a novel and ultimately far more efficient set of techniques using fossil fuels. The evidence for this oft-repeated tale is far from convincing.[34] As a general statement, however, it suffers from the logical difficulty that the scarcity of natural resources and their abundance cannot *both* be regarded as stimulating factors for technological progress. At most, one can say that nature worked as a "focusing device," to use Nathan Rosenberg's felicitous term. Given a certain level of technological creativity, nature would direct this creativity in a certain direction, what Kuznets (1965, p. 91) has called a national bias in technological progress. Thus coal-rich Britain would focus on Newcomen engines, while coal-poor Switzerland would find economic success in precision-intensive low-energy industries such as watchmaking and engineering. Many other economies, rich or poor in resources, lacked the technological creativity and achieved little progress in this period. For a focusing device to work, there has to be a source of light.

Geography and physical endowment, like most other factors, are rarely either sufficient or necessary. Britain's geographic good luck was that it was an island and thus had not been successfully invaded since 1066. Being an island also provided it with access to a cheap form of transportation (coastal shipping). Yet being an island does not seem to have done much for Ireland, and good internal transportation was not very helpful to the Dutch economy in generating a phenomenon similar to the Industrial Revolution. Moreover, geography had to be aided by capital and technology. Patrick Verley (1997, p. 219) has recently recycled Babbage's (1835) calculations showing that in terms of naturally navigable rivers, England and France had similar proportional endowments, but once canals are taken into account, England (including Wales) had more than twice the internal waterways per square mile in 1820 and more than three times per capita. Britain's advantage in mineral wealth is equally problematic. Britain had coal and iron, but coal and iron were traded commodities; in 1794-1796 it imported £852,000 worth of iron and iron ore, mostly from Sweden. In the second half of the nineteenth century, it imported high-quality haematite ores from Spain. Coal, too, was traded, though its volume expanded only after 1830. Above all, it should be recalled that much of the Industrial Revolution depended on cotton and that raw cotton was entirely imported. Trade, it should always be remembered, liberates nations from the

---

[33] See also, for example, Parker (1979, p. 61). Coleman (1983, p. 443) even goes so far as to conclude that coal and iron were of greater consequence in determining the pattern of British industrialization than the existence of domestic industry. In making this statement he fails to apply to his own hypotheses the strict empirical standards he demands from others. E. A. Wrigley (1987; 1988, essay 4) has emphasized the importance of coal in the British Industrial Revolution, although his treatment is far more judicious.

[34] For an effective refutation of this argument, see Flinn (1959, 1978), Hammersley (1973). If it is true, as Hammersley (p. 609) notes, that wood as a crop could only use what to the landowner was marginal land and yielded returns far below those on pasture and other crops, it must have been the case that the "scarcity" of timber even in Britain was not too acutely felt. In his excellent survey of the issue, John Harris (1988) points out that the switch from charcoal to coal-based fuels in the iron industry in the second half of the eighteenth century is often believed to be the first such transition whereas in fact it was "virtually the last." Industries such as soapboiling, brewing, and glassmaking had switched to coal centuries earlier, and home-heating (the largest use for fuel) had become dependent on coal much earlier as well. While the iron industry itself may therefore not have been seriously constrained by the putative scarcity of charcoal, the benefits of abundant coal for Britain were larger than can be approximated by Darby's famous new technique of iron smelting. Yet this timing pattern also suggests that the nexus between Britain's fortunate endowment of coal and the Industrial Revolution is more complex than simple-minded models of geographical determinism suggest.

arbitrary tyranny of resource location. On the Continent, too, the evidence is mixed: Belgium, the first nation to adopt Britain's techniques, shared with her a wealth of iron and coal; Switzerland, a close second, had neither. Buying coal and iron from other economies added to industrial costs, but such additions were, on the whole, sufficiently small to be dwarfed by other cost differences. In other words, it is possible to accept Wrigley's (1987) view that substituting coal for wood was an important part of the economic transformation of Britain, without attributing undue significance to the geographical accident of the presence of coal in Britain. Coal had substitutes; coal-poor nations like the Netherlands and Ireland relied on peat for fuel, while the mountainous areas of Europe relied on water power for energy and flat windy areas in the Low Countries relied on wind power.. Such substitutions involved costs, of course, but the examples of Switzerland and New England prove that water power could provide an adequate energy base for a mechanized industry.

It could be maintained, however, that there were more subtle links between location and technological change. Small differences in resource endowment could set into motion chain reactions and steer a nation along a technological trajectory quite different from one that would have been followed in the absence of those resources. Britain's use of coal did not only help by providing cheap fuel; it focused Britain's attention on the solution to certain technological problems: pumping, hoisting, and mineral-exploration, which then spilled over to other industries (Cardwell, 1972, p. 73). Shipping, too, generated externalities in sawmills, carpentry, instrumentmaking, sailweaving, and so on. Yet in a deeper sense such mechanical descriptions are unsatisfactory since they describe opportunities; but clearly these opportunities were neither necessary nor sufficient conditions for success. Maritime Holland was not able to use its shipbuilding sector as a gateway into the Industrial Revolution.

In recent years, a growing number of scholars have followed the lead of Eric Jones (1988) in arguing that the Industrial Revolution was the culmination of a long process of modernization that started in Britain many centuries before (though opinions vary to when, exactly, this process started. The most influential economic historian of British medieval agriculture (Campbell, 1997) maintains that by the thirteenth and fourteenth centuries Britain was a market economy in which production decisions were sensitive to factor and commodity prices. Gregory Clark, in a number of papers (1997) as well as the essay below, has argued that medieval agriculture was as productive and sophisticated as British agriculture was on the eve of the Industrial Revolution and that markets for grain were well-functioning. Foreign travelers visiting Britain commented in living colors about the luxury and extravagance of British living standards in the last third of the eighteenth century (McKendrick, 1982a, pp. 9-10). Graeme Snooks (1994) has argued forcefully that economic growth was not unique to the period of the Industrial Revolution and that by the late seventeenth century Britain was an advanced and sophisticated economy.[35] In pointing this out, these scholars are joining the venerable company of MacFarlane (1978) who was one of the first scholars to pinpoint the beginning of Britain's modernity to the late middle ages. It is clear by now that far from being a "traditional" and "static" society, Britain was on the eve of the Industrial Revolution a country of sophisticated markets, in which profit-hungry homines economici did what they are supposed to do to help a country develop. Yet while this does explain Britain's wealth on the evolution of the Industrial Revolution, it raises as many difficulties as it solves. Is it so obvious that an urbanized, literate, market-oriented society leads inevitably to an Industrial Revolution? The Dutch economy, as De Vries (1973) and more recently De Vries and Van Der Woude (1995, pp. 798-806; 1997, pp. 693-710) have argued, had

---

[35]Snooks's (1994) belief in pre-modern growth is based essentially on his comparison between the income per capita he has calculated from the Domesday book (1086) and the numbers provided by Gregory King for 1688. While such computations are of course always somewhat worrisome (what, exactly, does it mean to estimate the nominal income of 1086 in the prices of 1688 given the many changes in consumption items?), the order of magnitude provided by Snooks (an increase of real income by 580 percent) may survive such concerns. Medievalists tend to agree with the occurrence of economic growth in Britain, though their figures indicate a much slower rate of growth, about a 111 percent growth rate between 1086 and 1470 (Britnell, 1996, p. 229), which would require more economic growth in the sixteenth and seventeenth centuries than can be justified to square with Snooks's numbers. Engerman (1994b, p. 116) assesses that most observers will agree with Snooks's view that by 1700 England had a high level of per capita and was in a good position to "seek the next stage of economic growth." Yet clearly he is correct in judging that "modern" economic growth (prolonged, continuous, rapid) did not begin until the early nineteenth century.

many elements of modernity and yet turned out to be one of the last economies to jump the bandwagon of modern manufacturing in Western Europe, whereas Switzerland, a relatively remote and simple highland economy had by 1850 a progressive modern sector.

Nor is it so obvious that income and wealth are positive feedback processes in which the rich get richer so that the Industrial Revolution can be seen as an example of economic divergence. Certainly, within the relatively small group of Atlantic economies, the past two centuries have shown a process of *con*vergence in which the backward economies managed to catch up with the leaders, so that within this group a process of income compression is clearly visible. Even less can we speak of a connection between income and technological creativity. Rich capitalist countries may be technological leaders for a long period, as Britain undoubtedly was. Yet more often than not, such leadership was eventually lost as it was in Britain's case and later in Germany's. Where Historical Accident did play an important role was in the coincidence of the British Industrial Revolution with the French Revolution and the Napoleonic Wars. Whereas the wars probably slowed growth and technological progress down everywhere, they did so far more seriously on the Continent than in Britain. Western technology emerged in a Western European club of economies, and while Britain was the undoubted leader, in its absence some other European economy would have played this role. The quarter of a century of turmoil and destruction that plagued the Continent after 1790 deepened and accentuated Britain's leadership and were especially hard on the Dutch economy, Britain's closest economic rival (Mokyr and Buyst, 1990). Yet it remains true that its technological leadership was already well-established by 1790, and while the diffusion of technological change and Continental industrialization may well have been faster in the absence of the French Revolution and its aftermath, there still would have been a British Industrial Revolution.

b. Technological Creativity[36]

If it is agreed that at the base of the Industrial Revolution lay something we call technological creativity, some speculation about the factors responsible for it is in place here. To start with, Britain seems to have no particular advantage in generating macroinventions; a large number of them were generated overseas, especially in France. Steampower and cotton technology were British inventions, but many of the other examples cited previously were imported: Jacquard looms, chlorine bleaching, the Leblanc soda-making process, food canning, the Robert continuous paper-making process, gaslighting, mechanical flaxspinning.

Any period of successful technological creativity requires both fundamental breakthroughs and small, incremental, often anonymous improvements that take place *within* known techniques. The key to British technological success was that it had a *comparative* advantage in *micro*inventions. This may seem unorthodox to those who think of the milestones set by Richard Trevithick, Richard Arkwright, and Henry Cort, but it should be recalled that it is possible to have an absolute advantage in both areas yet a comparative advantage in one, although it is not altogether clear whether Britain had an *absolute* advantage in macroinventions.

Evidence for the statement that the British comparative advantage was in improvement and not in originality comes in part from contemporary sources. In a widely cited comment, a Swiss calico printer remarked in 1766 that for a thing to be perfect it has to be invented in France and worked out in England (Wadsworth and Mann, 1931, p. 413). In 1829 the engineer John Farey stated that the prevailing talent of English and Scotch people is to apply new ideas to use and to bring such applications to perfection, but they do not "imagine" as much as foreigners (Musson, 1975, p. 81). Continental Europeans felt frustrated, reflecting Leibniz's prophetic words, written in 1670: "It is not laudable that we Germans were the first in the invention of mechanical, natural, and other arts and sciences, but are the last in their expansion and betterment" (cited in William Clark, 1991). A test of the hypothesis that Britain had a comparative advantage in microinventions is in the establishment of net trade directions. Economies tend to specialize in the areas in which they have a comparative advantage. The British economy, roughly speaking, was a net importer of macroinventions and exporter of microinventions and minor improvements. We should of course look at this specialization as a broad central tendency, but in rough

[36] The following paragraphs draw heavily on Mokyr (1990a, 1992a).

lines the distinction stands up. Britain took its major inventions where it could find them, but whatever it borrowed it improved and refined.[37]

On the eve of the Industrial Revolution, Britain was neither a scientific leader nor could it boast of a particularly effective education system. As David Mitch explains in more detail in a later chapter, British education was at its best outside the schools, and Britain trained most of its mechanics and engineers by its age-old apprenticeship system without introducing much formal schooling. In a sample of 498 applied scientists and engineers born between 1700 and 1850, 91 were educated in Scotland, 50 at Oxbridge, and 329 (about two-thirds) had no university education at all (Birse, 1983, p. 16). Yet these people thirsted for technical, applied, pragmatic knowledge, the knowledge of how to make things and how to make them cheap and durable. A few of them were educated at Scottish universities or dissenting academies, but many were self-taught or had acquired their knowledge through personal relations with masters, libraries, itinerant lecturers, and mechanics institutes. By the middle of the nineteenth century, there were 1,020 associations for technical and scientific knowledge in Britain with a membership that Inkster estimated conservatively at 200,000 (Inkster, 1991, pp. 73, 78-79).

For Britain in this period, this system clearly delivered. It produced some of the finest applied engineers in history. As long as technological advances did not require a fundamental understanding of the laws of physics or chemistry on which they were based and as long as advances could be achieved by brilliant but intuitive tinkerers and persistent experimenters, Britain's ability to create or adapt new production technologies was supreme. Most inventors and engineers were dexterous merchants or enterprising craftsmen whose technical ideas were often the result of luck, serendipity, or inspiration even if the successful completion of the innovative process required patience, determination, and confidence.

Moreover, some of the industries in which Britain had specialized before 1760 required skilled mechanics. Clock and instrument making, shipbuilding, iron making, printing, wool finishing, and mining required a level of technical skill that came in handy when new ideas had to be translated from blueprints to models and from models to real commodities. John Wilkinson, it is often remarked, was indispensable for the success of James Watt, because his Bradley works had the skilled workers and equipment to bore the cylinders exactly according to specification. Mechanics and instrument makers such as Jesse Ramsden, Edward Nairn, Joseph Bramah, and Henry Maudslay; clock makers such as Henry Hindley, Benjamin Huntsman, and John Kay of Warrington (not to be confused with his namesake, the inventor of the flying shuttle, who was trained as a reed and comb maker), engineers such as John Smeaton, Richard Roberts, and Marc I. Brunel; ironmasters such as the Darbys, the Crowleys, and the Crawshays; chemists such as John Roebuck, Alexander Chisholm, and James Keir were as much part of the story as the "superstars" Arkwright, Cort, Crompton, Hargreaves, Cartwright, Trevithick, and Watt. Below the great engineers came a much larger contingent of skilled artisans and mechanics, upon whose dexterity and adroitness the top inventors and thus Britain's technological success relied. These unknown but capable workers produced a cumulative stream of anonymous and small but indispensable microinventions without which Britain would not have become the "workshop of the world." It is perhaps premature to speak of an "invention industry" by this period, but clearly mechanical knowledge at a level beyond the reach of the run-of-the-mill artisan became increasingly essential to create the inventions associated with the Industrial Revolution. Dozens of scientific journals and the published transactions of scientific societies had appeared by 1800, most of them after 1760 (Kronick, 1962, p. 73). A widespread thirst for knowledge about "natural philosophy and its relation to the useful arts" penetrated Britain down to the small towns of the kingdom where itinerant lecturers were in much demand. The people who worked in applying the principles of physics, chemistry, and biology in their daily work were thirsty for innovations. In this milieu *micro*inventions,

---

[37] The case of chlorine bleaching is revealing here. The Swede Karl Wilhelm Scheele and the Frenchman Claude Berthollet clearly produced the original breakthrough, but the commercial value of the idea was recognized by James Watt (whose father-in-law, James McGrigor, was a Glasgow bleacher), and a series of British chemists and entrepreneurs set out to improve on the original invention (Musson and Robinson, 1969, pp. 251-337). The definitive improvement came when a Scottish bleacher Charles Tennant replaced potash with slaked lime as the solution in which chlorine was absorbed. Chemical bleaching, a Continental macroinvention, was made into bleaching powder, a British improvement.

the gradual improvement of pathbreaking ideas, will prosper. In the early stages of the Industrial Revolution this ability was the key to Britain's technological success.

It is of course a truism that advantages in skilled labor were a matter of degree, not an absolute. France, Germany, and the Low Countries had their share of able and innovative engineers. But degree is everything, and in the early nineteenth century Britain tried, in vain, to keep the secret of its success by prohibiting the exportation of machines and the emigration of skilled mechanics. Yet as it had imported macroinventions, it exported microinventions and the people who implemented them. The engineers who spread the new technologies to the Continent after 1800 had names like Cockerill, Hodson, Ainsworth, Douglas, and Holden. Insofar as trade patterns reveal comparative advantage, these patterns reveal Britain's technological superiority. Explaining this superiority is a different matter: Landes (1969, pp. 61-64), who was one of the first to call attention to Britain's advantage in mechanics and technicians, spoke of the question of British mechanical skills as "mysterious." Clearly, any explanation will have to take us beyond the narrow boundaries prescribed by economic science.

c. Social and Institutional Factors

It is easier said than demonstrated that Britain had the "right kind of society" to have an Industrial Revolution. After all, what exactly do we mean by social preconditions to industrialization and how do we demonstrate the proposition that they were important? One way to approach the subject is through the concept of a "hierarchy of values." Each society defines in some way the criteria of success. Success means access to certain nonmarket goods such as political offices, membership of social clubs, being plugged into information networks, and in general earning respect from people whose opinions matter. Social status and prestige are always and everywhere *correlated with* economic success but are almost never *identical* with it. In many societies the causation ran from non-economic success to enrichment; victorious Roman generals were rewarded by remunerative governorships. One key to the economic success of a society is essentially the degree to which social respect not only correlates with economic success but is caused by it.[38]

The most complete and persuasive attempt to provide a social explanation of the Industrial Revolution based on this idea has been provided by Perkin (1969). Perkin dates the creation of the type of society that was most amenable to an Industrial Revolution to the Restoration of 1660 and the social and political changes accompanying it.[39] He points out that the principle upon which society was established following the Civil War was the link between wealth and status. Status means here not only political influence and indirect control over the lives of one's neighbors but also the houses to which one was invited, the partners that were eligible for one's children to marry, the rank one could attain (that is, purchase) in the army, where one lived, and how one's children were educated. In Perkin's view, the quality of life was determined not just by "consumption," as usually defined by economists, but by the relative standing of the individual in the social hierarchy. Whether this social relativity hypothesis is still a good description of society is an open question, but a case can be made, as Perkin does, that it is an apt description of Britain in the eighteenth century. Perkin cites a paragraph from Adam Smith's *Theory of Moral Sentiments*, which economists -- always a bit selective in what they learned from the Master -- have been ignoring at their risk:

[38] Economic theorists have belatedly rediscovered this rather obvious fact and have tried to formalize it. See Cole, Mailath, and Postlewaite (1992, 1995). Their model of “social norms” distinguishes between a “wealth is status” norm and an aristocratic (“birth is status”) norm. They show, among others, that the former norm encourages savings to facilitate social climbing. When wealth is not directly observable, individuals may engage in conspicuous consumption, signaling their wealth, but such consumption by itself destroys part of the wealth.

[39] Some social historians argue that the changes started much earlier. Alan MacFarlane (1978, pp. 199-201) explicitly dates the beginning of English "modern society" to some point before the Black Death.

> To what purpose is all the toil and bustle of the world . . . the pursuit of wealth, of power, and preeminence? Is it to supply the necessities of nature? The wages of the meanest labourer can supply them. . . . What then is the cause of our aversion to his situation? . . . Do the rich imagine that their stomach is better, or their sleep sounder in a palace than in a cottage? The contrary has so often been observed. . . . What are the advantages [then] by that great purpose of human life which we call bettering our condition? . . . It is the vanity, not the ease of the pleasure, which interests us. But vanity is always founded upon our belief of our being the object of attention and approbation. The rich man glories in his riches, because he feels that they naturally draw upon him the attention of the world. . . . Everybody is eager to look at him. . . . His actions are the objects of the public care. Scarce a word, scarce a gesture can fall from him that is altogether neglected. In a great assembly he is the person upon whom all direct their eyes. . . . It is this, which . . . renders greatness the object of envy and compensates . . . all that toil, all that anxiety, all those mortifications which must be undergone in the pursuit of it (Smith, 1759, pp. 50-51).

In Perkin's own words, "To the perennial desire for wealth, the old society, [i.e., Britain after 1600] added more motivation which gave point and purpose to the pursuit of riches. Compared with neighbouring and more traditional societies it offered both a greater challenge and a greater reward to successful enterprise. . . . the pursuit of wealth *was* the pursuit of social status, not merely for oneself but for one's family" (Perkin, 1969, p.85).[40] Examples are not hard to find: The riches accumulated by Richard Arkwright in cotton spinning bought him not only all the comforts that money could buy but also a knighthood and the office of sheriff of the County of Derby. Other cotton manufacturers who rose to high office included Robert Peel, Sr., who became an MP and whose son became prime minister. Brewers, paper makers, potters, and iron masters became barons, earls, MP's, and castle dwellers. Men of business could, through money, "advance in rank and contend with the landlords in the enjoyments of leisure, as well as luxuries," as Malthus (1820, p. 470) put it.[41]

Perkin's insight is important because it underlines a basic point often overlooked by economists trying to understand entrepreneurial behavior. It is almost always true that an easy opportunity to earn money will not be passed over by a rational individual. Moreover, if there is a divergence of opinion about the expected profitability of an opportunity, one should expect the optimists to replace the pessimists. Unexploited opportunities to quick gains will rapidly disappear. There were opportunities to make money during the Industrial Revolution, but few were quick and easy. Almost all major entrepreneurial figures took enormous risks, worked long and hard hours, and rarely enjoyed the fruits of their efforts until late in life or enjoyed them vicariously through their descendants. Entrepreneurship will be more forthcoming if the rewards of money exceed the costs of risk bearing, hard work, and postponed gratification. Perkin's thesis stresses the benefit side in this equation; in Britain money bought more than just comfort. Money acquired in commerce or industry was less tainted by the stigma of being "nouveau riche." The example set by the elite (the landowning gentry and aristocracy) profoundly influenced the values and attitudes of those who aspired to be like them. In Britain, far more than

---

[40] Perkin anticipated here the interesting work of Fred Hirsch (1976), who, although not concerned with history, sets up a framework that complements Perkin's. Hirsch distinguishes between material goods--i.e., ordinary commodities--and positional goods of which there are by definition a constant amount. Examples of the latter are social prestige, political power, and symbols indicating one's relative position. Markets for material goods tend to be well developed, so material wealth provides easy access to them. Markets for positional goods are less well developed. The more efficient the markets for positional goods, the easier it is to acquire them by the means of acquiring wealth or to lose them by the lack thereof. Therefore, relatively efficient markets for positional goods should strengthen the incentive to get rich (increase the marginal utility of income) and make the toil and risks of entrepreneurship more worthwhile.

[41] Local studies confirm the importance of wealth as a determinant of status. Urdank, in his study of Gloucestershire, found that "between 1780 and 1850 wealth had become a more obvious criterion for defining status than in the past, so much so that men with the humblest occupations might call themselves 'gentlemen' if the size of their personal estates seemed to warrant the title" (Urdank, 1990, p. 52).

on the Continent, a materialist element had come to dominate these values. As Landes (1969, p. 70) put it, "The British nobility and gentry chose to meet the newcomers on middle ground: they affirmed their distinction of blood and breeding; but they buttressed it with an active and productive cultivation of gain."

Still, some empirical questions have to be answered before the connection between wealth and status can be accepted as one explanation of England's success.[42] Was the correlation between wealth and social status stronger in Britain than elsewhere? Holland was an urban, capitalist, bourgeois society, indicating that having the "right kind of society" is not a sufficient condition for a successful Industrial Revolution.[43] But what about France? In the eighteenth century aristocratic titles could be bought, and much of the nobility was a *noblesse de robe*, i.e., of bourgeois origins. Was the aversion to parvenus among the upper class stronger in France than in England? Although the latter question cannot readily be answered, there were two important differences between the two countries in this respect. First, in France, too, money could enhance social status, but the respectable local country gentleman who ran the affairs of the parish was a wholly British institution. Second, in France social status was often literally bought. The price of a noble title reflected a tax exemption, so that the sale of titles was not a one-way street by which the crown soaked up wealth. But nobility implied high standards of consumption in the noblesse oblige tradition. In England, by contrast, wealth was correlated with influence and respect, but one did not necessarily have to part with the former to attain the latter.

Furthermore, Perkin's logic implies an almost dialectical dynamism of the supply of entrepreneurship. If merchants and manufacturers made money in order to buy themselves or their descendants the good life of the country squire, the ranks of the entrepreneurial class would be constantly depleted. Upward mobility by means of wealth thus also led to the eventual destruction of the entrepreneurial class. Having attained their new status, the new elite tended to slam the door shut to further entrants. This "gentrification" of the commercial and industrial class, which has been blamed for the decline of Britain's leadership in the Victorian age (Wiener, 1981), seems a logical extension of Perkin's thesis. Because the debate on the "failure" of Victorian Britain lies outside the scope of this volume, this implication cannot be pursued here.

Society is, of course, more than attitudes and mind-sets. Its importance lies above all in the institutions within which economic activity takes place. Some institutional setups are more suitable for technological change than others, and although institutions eventually may respond to economic and political needs and pressures, these responses are sufficiently sluggish to allow us to point to institutions as a "causal" factor in economic development. Institutions defined property rights and thus the rate of return on inventive and entrepreneurial activity. This has been stressed by North (1981, 1990). In North's interpretation, property rights and incentives are the crucial elements in the story. He stresses (North, 1990, p. 75) that patent laws and other institutions raised the rate of return on innovation and thus stimulated the process of technological progress. Britain's patent law dates back to 1624, whereas France and most of the rest of the Continent did not have such laws on the books until after 1791.

The exact role of the patent system in Britain's Industrial Revolution is hard to determine. A patent is only one way of encouraging a potential inventor to spend time and money on the uncertain road to success. The French government, for instance, awarded pensions through the Royal Academy and through so-called *privilèges* (administered by the king), which were also intended to encourage invention (MacLeod, 1991). North overrates the effective protection that the British patent system provided to inventors; court decisions in infringement cases tended to be unsympathetic to inventors, and patents were overthrown on minor technical points such as scribes omitting one line. In some cases, financial success came without patent protection, as in the case of Richard Arkwright. The court's invalidation of his patent did not stop him from becoming extremely rich. In other cases, when inventions failed to be patented or when patents were lost because of technicalities, inventors were rewarded by Parliament in recognition of their social value. The mule's inventor, Samuel Crompton, and

---

[42] Perkin's further attempts to explain the timing of the Industrial Revolution in terms of population growth and demand are far less successful. Some of these issues will be dealt with later in this chapter.

[43] For economic explanations of the Netherlands's failure to industrialize, see Mokyr (1976a) and Griffiths (1979).

the power loom's inventor, Edmund Cartwright, were both the beneficiaries of Parliament's gratefulness. Moreover, patents and infringements of them led to endless court battles that sapped the energy and resources of technologically creative people. Arkwright and his sometime partner, Jedediah Strutt, spent much time in courts defending their patents. Some innovators, such as John Kay, the inventor of the flying shuttle, and the Fourdrinier brothers, who pioneered the paper-making machine, were ruined by litigation. In many cases, inventors decided to protect their monopoly rents by keeping their inventions secret. If "reverse engineering" was not likely or if the inventor could make his money by employing his machines to produce a final output rather than by selling capital goods, this was often tried. Yet secrecy had its risks: Industrial espionage was an ever-present danger.[44]

The effects of patents on the rate of innovative activity is further clouded by the fact that the patentability of innovations differed a great deal from industry to industry. Christine MacLeod has estimated that nine out of ten patents arose in industries that saw little innovation and concludes that patents were related to technological change in an erratic and tangential manner and were more closely associated with "emergent capitalism" than with inventiveness (1988, pp. 145, 156-157). Moreover, patent protection, as is well known, is a double-edged sword. If a patentee is a monopolist, the invention's diffusion will be retarded and the industry will grow at a slower pace, unless the inventor's firm can expand as fast as the industry as a whole. The fundamental dilemma in the economics of technological change is that there is a trade-off between generating an invention and its diffusion.[45] The more monopoly protection is used to encourage invention, the slower its adoption and thus its social benefits. Patents imply the choice of a particular point on this trade-off; so do alternative arrangements. Moreover, patents may have had a mixed effect on invention itself; in some cases owners of wide-ranging and vague patents used their power to close avenues they deemed undesirable or potentially competitive. The best-known example of that in the period of the Industrial Revolution is Watt's use of his patent to resist the development of high-pressure steam engines.

All the same, the importance of the patent system for Britain's technological success cannot be wholly dismissed by these objections. As Adam Smith was the first to point out in his *Lectures on Jurisprudence*, patents alone preserved some automatic correlation between the value of an invention and the return received by the inventor. The French system of rewards administered by a governmental committee made the return on invention dependent on political clout more than on the test of the market (Gillispie, 1980, pp. 459-478). Moreover, incentives refer to potential inventors' *ex ante* expectations of being financially rewarded if they were successful. Disappointments and lawsuits were relevant to further technological progress only to the extent that they discouraged others. By definition, each patent is inherently different from every other one, and so the failure of an inventor to secure a return on his efforts may not have necessarily indicated to others that their fate would be the same. The desire to patent new inventions did not weaken during the Industrial Revolution. Goethe may have been somewhat naive when he wrote that the British patent system's great merit was that it turned invention into a "real possession, and thereby avoids all annoying disputes concerning the honor due" (cited in Klemm, 1964, p. 173). Yet in 1845 the Swiss industrialist Johann C. Fischer concluded that "the system of patents so early introduced there may well have . . . been responsible for manufactured goods possessing so high a degree of perfection." Britain's greatest post-1830 inventor, Henry Bessemer, believed that "the security offered by patent law to persons who expend large amounts of money in pursuing novel inventions, results in many new and important improvements in our manufactures" (Bessemer, [1905] 1989, p. 82). Not all inventors concurred with this view, but if enough of them saw it this way, the British patent system deserves some credit. H. I. Dutton

[44] Richard Roberts, one of the leading engineers of his time, felt that "no trade secret can be kept very long; a quart of ale will do wonders in that way" (cited by Dutton, 1984, pp. 108-111).

[45] The efficient solution maximizing the social savings could be attained if the patentee could license his patent out and earn royalties equivalent to the monopoly rent. Yet setting the correct prices and monitoring the arrangements were a major difficulty. MacLeod (1991) concludes that only after 1800 did British patentees learn to exploit licenses more profitably, and even then only a tiny minority mastered the art at the cost of extensive litigation.

(1984, p. 203) has argued that for many inventors patents were the only means by which they could appropriate a sufficient return for their effort and that patents thus provided security in an exceptionally risky activity. The patent law was often poorly defined and the courts unfriendly to inventors, but it remained in most cases the best incentive for inventive activity. Dutton argues that the patent laws were a "slightly imperfect" system that created an ideal system in which there was enough protection for inventors to maintain an incentive for inventions, yet was not so watertight as to make it overly expensive for users. If inventors systematically overestimated the rate of return on inventions by not fully recognizing the weaknesses of the patent system, they would have produced more innovations than in a world of perfect information.

d. Government and Politics

British political institutions differed greatly from those of most European countries, and recent thinking by economists has tended to place considerable emphasis on political elements. Some of the differences are obvious: Despite the fact that the Industrial Revolution coincided with two major wars, there was no fighting on British soil, and except for a few serious but localized riots and an abortive uprising in Ireland, Britain was spared the turmoil and turbulence of the Continent after 1789. The need to allocate resources to the war effort involved a substantial effort on the part of Britain, and the disruptions of trade and the disequilibria caused by the wars and blockades clearly slowed down the development of the British economy (Crouzet, 1987; Mokyr and Savin, 1976). Yet as already noted, these disruptions were far more deeply felt on the Continent, and the wars widened the gap between Britain and its main competitors in Europe.

Douglass C. North (1981, pp. 147, 158-170) has argued that the British Industrial Revolution was facilitated by better-specified property rights, which led to more efficient economic organization in Britain. The link between property rights and economic growth consists of the greater efficiency in the allocation of resources resulting from the equalization of private and social rates of return and costs. Property rights in innovation (patents and trademarks), better courts and police protection, and the absence of confiscatory taxation are examples of how the same phenomenon could raise the rate of innovative activity and capital accumulation.[46] North points out that well-specified property rights are not the same as laissez-faire. The former were by far more important because they reduced transaction costs and thereby allowed more integrated markets, higher levels of specialization, and the realization of economies of scale. Britain on the eve of the Industrial Revolution was far from a laissez-faire economy, but the net effects of the policies and regulations on the Industrial Revolution remain a matter of dispute. What is clear is that by the time of the Industrial Revolution Britain's government was one of, by, and for private property. Such property rights should be contrasted, not with chaos and anarchy, but with traditional and customary rights, often disputed, undocumented, and hard to establish. O'Brien (1991, p. 6) insists that in the eighteenth century the British government came down hard and persistently “in favour of property and against customary rights.” Yet as the case of the Dutch Republic demonstrates, a well-defined system of property rights too, was not sufficient cause for an Industrial Revolution.

More recently, North and Weingast (1989) survey the institutional changes that occurred in Britain in the wake of the Revolution of 1688, in which wealth holders increased their grip on power, and the government was put on a sound fiscal footing and committed itself to respect the existing distribution of property rights. They pose their question starkly: Had there been no Glorious Revolution in 1688, or had the Stuarts won, would there have been an Industrial Revolution? (p. 831). Although they wisely confess ignorance as to how to set up the counterfactual, they point to secure contracting and property rights as a precondition for specialization and impersonal exchange. Without denying the importance of secure contracts as a precondition for allocative

---

[46] Confiscatory taxation during the French Revolution took three main forms in Europe at this time. First, there was outright confiscation of property, such as the Church lands and the assets of *emigrés* expropriated during the French Revolution. Second, raising armies by conscription, as practiced by France, constitutes a de facto confiscation of labor. Third, the French government (and subsequently the Dutch) defaulted on their debts by reducing interest payments on debts by two-thirds. Moreover, many innovators who had been voted pensions by the *ancien régime* were denied their payments, and some of them, like Nicholas Cugnot, the inventor of a steam-powered wagon, died in poverty. Nicholas Leblanc, the inventor of the soda-making process, tried in vain to make the revolutionary regime recognize his rights on his invention and in the end committed suicide.

efficiency, one could object that the Industrial Revolution was not first and foremost an example thereof. It was an example of Schumpeterian disequilibrium, in which the main dynamic elements came from innovation and rebellion against the status quo. Invention and change may well have come at the expense of an efficient allocation of resources and more static equilibrium conditions. Moreover, the impact of financial markets, the development of which is emphasized by North and Weingast, on the Industrial Revolution is still very much the subject of debate. Finally, it seems unwarranted to imply that before the Glorious Revolution contracts and property rights in Britain were insecure. By taxing according to prespecified and well-understood rules, and by gradually abandoning the Tudors' and Stuarts' reliance on monopoly rights as a source of crown revenues, the post-1689 regime continued a trend that had begun long before and was certainly well established by the Restoration of 1660.

What kind of government helped bring about something like the Industrial Revolution? O'Brien (1991) carefully credits them with sustaining legal and political conditions which turned out on balance to be conducive to bring about "the most efficient industrial market economy in Europe." Yet any policy objective aimed deliberately at promoting long-run economic growth would be hard to find in Britain before and during the Industrial Revolution. To be sure, certain statutes aimed at encouraging progress, from patents to prohibitions on the emigration of artisans and the exports of machinery remained on the books until deep into the nineteenth century. But many of these acts were directed toward increasing the economic rents of a successful political lobby and their overall impact on technological progress at best ambiguous. In terms of its spending and its attention, the British government was clearly still largely mired in colonial and foreign policies. In Britain the public sector by and large eschewed any entrepreneurial activity. During the heyday of the Industrial Revolution, even social-overhead projects that in most other societies were considered to have enough public advantages to warrant direct intervention of the state were in Britain left to private enterprise. Turnpikes, canals, and railroads were built in Britain without direct state support; schools and universities were private. The promotion of the "useful arts" (that is, applied science and technology) was largely left to voluntary organizations and local societies. Even the less invasive forms of state support, like the policies of William I of Orange in the Low Countries or the Saint-Simonians in France during the Second Empire, were notably absent in Britain. Until the end of the nineteenth century, the British government clearly was reluctant to invade what it considered to be the realm of free enterprise.

Providing a conducive environment in which business could operate to maximum effectiveness might seem a reasonable task for a modern economist and clearly contemporaries realized this.[47] Yet as far as contract enforcement and dispute arbitration were concerned, readers of *Dombey and Son* will not be surprised by O'Brien's (1991) assessment that "the English legal system did not offer speedy, cheap, and economically efficient" solutions to commercial disputes. Much of the system was, no doubt, self-enforcing. Reputation, moral codes, fear of stigma and religious scruples could not altogether prevent the occasional Uriah Heep from behaving opportunistically, yet they must have been sufficiently rare so that they did not constitute a brake on the economy. Informal arbitration within trade associations, chambers of commerce and a variety of other institutions cleared most disputes. When the state failed in providing public goods, spontaneous corrective action was common. Middle class associations to help in the apprehension of criminal and private detective agencies start appearing after 1770 (O'Brien, 1994, p. 217). Similarly, the private sector was able to correct the errors of the **public** one in the other area widely regarded today as a main function of the State: the supply of money. The management of the money supply in the eighteenth century is widely thought to have been inadequate with much inconvenience arising from the shortage of coins of relatively small denomination and the supply of legal money to be inadequate for the needs of an expanding economy (Ashton, [1955], 1972, p. 167). Copper coins

[47] Adam Smith in his chapter on the "Expense of Justice" in his *Wealth of Nations* realized that "the acquisition of valuable property ...necessarily requires the establishment of civil government. Yet he missed the point made by modern economics when he noted that "the benefit of the person who does the injury [to property] is often equal to the loss of him who suffers it." The social deadweight losses of uncertain and poorly enforced property rights imply that the gains are lower than the losses because injuries and uncertainty will affect the allocation of resources.

were particularly in short supply. In the second half of the eighteenth century, lower denomination coinage was largely left by the government to the private sector (Sargent and Velde, 1998). Enterprising industrialists and entrepreneurs alleviated the problem by creating more means of exchange, both imitated coinage and inside money (such as Bills of Exchange and banknotes). Indeed, it is significant that the small change shortage was eventually relieved by the adaptation of steam power to minting by no less a figure than Matthew Boulton in 1787, which spread quickly among private minters and eventually landed Boulton a contract to make copper coins for the government. In short, it may be that the greatest merit of the Hanoverian State was that while it did not do an outstanding job in providing the kind of public goods and institutional infrastructure needed in an expanding economy, it did not prevent the private sector from stepping in.

The success of Britain in the late eighteenth century is perhaps surprising to those who firmly believe that taxes and government debts are a guarantee of economic disaster. In 1788, British GNP per capita is estimated to have been about 30 percent higher than that of the French, though such comparisons are inherently hazardous. What is perhaps more surprising is that the tax burden in Britain was almost twice what it was in France: 12.4 percent of GNP as opposed to 6.8 percent. Moreover, the British national debt as a proportion of GNP exceeded that of the French by more than threefold; yet because French finances were much less sound than the British, the annual debt service ratio was comparable (all figures from Weir, 1989, p. 98). These figures do not explain the Industrial Revolution in Britain, but they should serve as a warning for simple-minded explanations that view high taxes and government debts as a prescription for economic disaster. Despite its high taxes and a government debt that climbed from 5 percent of GNP in 1688 to 200 percent of it in 1815, Britain had a viable and strong economy, strong enough to withstand a quarter century of fiscal stress following the French Revolution.

Different in emphasis but equally unequivocal in its certainty about the role of politics in Britain's Industrial Revolution is the view advanced by Mancur Olson (1982). Olson's theory of economic growth is based on the idea that political bodies are subject to pressure groups pursuing the economic interests of their members, even if they come at the expense of society as a whole. Olson is thus led to associate periods of economic success, such as the Industrial Revolution, with the comparative weakness of such pressure groups. Britain during the Industrial Revolution, maintains Olson, was relatively free of class differences and by comparison a socially mobile society so that loyalty to a particular pressure group was not yet very strong. The Civil Wars of the seventeenth century, moreover, had created a stable nationwide government, which made Britain into a larger jurisdictional unit in which it was more difficult to organize pernicious pressure groups (Olson, 1982, pp. 78-83, 128).

Despite a number of inaccuracies, Olson's insight that technological progress depended to a great extent on the political environment is valuable.[48] As I have pointed out elsewhere (Mokyr, 1992b), technological progress almost inevitably runs into resistance by vested interests who stand to lose some of their rents as a result of the revaluation of physical and human capital. It is natural and rational for these groups to organize and try to resist the changes. Because that resistance by definition has to use nonmarket mechanisms, the government plays a pivotal role here. First, the technologically conservative forces might try to use existing organizations, such as guilds or even the government itself, to pass and enforce regulations and legislation inimical to technological change. Second, they may try to use extralegal methods, such as violence, to try to suppress innovation. The attitude of the authorities is thus crucial in determining the outcomes of these struggles. On the whole, the British government during the Industrial Revolution consistently and vigorously supported innovation. Many of the obsolete laws and regulations that encumbered progress (for example by mandating precise technological practices in detail) were revoked. Labor organizations ("combinations," in the language of the day) that were seen

---

[48] Olson writes (1982, p. 128) that the English Civil Wars "discouraged long-run investment" (a possible but wholly undocumented inference) but that "within a few decades after [the Civil War] it became clear that stable and nationwide government had been re-established in Britain [and] the Industrial Revolution was under way." "Under way" is, of course, an ambiguous phrase, but between the Restoration and the beginning of the Industrial Revolution, as commonly defined, a century or more (and not "a few decades") had passed.

as threatening the advance of technology were made illegal and had little effect. In 1809 Parliament revoked a sixteenth-century law prohibiting the use of gig mills in the wool-finishing trade, and five years later it did away with one of the pillars of regulation, the Statutes of Artificers and Apprentices. Violent protests, such as the Luddite riots, were forcefully suppressed by soldiers. As Paul Mantoux put it well many years ago, "Whether [the] resistance was instinctive or considered, peaceful or violent, it obviously had no chance of success" (Mantoux, 1928, p. 408). Challenges to law and order that could not be settled by local authorities were dealt with effectively and harshly.

Was Britain a laissez-faire economy, and does the Industrial Revolution therefore stand as a monument to the economic potential of free enterprise? In absolute terms, Britain certainly was not a pure laissez-faire economy. A large number of regulations, restrictions, and duties were on the books. But absolutes are not very useful here. Compared with Prussia, Spain, or the Habsburg Empire, Britain's government generally left its businessmen in peace to pursue their affairs subject to certain restraints and rarely ventured itself into commercial and industrial enterprises. Seventeenth-century mercantilism had placed obstacles in the path of all enterprising individuals, but British obstacles were less formidable than those in France. More regions were exempt, and enforcement mechanisms were feeble or absent. One such enforcement mechanism, widely used on the Continent, was the craft guild, yet by the time of the Glorious Revolution of 1688, the craft guild in Britain had declined and lost most of their political clout (Nef, 1957, pp. 26, 32). Market forces were more powerful than politics, even if they were constrained to operate within a framework of laws and institutions produced by political forces. Mercantilism and regulation in eighteenth-century Britain was alive and well, yet it never took the extreme forms it took in France under Colbert and in Prussia under Frederick the Great.

The general consensus among historians today is that the regulations and rules, most of them relics from Tudor and Stuart times, were rarely enforced. As the economy became more sophisticated and markets more complex, the ability of the government to regulate and control such matters as the quality of bread or the length of apprentice contracts without an expanding bureaucracy effectively vanished (Ashton, 1948, p. 95). The central government was left to control foreign trade, but most other internal administration was left to local authorities. Internal trade, the regulation of markets in labor and land, justice, police, county road maintenance, and poor relief were all administered by local magistrates. Although in principle these authorities could exercise considerable power, they usually elected not to. This de facto laissez-faire policy derived not so much from any libertarian principles as from the pure self-interest of people who already had wealth and were making more. By ignoring and evading rather than altogether abolishing regulations, Britain moved slowly, almost imperceptibly toward a free-market society. Except for its strictures against the state's intervention in foreign trade, *The Wealth of Nations* was a century out of date when it was published: What it advocated had already largely been accomplished (Perkin, 1969, p. 65).[49] Some regulations were more difficult to ignore than others. The usury laws, which set a ceiling on all private interest rates, are thought by some historians to have had considerable impact on the allocation of resources (Ashton, [1955] 1972, pp. 27-28; Williamson, 1984). There is, however, evidence indicating that the usury laws were sufficiently evaded to limit their impact on the economy.[50]

Even when mercantilist regulations were enforced, their net effects were ambiguous. The silk and light woolen industries tried to stop the import of cheap Indian cottons. This resulted in the Calico Act which

---

[49] The Statute of Artificers (of 1563), for instance, so detested by Adam Smith, required that workers serve a formal apprenticeship before their employment in a trade. Yet in 1777 the calico printers admitted that fewer than 10 percent of their workers had served because "the trade does not require that the men they employ should be brought up to it; common labourers are sufficient" (Mantoux, 1928, p. 453).

[50] Although the usury laws were not capable of holding down private interest rates to 5 percent at all times, they distorted the capital market to a substantial degree. A Parliamentary Select Committee concluded in 1818 that "the laws regulating or restraining the rate of Interest have been extensively evaded and have failed of the effect of imposing a maximum on such rate. . . . Of late years, from the constant excess of the market rate of interest above the rate limited by law, they have added to the expense incurred by borrowers on real security" (Great Britain, 1818, vol. VI, p. 141). See also Pressnell (1956, pp. 95, 318, 368, 428) and Cottrell (1980, pp. 7-8, 13).

prohibited the importation and sale of printed white calicoes, passed in 1721 and repealed in 1774, and a host of other measures and countermeasures. The maze of protection and subsidization was the confusing outcome of political pressures and counterpressures by interest groups that tried to keep out competition and keep in complements. Because fustians looked much like calicoes, the prohibition was widely evaded, although it remained a nuisance.[51] It has been argued that the mercantilist laws that prohibited the importation of calicoes stimulated the British cotton-printing industry and that high taxes and tariffs on white calicoes encouraged domestic production (Wadsworth and Mann, 1931, p. 144). More recently it has been argued that by encouraging fustians these regulations constituted a "legislative assistance that was important for the mechanization of Lancashire's growing industry," so that "British pragmatism appears in retrospect more productive than Dutch free trade or French style mercantilism" (O'Brien, Griffiths, and Hunt, 1991, pp. 415, 418). Yet evidence for any direct link between the protectionist measures taken and the technological breakthroughs in cotton is absent. What we know with certainty is that mercantilist bounties and encumbrances to trade distorted the operation of the free market, and as soon as Arkwright's patent was secured and his machines producing, he petitioned for repeal of the Calico Act and was granted it in 1774. Most of the important inventions in cotton, including the mule, cylindrical printing, the power loom, and the carding machine, followed in the decade after the repeal of these acts. Until more evidence is forthcoming, it seems reasonable to conclude that technological progress occurred *in spite of* rather than *thanks to* the meddling of a special-interest-driven Parliament in the price mechanism.

The Bubble Act, passed in 1720, required a private act of Parliament to establish a common-stock corporation. However, modern scholars have increasingly realized that this impediment, too, was more an inconvenience than a real obstacle to business activity (Cottrell, 1980, p. 10).[52] Even after the Bubble Act was repealed in 1825 and all remaining obstacles to joint-stock company formation were removed in the Joint Stock Companies Act of 1856, there was no sudden rush to create joint-stock corporations. The prohibition on incorporation was a less formidable obstacle to technological progress and industrial growth than might appear. The same applies to the restrictions on the export of textile machinery and the emigration of artisans (Jeremy, 1977; Jeremy, 1981, chap. 3). Business organization law remained a "mound of case law not tidied up until the end of the nineteenth century." Partnerships, the normal mode for business associations, increased the vulnerability of business, could only be bought and sold with unanimous consent, and could not sue as an Association without a private act of Parliament (O'Brien, 1994, p. 234). Yet the history of the Industrial Revolution is full of remarkably symbiotic relations between partners, and while some individual enterprises may have suffered from cumbersome institutional relics, the system as a whole seems to have found workable (if not always cheap) ways around it.

Not all government intervention was equally ineffective, of course. A few government monopolies, such as the East India Company, survived well into the nineteenth century. Moreover, free trade remained a far cry from reality until well into the nineteenth century. During the Napoleonic Wars, tariffs were raised to unprecedented heights (peaking at 64 percent of the value of imports in 1822). A slow trend toward lower tariffs began in 1825, culminating in the abolition of the Corn Laws in 1846 and the repeal of the Navigation Acts, which had severely limited foreign freighters from carrying British goods, in 1849-1854. Yet in the first half of

---

[51] By 1736 fustians were explicitly exempted from the Calico Act, and by this time they contained two-thirds cotton and one-third linen, so that fustians "replaced Indian calicoes as the prime threat to light woollens and silks" (O'Brien, Griffiths, and Hunt, 1991, pp. 414-415).

[52] The Bubble Acts could be evaded by organizing companies under a trust deed, a legal form widely used in the woolen cloth industry in Yorkshire (Hudson, 1983).

the nineteenth century, Britain's trade was more restricted by tariff legislation than France's (Nye, 1991a). To be sure, tariffs and navigational restrictions were widely evaded, too.[53]

Another area in which government intervention was important and the law far from a dead letter was poor relief. Here the difference between Britain and the Continent is striking. Nowhere in the world can one find a well-organized, mandatory poor relief system like the English one. The Old Poor Law, sometimes erroneously referred to as "Speenhamland" (in fact, the Speenhamland system of allowances in aid of wages was used in a minority of counties), has had a notably bad press. Three major criticisms have been raised against it. One was the Malthusian complaint that outdoor relief subsidized childbearing and thus increased the birth rate. A second criticism, already mentioned by Adam Smith ([1776] 1976, p. 157), was that the Old Poor Law (and particularly the Settlement Acts) encumbered the free movement of labor and thus hindered its reallocation in a society in which labor markets played an ever-increasing role (Polanyi, [1944] 1985, pp. 77-102; Ashton, 1948, p. 111). Finally, the standard complaint against the Old Poor Law was that it impaired the incentive to work by distorting the leisure-income trade-off, or, in the language of the time, encouraged indolence and sloth.

These criticisms have not fared well in recent years. Indeed, it seems likely that the effects of the Poor Laws on the Industrial Revolution were not nearly as negative as used to be thought. The demographic argument against them has been criticized by James Huzel (1969, 1980). More recently, however, the important work of Boyer (1990) has vindicated Malthus's approach. The use of multivariable regression shows that the introduction of child allowances after 1795 did have an important effect on birth rates.[54] Whether the Old Poor Law was somehow responsible for the creation of an army of able-bodied paupers is still unclear and awaits further research. In the absence of any a priori idea of the effect of the increase in birth rate on the Industrial Revolution, however, it is unclear what the long-term economic implications of this higher birth rate were. Moreover, even in the absence of a poor law, population would have grown, and its demographic effects were the most pronounced in the south of England.

As to the geographical immobility imposed by the Settlement Acts, these were to some extent alleviated by the Poor Law Removal Act of 1795 (35 Geo. III (1795) c. 101), which expressly forbade the ejectment of poor immigrants unless they actually became chargeable to the parish. Even before 1795 the system was "by no means such a check on mobility of labour as some of the older writers . . . supposed," because as the option to evict was exercised in a haphazard and casual way (Styles, 1963, p. 62). Some contemporary opinion agrees with this finding. Sir F. M. Eden, whose opinion according to Redford was "as weighty as that of Adam Smith," thought that the Settlement Laws were too weakly enforced to constitute the hindrance to mobility alleged by Smith (Redford, 1964, p. 85). Perhaps the primary mechanism by which the Settlement Acts discouraged migration was their sheer complexity and the uncertainty that irregular enforcement implied for anyone contemplating migration. Since migration was, however, a risky undertaking under any circumstances, it is far from obvious

---

[53] Smuggling was widespread, as can be verified from the fact that at times, when tariffs were reduced substantially, imports increased by a much larger proportion than the reduction of the tariff and a reasonable guess about the elasticity of demand would imply. For example, when the tariff on coffee was reduced by two-thirds in 1808, imports into Great Britain increased from 1.07 million to 9.3 million lbs. in 1809.

[54] The observed birth rate rose by 14 percent, according to estimates of Wrigley and Schofield (1981), between about 1780 and about 1820. Boyer estimates (1990, p. 170) that in the absence of child allowances, the birth rate would actually have declined by 6.4-9.2 percent. He concludes that allowances in aid of wages did to some extent "create the poor which they maintain" (p. 142). The numbers he provides imply that in the absence of the poor laws, English population would still have been larger in 1826 than it was in 1781, but it would have grown at a much slower rate after 1795. A rough computation suggests that on Boyer's assumptions the population of England and Wales in 1826 without a poor law would have been 9.78 million instead of the 12.4 million estimated by Wrigley and Schofield. Solar (1995) suggests that because the benefits were financed by poor-rates paid by local landlords, they had an incentive to try to reduce the number of potential recipients by discouraging large families, though it is unclear how successful such policies might have been. From a different perspective, McCloskey (1973) has also argued that the wage supplements paid under the Old Poor Law were likely to have reduced the supply of labor and thus may have raised wages, though the magnitude of this disincentive-to-work effect is unclear and the evidence for it rather weak.

to what extent the Old Poor Law made things worse.[55] More to the point, Boyer's analysis shows that the overall magnitude of the Poor Law's effect on labor mobility bemoaned by Polanyi was neglibly small.[56]

As to the work-incentive effect stressed by T.R. Malthus and his followers, research carried out by Blaug in the 1960s has recently been reinforced by the work of Pollard (1978, pp. 109-110) and George Boyer (1990). They argue that the causality runs the other way: Wage-support payments were made in areas that suffered from seasonal unemployment and the decline of cottage industry, which explains the association of Speenhamland with the agricultural areas of England. Boyer's regressions provide little support for the hypothesis that outdoor relief caused an increase in voluntary unemployment, although it was not possible to estimate the relation between the two directly (Boyer, 1990, p. 142-143). The effect of poor-law variables on male labor income was statistically insignificant, which it could not have been if poor relief had been treated as a substitute for labor income.

Indeed, it could be maintained that the Poor Laws, despite their obvious flaws (in particular their nonuniformity), may have had some overall positive effects on the Industrial Revolution. A comparison with Ireland, which had no formal system of poor relief prior to 1838, bears this out (Mokyr, 1983). The social safety net provided by the Poor Laws allowed English individuals to take risks that would have been imprudent in Ireland, where starvation was still very much a possibility. In societies without such laws, self-insurance in the form of large families and liquid assets was widely held, whereas in England even the worst case rarely implied actual starvation. In a recent paper, Solar (1995) extends this argument to the creation of a wage-labor force. The main obstacle to the creation of a wage-labor force was the attachment of the rural population to land. Land served not only as a source of income but also as a form of insurance -- in times of duress it could be mortgaged or sold. It was also a form of old-age insurance; its inheritability made it a bargaining chip with which parents could persuade their children (or other heirs) to look after them in their old age (see also Guinnane, 1991). The existence of the British Poor Law provided a substitute for land for insurance purposes and thus reduced the need to cling to land at all costs, thus contributing to the creation of a proletariat needed for the factories and the railroad. The magnitude of this effect is of course not known, but it makes sense as economic analysis.

The Speenhamland system, by subsidizing workers in the off-season, assured a regular labor force during the busy seasons in agriculture (Boyer, 1990). A similar argument may be made for manufacturing: Workers could be laid off during periods of business slumps without fear of having the labor force emigrate or starve. Irish employers, on the other hand, complained about having to continue to pay their workers during slumps or risk losing them (Mokyr, 1983, p. 227). In addition, the practice of pauper apprenticeships and the recruitment of factory workers from workhouses run by local Poor Law guardians provided an important source of labor for the factories, especially in rural and small-town mills before 1800.[57] All this is not to argue, of course, that the Poor Laws somehow "caused" the Industrial Revolution. But it seems that a case can be made that their net effect was not nearly as negative as has been maintained and that they may have had hitherto unsuspected beneficial effects.

Another political difference between Britain and most other European countries was the lack of centralization of political power. Britain's system of government left most of the day-to-day management of

---

[55] In 1832 out-migration was more important in Speenhamland parishes, which paid allowances in aid of wages or child allowances, in Kent than in non-Speenhamland parishes (Huzel, 1980, pp. 375-378).

[56] The fact that the British Poor Law was a *national* system rather than a patchwork of local systems, as on the Continent, may have *increased* geographic mobility by reducing the uncertainty involved in migration (Solar, 1995).

[57] Some of the transactions between Poor Law authorities and mill owners resembled nothing as much as slave trade; e.g., the purchase of seventy children from the parish of Clerkenwell by Samuel Oldknow in 1796 (Mantoux, 1928, p. 411). Pollard ([1965] 1968, pp. 194-195) cites the sanctimonious claim by some notorious users of child labor that these pauper apprentices were "more expensive" than paid labor and that they were employed out of civic duty. For a similar view, see Collier (1964, p. 45). Recruiting agents were often sent to scour the surrounding countryside in search of workhouse labor, and some of these children were brought in from the other end of the country, which indicates that for some industrialists pauper apprentices were indeed a cheap and satisfactory form of labor.

affairs to local magistrates, who were, on the whole, respectable residents for whom administration was a form of leisure activity. Whether this government by amateurs was an effective way of providing government services is another matter, but one effect was the relative unimportance of London as an administrative and cultural center when compared to Madrid, Paris, St. Petersburg, or Vienna. In France, for example, Paris traditionally drained large amounts of talent from the provinces, and provincial centers of learning and technology were of small importance compared to those in the capital. This rural-urban brain drain would not have mattered, of course, if industrialization could have been concentrated near the capital of the country. Interestingly, this did not happen anywhere. Neither Brussels nor Paris, nor Berlin, nor Amsterdam, nor any other major capital city in Europe became a center of modern industry. Although some manufacturing activity developed around the capitals, the main centers of modern industry were usually elsewhere. As a result, a highly centralized state in which the capital city drained the countryside of ambitious and able men, strongly attracted to "where the action is," operated at a disadvantage compared to a decentralized state like Britain.[58] In Britain the situation was radically different; provincial institutions like the Manchester Literary and Philosophical Society or the universities of Glasgow and Edinburgh, located near centers of industry, were of central importance to the technological developments of the eighteenth century. Wrigley (1967) has argued more or less the opposite, ascribing to London a major role in creating the conditions leading to the Industrial Revolution. The size of London relative to England's population and its enormous needs in terms of food, fuel, and other products seem to support his claim. Sheer size, however, is not necessarily an advantage. A top-heavy capital might just as well be viewed as imposing a major cost on the country. Wrigley's argument seems better suited to explain commercial development before 1750 than industrial development thereafter. During the Industrial Revolution, indeed, the demographic predominance of London declined somewhat. Between 1650 and 1750 London's share of English population rose from about 7 percent to 11.8 percent. By 1800 this percentage had declined to 10.5 percent.[59] All the same it would be wrong to ignore the importance of London; after all, it was a major industrial town in which much of Britain's beer was brewed, its silk thrown, its books printed, and many of the sophisticated machine tools made by Bramah and Maudslay were first conceived.

Some historians have argued that the British government stimulated the Industrial Revolution by creating a demand for military products, which led to rapid technological change in some industries (McNeill, 1982, pp. 210-212). It is true that some of these externalities can be identified. Cort's puddling-and-rolling technique was completed when its inventor was working on a contract for the Admiralty. Wilkinson's lathe, which bored the accurate cylinders needed for Watt's steam engines, was originally destined for cannon. The correct test for the net impact of military demand is, however, the question whether in the absence of military demand these innovations would have been substantially slower in coming. On that issue most scholars are wisely cautious. Moreover, what little innovation that can be directly attributable to the war had few civilian spillover effects. A case in point is the well-known Portsmouth manufacture of wooden blocks for pulleys to be used on naval vessels, designed by two of the greatest engineers of the time, Marc Brunel and Henry Maudslay. Despite the precocity of this plant, which pioneered interchangeable parts as well as continuous flow processes, it was too specializd to have spillover effects on the civilian sector. Scholars largely agree that favorable external effects were relatively small and that on balance the economic impact of the wars between 1756 and 1815 were negative (Trebilcock, 1969, pp. 477-478; Hyde, 1977, pp. 112-116). Moreover, any hypothesis of a substantial positive effect of the government's war-related activities on technological progress encounters a difficulty: If military efforts created major technological externalities, why did France and other Continental countries not benefit from them to the same degree that Britain did? Research on the French iron industry, for example, shows that the revolutionary and Napoleonic wars did little to stimulate technological progress (Woronoff, 1984).

---

[58] See Cardwell (1972, p. 126) for a similar argument. Interestingly, Ireland, with its centralized government in Dublin, conforms more to the Continental than the British model.

[59] The London population estimates are from Wrigley (1967, p. 44). English population data (less Monmouth) are from Wrigley and Schofield (1981).

To summarize, most economic historians would agree that politics was a positive factor working in Britain's favor, although the magnitude of the effect, as well as its *modus operandi*, are still in dispute. The appropriate standard of judgment should be a comparative one, and it seems hard to disagree with the proposition that the specific form of government that had emerged in Britain created an environment that was more conducive to economic development than elsewhere. Some oppressive mercantilist laws were on the books, but most were successfully evaded. Britons were heavily taxed, but taxation was never allowed to become arbitrary and confiscatory. Most important, the right to own and manage property was truly sacrosanct, contrasting sharply with the confiscations and conscriptions on the Continent during the French Revolution and the Napoleonic era. Personal freedom -- with some exceptions -- was widely accepted in Britain. True, the Acts of Settlement remained on the books until 1834, but they were by no means as restrictive as the harsh requirements on the books in France and in Prussia, where workers were required to have *cahiers* or *Wanderbücher* in which their employment was recorded and which required them to ask for passes for journeys within the country. Serfdom was still very much in existence east of the Elbe in 1815. The cathartic revolutionary medicine administered to the Continent between 1789 and 1815 by the French was needed to prepare the rest of Europe for the modern age. But the medicine's immediate side effects were so painful that most of the Continent required many years and even decades to recover from the treatment and start to threaten Britain's lead. Britain did not need this harsh shock treatment, since it alone had learned to adapt its institutions to changing needs by more peaceful means, and the English Channel had sheltered it from undesirable political imports.

Britain's political stability contrasts sharply with the history of France, with its four major revolutions in the eight decades following 1789. But was political stability always an asset on the path toward modernization? If investors are wary of investment in politically unstable environments, political stability was an advantage and its absence had a negative effect on industrialization. But how important was that effect? The economic performance of powerful autocratic and "stable" regimes in Russia and Turkey was disappointing to say the least. Moreover, Olson (1982) has insisted that political stability is in fact a rather mixed blessing, because it permits the crystallization of pressure groups whose activities are, in Olson's view, the archenemy of economic development. It is thus unclear how much of the difference in economic development can be attributed to this factor.[60] Still, it is no exaggeration to say that nowhere in the world was property perceived to be more secure than in Britain. Such security is important in part because it included intellectual property rights, such as patents and pensions awarded in recognition of breakthroughs. Moreover, much technological progress required capital goods in which they were "embodied," from the machinery itself to buildings and sites. Clearly, security of these assets from taxation, confiscation or private tresspass was necessary if such investments were to be sustained.

Finally, British society exhibited a degree of tolerance for deviant and heterodox ideas that was unusual, though not unique. Although tolerance was quite different from equal rights, Britain developed in the seventeenth century the ability to accomodate a high level of acceptance of different modes of thinking. The intolerance on the Continent toward dissidents led to the hemorrhage of technical talents from the southern Netherlands and France to countries where they were more welcome. As Landes (1983, p. 219) recounts it, after 1685 (when the Edict of Nantes was revoked) French industry was "crippled by the exodus of some of its best practitioners fleeing a wave of anti-Protestant bigotry and persecution." In many industries, France's loss was England's gain. The Belfast linen industry was, if not founded, certainly enhanced and developed by the technical know-how of Huguenot refugees, especially Louis Crommelin. Nicholas Dupin was an active promoter of companies and operated a number of paper mills in England. The great hydraulic engineer and lecturer John (Jean) Desaguliers, too, came from a Huguenot family as did Denis Papin, who had as much ground for claiming to be "the" inventor of the atmospheric engine as anyone. Crouzet, who has studied the financial activities of these refugees, states that the "persecution of the Huguenots [was] not only a crime, [it] was also a blunder, as

---

[60] The revolutions in France may have increased the perceived insecurity of property and inhibited capital formation. Similarly, the continuous struggle between landlord and peasant in Ireland before the famine reduced the attractiveness of Ireland as a site for industrial capital (as is the case today in Ulster). The Civil War in Spain (1832-1839) and the Miguelite Wars in Portugal (1828-1834) had similar effects in the Iberian Peninsula.

France was impoverished by a brain drain which brought wealth to her rivals and enemies" (Crouzet, 1991, p. 224). The direct impact of these individuals on the aggregate economy may not have been vast, but that is less important than their significance as a symptom of the open-minded attitude of agreeing to disagree that flavors the British enlightenment. Such open-mindedness is essential if new technological ideas are to compete in the marketplace on their economic merits. The differences between Britain and the Continent were not absolute here either. At times Britain turned on its most innovative spirits, as it did to the inventor of the fly shuttle, John Kay, who ended up having to flee to France, and as it did to the great chemist Joseph Priestley, whose unpopular political views caused a mob to burn down his house and forced him to flee to the United States. On the whole, however, the atmosphere in Britain was comparably comfortable for rebels and deviants, of which inventors in some sense are a subspecies.

e. Demand vs. Supply

A large and venerable literature links, in one form or another, the British Industrial Revolution to the growth of the home market, the expansion of consumer demand, and the growth of a "consumer revolution." From the point of view of economic analysis, technological change, capital accumulation, and the rise of the factory are primarily supply-side phenomena. Demand-side factors are more difficult to integrate into the story. Yet economic historians, beginning with a famous paper by Gilboy (1932), have always felt intuitively that demand should be given a parallel role. In price theory it is typically assumed that demand and supply move independently of each other, so that an increase in demand means a movement *along* the supply curve. Any argument that links the Industrial Revolution with changes in demand relies on models that postulate a shift of the supply curve as a response to an increase in demand. North, relying on the work of Kenneth Sokoloff, has recently concluded that innovation and technological change are primarily determined by the "size of the market" (1990, p. 75; cf. Sokoloff, 1988). Less cautious writers have gone further and simply asserted that a "consumer revolution" was a necessary condition for the Industrial Revolution to occur. Thus in an influential paper stating the most extreme position on this question, Neil McKendrick (1982a) writes that "the Consumer Revolution was the necessary analogue to the Industrial Revolution, the necessary convulsion on the demand side of the equation to match the convulsion on the supply side."[61]

As I argued in a paper first published in 1977 (Mokyr, 1985b), supply and demand are not symmetrical in long-term economic change. In a historical event like the Industrial Revolution, demand factors can only play a role under certain assumptions that have to be examined carefully. To start with, it is important to distinguish between economic changes that affect economic growth in a fixed technology (for example, the expansion of trade due to growing markets) and those that actually change the techniques in use. While the two may be related at some level, they can be treated logically as distinct and the causal link between them has to be demonstrated. Secondly, if output increased and technology possibly changed because of a rise in demand for industrial goods, it has to be made clear *why* demand increased in the first place. Changes in demand are not exogenous to an economic system -- they occur for well-understood reasons. Population, of course, began to increase rapidly after 1750, but this was a worldwide phenomenon and it seems far-fetched to link it directly to the Industrial Revolution. In a technologically static world, population growth (as the Classical School firmly believed) would lead to declining living standards. Hence, population growth in and of itself would increase the demand for food products more than the demand for manufactured goods, and the combination of growing population, bad harvests, and disruption of foreign supplies led to sharply higher agricultural prices, hardly a stimulus for industrial demand.[62] Export demand, too, although of some importance in some industries, does not seem to have been the crucial element in the Industrial Revolution that some scholars have claimed. The role of foreign

[61] For a critique of McKendrick's view, see for example Fine and Leopold (1990) and John Styles (1992).

[62] The demand for agricultural goods was inelastic, so that increases in agricultural prices meant that a larger amount of income was spent on agricultural goods, reducing the amount left for manufactured goods.

trade in the Industrial Revolution, however, is sufficiently interesting and controversial to merit a separate discussion.

Secondly, the *modus operandi* of demand-side factors has to be specified and documented. For instance, an increase in aggregate demand due to, say, a rise in the propensity to consume or an autonomous growth in investment will only have an effect if the economy has large underutilized resources that can be brought into production. Such a Keynesian scenario may indeed have been of some importance. Evidence for large amounts of underutilized resources that were brought into production as aggregate demand expanded in the second half of the eighteenth century is, however, lacking. Or, if there were strong positive external economies between firms so that a sharp increase in demand led to an industry-wide decline in costs, demand would be directly linked to higher productivity. Yet the existence of such externalities is notoriously hard to demonstrate.

Thirdly, McKendrick's observations that the Consumer Revolution was somehow correlated with the Industrial Revolution seems open to a level of historical criticism from which it will not easily recover. Work by Lorna Weatherill (1988) suggests that if there was a Consumer Revolution at all, it peaked in the period 1680-1720. The long lag between that event and the Industrial Revolution makes any causal connection between the demand and changes in industrial technology difficult to support.[63] Equally damaging is the fact that consumer revolutions were taking place elsewhere in Europe. Seventeenth century Holland was, of course, the most obvious example thereof, but Cissie Fairchilds (1992) has employed probate records to show that France, like England, experienced a consumer revolution albeit fifty years later. The goods that the French bought were different, but on the whole the absence of an Industrial Revolution following the French increase in mass consumption leads Fairchilds to conclude that the two revolutions were largely independent of each other and that the changes in technology were shaped by supply, not demand side elements. In a recent paper Horrell (1996) has employed household budget studies to test whether an increase in home demand between 1801 and 1841 indeed did take place at all. She finds indeed an increase in aggregate demand, but that many of the changes associated with the Industrial Revolution such as increased urbanization and a declining subsistence sector led to a retrenchment of working-class demand into the products of traditional industries and reducing demand for the new industries. The increase in middle class demand was far more substantial and clearly created large markets for the new products. Yet, as Horrell concedes, this is not at all tantamount to a demonstration that such an increase in spending on nonessential items fed back into the processes that produced the increase in income. In a growing economy somebody has to earn and spend the increased incomes. The "demand" hypothesis suggests that such spending helps increasing incomes per capita even more. It is this part of the story that remains unpersuasive.

The notion that somehow technological change takes place when the demand for it "arises" is thus clearly fallacious. Some scholars refuse to abandon the concept.[64] As T. S. Ashton argued long ago, invention was the mother of necessity, not the other way around (1948, p. 62).[65] All the same, it seems natural to pose the question whether technological change will occur without some prior knowledge that the goods produced will sell. Will it be possible to "find people with income and demand schedules capable of absorbing the increased output?" (Eversley, 1967, p. 211). It should be noted that unless the good produced is totally new and has no

---

[63]Among the goods the consumption of which increased according to the probate records were knitted goods, pottery, pipes, clocks, mirrors, and fancy textiles.

[64] Braudel (1984, p. 566) writes flatly that "the efficient application of technology lags, by definition, behind the general movement of the economy; it has to be called on, sometimes several times, to meet a precise and persistent demand." Jan De Vries (1994, p. 255) notes that "the interest in a demand-side appreciation of early industrialization, beaten back in economic history, emerged again among social historians, among whom sightings of a 'consumer revolution' gained credence and has now found a comfortable home among cultural historians, where the triumph of the will of the consumer can overcome any scarcity."

[65] Economists and historians alike have treated the common wisdom that necessity is the mother of invention with contempt. For some examples of this literature, see Mokyr (1990a, p. 151, n. 1).

suitable substitutes (for example, aspirin), invention usually occurs in markets that already exist. When he improves an existing good or produces it at a lower price, the innovator taps into a market he already knows. By innovating, he undercuts his competitors or those selling close substitutes. The invention of the puddling-and-rolling technique or the continuous paper-making machines, for instance, can be represented as supply curves shifting to the right, with the market sliding down existing demand curves. An *autonomous* and *prior* shift of the industry demand curve is not an essential part of the story. Modern studies of contemporary technological progress have often claimed considerable evidence for demand-led technological change, but these studies are often flawed and biased. In their demolition of many of these studies Mowery and Rosenberg (1979, p. 142) note that "the demand-pull approach reflects an insufficient appreciation for the iunnumerable ways in which ... very small changes in production technology are continuously altering the ... structure of production cost."

Still, this does not mean that demand played no role in generating technological change. Adam Smith himself noted that the division of labor was limited by the extent of the market and strongly believed that the division of labor itself was the main agent of technological progress. He thought that highly specialized workmen were more likely to come up with inventions.[66] Innovation usually involved substantial fixed costs, and thus a minimum level of sales was expected by the innovator. In 1769 Matthew Boulton wrote to his partner James Watt, "It is not worth my while to manufacture your engine for three counties only, but I find it very well worth my while to make it for all the world" (cited by Scherer, 1984, p. 13). Some minimum level of demand was thus necessary to cover the fixed costs of research and development. An expansion of demand, through the integration of markets or through a growth in population and income or through an increase in export demand, could thus have stimulated invention.

In fact, however, fixed costs, including those of R and D, remained relatively small in most industries, as the large number of firms indicates. The costs of invention were small relative to the costs of production. Although men like Crompton and Trevithick worked for many years on their inventions, these costs would still have been covered in a much smaller market. This was true whether industry demand was stationary, expanding, or even contracting. It is, of course, true that in a highly fragmented economy, with high transport costs or internal barriers to trade, the competitive model does not hold. Szostak (1991) maintains that the increase in demand engendered by improved transport led to regional specialization and an accelerated rate of technological progress. Yet, as he realizes, a more integrated economy is not quite the same as an expansion of market demand, even if for the individual producer they may be indistinguishable. In Szostak's model, the *primum mobile* is an improvement in transportation, itself a supply-side phenomenon.

Where changes in demand can and do matter is when demand shifts from an industry that is relatively impervious to technological change to one that is not. It is, for example, quite clear that of the three large textile industries -- wool, cotton, and linen -- cotton fibers lent themselves best to mechanization (although worsted yarns were also well adapted to Arkwright's rollers). A change in demand in favor of cotton would increase output, and insofar as technological change was a function of the quantities produced (as in learning-by-doing phenomena), demand shifts could have affected the rate of technological progress. Demand for cotton, moreover, was price elastic, which means that for any given shift in supply a large increase in sales could be

---

[66] Smith supports this view by the story of a boy who, while operating one of the first steam engines, tied a string to the handle of a valve, allowing it to open and shut automatically. As Cannan points out in his notes, the story is apocryphal (Smith [1776] 1976, pp. 13-14). On the whole, Smith's ideas of the connection between the division of labor and technological change seem to be lacking in persuasion. He postulates that "the greater the number [of laborers in a workhouse], the more they naturally divide themselves into different classes and subdivisions of employment. More heads are occupied in inventing the most proper machinery for executing the work of each, and it is, therefore, more likely to be invented" (Smith, [1776] 1976, pp. 96-97). In some cases there may be merit to this argument. Some machines were made to mimic the motions of human arms, and the simpler the task, the more such imitation was possible. A division of labor between workers and engineers could create a special class of outsiders who could observe the production process and suggest improvements. Yet how much division of labor was necessary to create the conditions necessary for invention? It could just as well be argued that rigid specialization stifles the cross-fertilization between different activities that is the source of much technological creativity. Adam Smith's own career, incidentally, seems a good counterexample of his belief in the benefits of specialization (Brenner, 1987, pp. 109-110).

realized leading to further learning. Yet economic analysis sounds a warning bell: The elasticity of demand is important, but a single inventor in an existing competitive industry always faces a very elastic demand curve, much more so than the industry as a whole. All the same, the strong demand for cotton clothes, due in part to fashion, operated as a "focusing device," in Rosenberg's (1976) terminology, with inventors directing their energies to an industry that was expanding.

The same is true for the "leapfrogging" models proposed by Landes (1969) in which a sudden increase in the productivity of one activity (such as weaving) created a demand for improvement in the other, complementary activity (spinning). Sudden demand-induced imbalances may focus the attention of inventors on a profitable avenue, but they do not constitute a complete theory of technological change. Why are some "bottlenecks" solved by technological change while others have to be accommodated by massive reallocations of resources?[67] Markets for knowledge existed, to some extent, and a sudden surge in demand for technical knowledge might well have led to more technical innovations. Yet as Ian Inkster, in a recent criticism of this hypothesis points out, if this were the case we should have observed a *higher* price for this knowledge, which eventually would have choked off the rate of growth (Inkster, 1991, p. 69). Yet, if anything, the reverse was the case.

Economists interested in economic growth in the past few years have come to realize that the standard assumptions of constant returns and limited externalities do not necessarily hold in historical reality. As I have already noted, relaxing these assumptions leads to radically different insights into the dynamics of an economy. Minor shifts in demand could trigger the economy to move one way or another and thus could have been "causal" in the Industrial Revolution (O'Brien and Engerman, 1991). Alfred Marshall, more than any other neoclassical economist, realized the dangers that such production technologies implied for the static market equilibrium that is at the heart of standard economics. Yet it is important to note that although such models are likely to increase our understanding of historical change, they depend on certain conditions to hold, none of which have ever been satisfactorily demonstrated to have been of great import in Britain during this period: economies of scale, strong externalities, learning effects, and similar sources of positive feedback.[68] A major source of such effects is found in modern network technologies, in which standardization is required. Such network technologies are found throughout history but it is not till the Industrial Revolution that they became of central importance. Gaslighting, railroads, telegraph and later electricity, telephones, software and so on all came after the Industrial Revolution or in its later stages. They were not very important in the markets for consumer goods in about 1750.

A different approach to the "demand hypothesis" has been proposed recently in a duo of papers by Jan de Vries (1993, 1994) in which he argues that changes in preferences could be of importance in explaining some of the economic changes in eighteenth-century Britain. De Vries argues, essentially, that the period was characterized by two distinct but related events: a *supply*-driven *Industrial* Revolution and a *demand*-driven set of changes in household behavior that he calls an "industrious revolution." The idea focuses on the household as a decision-making process: The household can allocate its resources to production for the market or to household production. In premodern Europe, as is still true today, the existence of household work makes the concept of leisure hard to define. De Vries points out that market purchases and household production are

---

[67] Two examples will suffice: In the cotton industry, carding, spinning, weaving, and bleaching were all complementary, and improvements in one of these areas stimulated the others. Yet some activities defied mechanization: The planting and picking of cotton could not be mechanized, which had momentous consequences for the history of the southern United States. In coal mining, too, an increase in demand led to relatively few innovations in mining technology. Here markets replaced the innovation process: The proportion of workers in the coal and lignite mines in Europe subsequently increased everywhere, despite the rather obvious shortcomings of this employment.

[68] One tireless advocate of the role of demand in the Industrial Revolution (McKendrick, 1982a) speaks repeatedly of "mass markets," which suggests mass production, an important source of increasing returns. Yet as Styles (1992) has recently warned, applying modern terms of this nature to British manufacturing before 1850 -- manufacturing without interchangeable parts, without continuous flow processes, highly dedicated tools, or uniform standards -- is misleading.

imperfect substitutes for each other: Child rearing, food preparation, apparel making, and personal services can be purchased or homemade, but the products are not identical. An increased preference for the consumption of purchased goods requires cash, however, and thus implies greater labor force participation and market orientation. The resources thus reallocated were not idle before, nor were they absorbed by leisure, strictly speaking; they were simply deployed differently. The allocation between household and market depends simultaneously on preferences and on the relative efficiency of the household in producing for its own consumption or for the market.

The industrious revolution, in de Vries's view, was thus a change in allocation from production by, in, and for the household to a more market-oriented behavior. The net result was a vast increase in specialization on a microlevel: Workers came to produce, by and large, one or two products and buy everything else. It is tempting to attribute the shift to changes in preferences, although that would still not entirely justify Berg's enthusiastic claim (1994, pp. 134-35) that the shift in household behavior was necessary to the Industrial Revolution and her conclusion that "we must seek for the origins of the Industrial Revolution not just in women's labour but in women's wants and desires." Changes in preferences are notoriously hard to document. Moreover, while an exogenous change in preferences cannot be ruled out, such a redeployment could also have come as a result of technological changes.[69] First, better technology created and brought close to home some of the market-produced goods that the British consumer wanted to buy: cotton clothes, toys, adornments, tableware, kitchen utensils, clocks, books, and so on. As the array of goods that the consumer could buy increased, their quality improved, the uncertainty of their characteristics declined with standardization, and their price fell, the consumer would be more inclined to substitute cash income for housework. Second, the technological changes during the Industrial Revolution were biased in favor of production for the market. The factory was of course the obvious locus of the specialization of labor, but even those workers who remained at home found increasingly that they could do better by buying the goods they needed while producing for the market.[70] All the same, before 1800 or so the trend towards greater market participation accelerated and increased the effective labor input per worker by increasing the length of the labor year and intensifying the pace of work. The labor year could be extended in part by the reduction of seasonal unemployment through technical changes in the transport sector that made it possible to move materials and workers around with greater ease or through improvements in lighting technology that made it cheaper to work at night. Consumption of leisure declined as old and venerable institutions, such as "St. Monday" (the custom of taking Mondays off to recover from the weekend), were abandoned (Voth, 1998). There is also evidence collected by Clark that indicates that British workers worked at higher intensity than others (Clark, 1987a). Perhaps the most interesting alternative explanation of this phenomenon is the better level of nutrition that Britons enjoyed by this time, which permitted them to expend more energy in physical labor.[71] Whereas the Industrious Revolution hypothesis is demand based, the nutrition hypothesis should be regarded properly as supply-related insofar as it depends on increased availability of food. Below, I will discuss what kind of actual historical evidence exists to support De Vries's Industrious Revolution hypothesis.

Another aspect of the Industrious Revolution was an alleged increase in the participation of women and children in the labor market, which caused income as traditionally measured to increase (McKendrick, 1974). De

---

[69] De Vries points out that a change in preferences in favor of market-purchased goods would increase the marginal utility of money income. Yet reductions in the prices of these goods would have the same effect, because it is the <u>ratio</u> of marginal utility to the price of market goods that is the critical variable here.

[70] Some inventions, particularly those that revolutionized the household in the late nineteenth century, operated in the other direction. Thus the invention of the vacuum cleaner and the washing machine would lead to an increased production of these services by household members rather than buy them at the market.

[71] This point was first made in a pioneering paper by Freudenberger and Cummins (1976). For more recent work, see Fogel (1989).

Vries (1993, pp. 110-115) notes that the prominent role of woman and child labor during the Industrial Revolution represented a continuation and intensification of an already established trend toward greater paid labor force participation. As he points out, this movement did not start in 1750. It can be traced to the rise of market-oriented cottage industries ("protoindustrialization") in which women and children played a major role. As the Industrial Revolution progressed, the trend in labor force participation and contribution to household earnings seems to be subject to complex and often contradictory forces (Horrell and Humphries, 1992b). On the whole, the weight of the evidence suggest a rise in women and child labor beginning in the middle of the eighteenth century followed by a decline after 1815 or 1820, though the movements differed across regions and occupations. While the wives of elite workers in the formal sector such as factory operators and colliers could retreat to a quasi-middle-class homemaker's existence, those of outworkers and artisans experienced declining household income, forcing them to work harder. Yet at the same time the demand for their services declined. With income and substitution effects thus working in opposite directions, and with the labor force's structure changing, it is not surprising that the actual picture produced by the data is confusing.

Growing specialization and commercialization, an increasing reliance on the market, and the decline of "autoconsumption" preceded and accompanied the Industrial Revolution. As we have noted, to some extent these trends were themselves caused by the technical changes and to some extent they further stimulated additional technological changes. The idea of the "industrious revolution" is an important one, but it is not tantamount to restoring demand as a central factor in the economic changes that transformed the British economy. Much of the growing reliance on the market was supply driven, and although changing preferences toward market-produced goods buoyed demand for the products that the new technologies supplied, the contemporaneity of these two trends was only partial and to some extent accidental.

f. Foreign Trade

On the eve of the Industrial Revolution, Britain was in many ways an open economy. It exported close to 15 percent of its GNP. Exotic goods, brought in from Asia, South America, and Africa, were widely consumed. Grain moved into the country in years of scarcity and out in years of abundance. People, both emigrants and tourists, came and went. Capital moved in and out of the country with ease. Intellectuals corresponded with their colleagues overseas, and ideas -- technical and philosophical -- moved back and forth over the channel and the Atlantic. It seems natural that this openness would have been an advantage for Britain, setting it apart from such comparatively closed societies as Russia, Spain, or Turkey. Yet the mechanism linking this openness to the Industrial Revolution is far from clear. Part of the difficulty is that during most of the period of the Industrial Revolution political and military conflict disrupted the international economy. Between 1760 and 1815 only two short periods of peace (1763-1776 and 1783-1793) punctuated an otherwise long era of war, blockades, and embargoes. There is also a logical question how trade affected other variables such as industrial technology beyond the obvious consideration of the importation of essential raw materials.

The role of foreign trade in the British Industrial Revolution is hotly contested. Part of the confusion results from disagreements about what variable foreign trade is affecting. In principle, exports increase economic performance either through the employment of resources that otherwise would have been idle, and through the fact that these exports are exchanged for imported goods that the economy cannot produce as cheaply (or not at all). If the economy is at full employment, and if the endowments and technology of the economy are quite similar to those of its trading partners, the gains from trade could be limited even if exports constituted a very large proportion of output. Conversely, even a small level of international trade can have a huge impact on growth if it supplies a crucial missing ingredient to the economy. Hence any inference regarding the "importance" of exports based on what proportion of output was exported is highly suspect. In any event, such Smithian gains are inherently static; the Industrial Revolution constituted a change in the technology and institutions of production, and linking these directly to the level or rate of change of foreign trade is not transparent.

Some of the most prestigious scholars in the field have vehemently denied any essential role for exports. Harley (1994, p. 306) calculates the gains from trade as the difference between what Britain ended up paying for the goods she imported and what she might have had to pay had she been self sufficient, and concludes that this might have been on the order of 6 percent of National income by 1860, not a trivial sum, perhaps, put dwarfed by the growth of income in the previous century. Thomas and McCloskey (1981) start their essay by citing Deane and Cole to the effect that overseas trade was of central importance to the expansion of the economy and then add an ominous "we shall see," arriving ultimately at the conclusion that "the strongest effect between commerce abroad and industry at home was from industrialization to commerce, not the reverse. Trade was the child of industry." Trade theorists such as Charles Kindleberger (1964, pp. 264-266) and Ronald Findlay (1982) have come to the same conclusion. Many traditional historians are also of the same opinion, including the leading modern scholar of British overseas trade in this period, who writes:

> I share the view that overseas trade did not have an important *direct* role either in bringing about the Industrial Revolution or in supporting the first stages of its progress. . . . The initiative came from the supply side, from technical change. . . . Though a combination of changes made up the Industrial Revolution, the principal driving forces came from the nature of the inventions in the textile industry . . . and the efficacy of these inventions, which lifted the market for these inventions, at home and abroad, to an entirely new level. . . . Overseas trade made little contribution to the advent of the Industrial Revolution itself and was not essential in the early stages of its development. Its importance reappeared in the further expansion of the mature industrial economy (Davis, 1979, pp. 62-63).

Yet foreign trade as an essential impetus to the growth of the British economy is a tenacious concept. A recent paper by O'Brien and Engerman (1991) has tried to revive its importance by criticizing the assumptions made by economists who minimize the role of foreign trade. They appear to favor Adam Smith's "vent for surplus" theory of exports and even mercantilist ideas of "employment-creating" exports over the Ricardian notions of comparative advantage. They conclude that "domestic exports may be designated . . . as clearly important and necessary components of industrial growth that occurred in Britain in the eighteenth century" (p. 207). Javier Cuenca (1997, p. 16) has recently argued that “overseas demand in general provided the opportunity and the stimulus for technological innovation as the industry reached the limits of growth within a protected domestic market.”

At least some of the sharp differences of opinion that arise between O'Brien-Engerman and their opponents result from different formulations of the question. Foreign trade was necessary if Britain was to import goods she could not produce for herself or could produce only at enormous cost. Tropical groceries (sugar, tobacco, spices, tea), European foodstuffs (wine, dried fish, corn in years of high prices), and raw materials (timber, hemp, high-quality ores, tar, and of course raw cotton) had to be brought in from overseas. O'Brien and Engerman (1991, pp. 201-202) point out that for this reason, in a closed economy Britain's real income would have been substantially lower, though it is hard to know precisely by how much without specifying what the next best substitutes were. The first difference between an open and a hypothetical closed economy was the "gain from trade," and it was of course large because trade occurred in large part with economies whose factor endowments were radically different from Britain's. Harley’s calculation is important in underscoring the dangers of indispensability theorems in economic history, but it is difficult to see how Britons could have produced the raw materials for their textile industry and the tea and sugar for their breakfast from domestic resources.

Was the *growth* of exports an "engine of growth" in the period of the Industrial Revolution? The question seems somewhat moot, given that there is a growing consensus that growth itself was comparatively modest before 1831. The intellectual resources that have been dedicated to explain British economic growth before 1830 by growing exports may have been misallocated now that it turns out that this growth was far less impressive than was hitherto supposed. One way of testing the relationship is to see whether domestic supplies grew faster than foreign demand, so that foreign demand was in this sense more a passive than an active factor. After 1800, when more data become available, we can be more certain that British supply increased faster than foreign demand, because Britain's net barter terms of trade worsened continuously (Thomas and McCloskey, 1981, p.

101).[72] At the same time, it seems plausible that Britain's single factoral terms of trade (in which the prices are weighted by the productivity of domestic factors of production) improved, so that the purchasing power of the average Briton to buy imports continued to rise due to growing productivity.

Manufacturing products, of course, were exported in large quantities, and taken together foreign markets would have been difficult to replace. The ratio of industrial exports to gross industrial output increased sharply, from 24.4 percent to 35.2 percent between 1700 and 1760, a period in which output was growing only slowly. What happened subsequently? Whereas Crafts's figures suggest a sharp decline in the subsequent decades (with a sudden and unexplained peak in 1801), Cuenca's new computations draw a rather different picture.

TABLE 1.2. Exports Growth, 1700-1851

Year	Total Exports (£ millions)	Industrial Exports as a % of Industrial Product (Crafts)	Industrial Exports as a % of Industrial Product (Cuenca)	
1700	3.8	24.4	13[a]	20[b]
1760	8.3	35.2	18	28
1780	8.7	21.8	25	33
1801	28.4	34.4	40	40
1831	38.9	21.9	49	45
1851	67.3	24.7	69[c]	n.a.

a Column using Crafts estimates of Industrial output
b Column using Cuenca estimates of Industrial output
c Actual point estimate (all other Cuenca data are 11 year averaged centered on date).
SOURCES: Computed from Davis (1979, pp. 88-89), Crafts (1985a, p. 132) and Cuenca (1997, table 1).

Table 1.2 suggests that the importance of exports to manufacturing during the Industrial Revolution was most crucial in its "adolescent phase" after 1780. All the same, if export markets were more than just a trigger, their relative importance should have increased and not declined as the Industrial Revolution progressed. If Cuenca's new data are even approximately correct, the data seem to support his view. A closer look at Cuenca's time series does leave some questions open. Taking his own industrial output series as a denominator (which seems the reasonable thing to do), it turns out that the 11-year average of "official values" of industrial exports to industrial output is essentially flat at about 40-42% between 1800 and 1826. Moreover, during the period in which technological progress was at its most feverish (1760-85), Cuenca's official values ratio series was rising equally slowly, except for the years following the peace of Paris (1783). At first glance, therefore, the timing suggests that the causality may be running from technology to exports and not the reverse. Yet Cuenca clearly is correct when he complains that the movement over time of industrial exports relative to industrial output *cannot* be taken as evidence against the causal role of export markets.

[72] The somewhat more uncertain calculations made by Deane and Cole (1969, pp. 319-321) for the eighteenth century show a worsening of the terms of trade for the later 1780s and 1790s as well. This leads them to conclude (p. 83) that the "accelerated growth of foreign trade in the second half of the eighteenth century was associated with an *adverse* movement of the terms of trade."

Many scholars have argued that foreign trade did more for growth than the aggregate statistics suggest and that exports were more important in certain key industries. Cotton, above all, depended for more that half of its sales on foreign markets, and insofar as the technology developed for cotton spilled over to domestic industries, the foreign sector's role is understated by the statistics. O'Brien and Engerman also suggest that the wealth accumulated by merchants through foreign trade was invested in British manufacturing and overhead capital, though no evidence is provided to support this point and indicate how large this investment was.[73] O'Brien and Engerman resuscitate the Rostowian notion of a "leading sector" and designate industry in this role. Because exports were so important to manufacturing and because manufacturing dragged the other sectors behind it, they maintain, exports were essential to the entire economy, and "the attention should remain focused upon those forces promoting increases in the production of manufactured goods" (1991, p. 208). Apart from the somewhat poorly defined concept of a "leading sector," the problem with their logic is that it is consistent with any set of facts and thus lacks power as an explanation. When exports stagnated in the 1760s and 1770s, just when a number of key industries were taking off, "domestic demand maintained the growth of industry," which proves that "interactions also flowed the other way" (p. 208).[74]

A different way in which exports could have led to growth is if export industries employed labor that would otherwise have been unemployed. Thomas and McCloskey base their thesis on the "unimportance of exports," on the simple notion that exports used up valuable resources that could have been used for the benefit of domestic consumers but are the inevitable price a country has to pay for the imports it enjoys. This assumes, however, a fully employed economy in which each factor is paid its opportunity cost. Many of the manufactures of Britain during the eighteenth century, however, were produced by rural industry, by men and women whose opportunity cost in the off-season was low. Insofar as export markets provided these workers with employment, an expansion of output can indeed be attributed to exports. In other words, in a closed economy the same employment levels might not be sustainable, so that one of the benefits of trade was an increased demand for labor. It is difficult to prove this point decisively, but O'Brien and Engerman are correct to point out that contemporaries were far from impressed by the success of the domestic economy's in maintaining full employment and were obsessed by the specter of unemployment. As we have already seen, there is some evidence that people in Britain worked longer hours in 1800 than they did in the middle of the century.

Turning to the *dynamic* question, as already noted it is much more difficult to connect the openness of the British economy with technological changes. It is transparent that technological advances will stimulate exports. But is there a feedback from rising exports to further technological progress? As Krugman (1995, p. 55) has pointed out, once we consider such a feedback effect the answers become much more controversial. Export demand may have been a consideration for some innovators, but almost every individual entrepreneur could cover his expenses by the domestic market. The growing dependence of the cotton industry on foreign markets was an ex post phenomenon, not something that *caused* technological change. Ralph Davis argues that cotton expanded overseas after it had earned its spurs in the domestic market and that the export-driven expansion of the industry in the 1790s simply called for a larger number of similar mills (Davis, 1979, p. 67). All the same, the microinventions that kept improving the quality and reducing the prices of the goods produced may have been a function of output and thus of the size of the market. Learning by doing and experience were the sources of productivity increase after the big breakthroughs had been made. Insofar as export markets permitted expanded sales, then, they led to productivity increases and lower costs. Export-oriented industries in the post World-War

[73] There is even less evidence for the statement that merchants "created and widened markets for British manufactured goods at home or abroad" (O'Brien and Engerman, 1991, p. 191), nor is there any suggestion as to exactly how merchants *create* markets as opposed to servicing them.

[74] The concept of a leading sector itself may prove to be more lasting than the "take-off hypothesis. Wijnberg (1992, pp. 165-167) defines a leading sector as an industry that is "technologically contagious," that is, in which the conditions for successful innovation such as low barriers to entry and appropriability of inventions spill over onto others. Such explicitly dynamic models are necessary to a consistent "demand-side" interpretation.

II Asian economies have often been "high-tech" and so a large export market may produce a stimulus to the adoption of frontier technologies. The unresolved questions remain, however: Is this connection between exports and technological progress also true for a nation that is generating the new technologies, and not only adopting them? To what extent would the domestic market have been able to replace the foreign markets? What was the elasticity of cost with respect to sales (that is, how strong, really, were the marginal learning effects of overseas sales)?

Even if the nexus between foreign trade and technological progress thus remains something of a mystery, the open-ness of the British economy was a central feature that determined her economic fate. Open-ness is not a yes-or-no variable: few economies have ever been hermetically closed and few have been "entirely open" (if that concept could be defined). While open-ness was thus a matter of degree, this degree was of great importance. One example is the role of agriculture in the industrialization process. In a recent paper, Matsuyama (1992) demonstrates rigorously an intuition long prevalent among economic historians, namely that the relation between agricultural productivity and the rate of industrialization depends on the open-ness of the economy. In a closed economy, manufacturing depends on productivity growth in agriculture to produce a surplus that will permit the reallocation of resources from farming to industry and to provide a market for manufactured products. It has often been thought that an "agricultural revolution" was a necessary precondition for industrialization. Yet in an open economy this is clearly false: food can be imported and paid for by industrial goods. In fact, in an open economy a highly productive agricultural sector signals to the economy that its comparative advantage lies in farming, thus losing the (unforeseen) advantages of industrialization. This is in fact what happened in the Netherlands between 1815 and 1870: an open, free-trade economy with a highly productive agricultural sector, the opportunity costs of labor was just too high to render manufacturing profitable (Mokyr, 1976a). In Britain, despite growing agricultural productivity (the dimensions of which are still heavily disputed) this did not happen. Imports from the Celtic Fringe and the Continent made up the British food deficit (Thomas, 1985). Indeed, Matsuyama's model implies that in an open economy the Industrial Revolution occurred not *because* but *despite* the growth in agricultural productivity.

The open-ness of the British economy also meant that technology was continuously stimulated by ideas from the outside. We have already seen the wide influence of French science and inventions on British technology. Throughout the period, close cooperation with French, German, and Swiss manufacturers led to the continuous exchange of technological knowledge. Arnold Pacey (1990, pp. 117-120) has argued that Asian stimuli were of primary importance to the Industrial Revolution. Indian calicoes and muslins could not be made in Britain using the laborious hand-spinning techniques of India, but they showed the British what could be done, and eventually Crompton's mule was able to produce yarns of Asian fineness. English entrepreneurs sent representatives to Smyrna to study the manufacturing of Turkey-red dye, and plants to produce it were set up in Manchester and Glasgow (Wadsworth and Mann, 1931, pp. 180-181). Technology was enriched by the infusion of foreign elements, and in the long run this exposure effect turned out to be one of the most lasting benefits of the open economy.

A separate issue often raised in this context is the impact of the British Empire. It seems somehow tempting for those who do not have much sympathy for British capitalism to link it with imperialism and slavery.[75] It is hard to see exactly how the imperial policies, which protected British merchants doing business overseas, could have had much impact on the Industrial Revolution beyond, perhaps, assuring favorable treatment in some markets. Empire and foreign policy seem to have conveyed at best a slight advantage. After all, Britain lost one of its richest colonies during the early stages of the Industrial Revolution, and yet after 1783 commercial relations with the young United States were none the worse for wear until complications in Europe drove the two apart again. India was an important market, but it never reached the size that would make it a sine qua non: In 1784-1786 Asia (that is, primarily India) absorbed 13.3 percent of British exports, a share that remained essentially constant until 1854-1856 (Davis, 1979, pp. 96, 100). To be sure, Asia did buy a larger than

[75] As Engerman (1972, p. 430) put it, in this version history becomes a morality play in which one evil (the Industrial Revolution) arises from another, perhaps even greater evil, slavery and imperialism.

proportionate share of the output of Britain's most dynamic industry, cotton, but as late as 1854-1856 it bought 22.5 percent of Britain's cotton exports. This is substantial, but Europe, the Near East, the United States, and Latin America, where Britain competed on an equal base, remained equally important markets. Outside Britain, Switzerland and Belgium, two nonimperial nations, were successful industrializers, whereas Holland and Portugal, which controlled a large and rich set of colonies, remained behind. In short, trade with the empire may have been central before the Industrial Revolution, but it lost much of its primacy in the years after 1780, when it might have been needed the most (Cain and Hopkins, 1980, p. 474).

The classic attempt to link imperialism and the slave trade with the Industrial Revolution is the Williams thesis. Eric Williams (1944) argues that profits from the triangular trade (between Western Europe, Africa, and colonial America) helped finance the early stages of industrial capitalism. In particular, Williams argues that the slave and sugar trades encouraged British industrial production and capital accumulation.[76] This thesis, which had long been regarded as discredited, has recently been resurrected, and a special issue of *The Journal of Interdisciplinary History* was dedicated to it (Inikori, 1987; Richardson, 1987; Solow, 1987; see also Inikori, 1989, Morgan, 1996). It can hardly be doubted that the West Indian sugar trade was highly profitable, as Adam Smith ([1776] 1976, p. 412) pointed out. Because the sugar trade depended on slave labor, the slave trade was, not surprisingly, profitable as well, although the high mortality of slaves and crew on these voyages raise some question as to the profitability of this activity (Morgan, 1996, p. 17). Commercial interests, shipbuilding, banking and insurance services, and industries catering to the triangular trade prospered, and the towns of Bristol and Liverpool consequently grew (Williams, 1944, p. 36, 62-64). Yet the links between Liverpool and Manchester do not prove Manchester's "tremendous dependence on the triangular trade" (p. 68) and recent work has not been very successful in substantiating Williams's claim that the profits from this trade "provided one the main streams of that accumulation of capital in England which financed the Industrial Revolution" (p. 52). The intuitive feeling that "the exploitation that really mattered was [that] of African slaves" (Solow, 1987, p. 737) is justifiable in that it surely mattered to the slaves themselves, as it did to Africa and to the areas to which slaves were shipped. Yet that does not necessarily mean that it "mattered" to the same degree to Britain and other European economies that were the main beneficiaries of the triangular trade system.[77] There is often a cruel asymmetry in the moment of injustices in the respective histories of the victim and the perpetrator, an asymmetry as clearly illustrated by the economic relations Britain had with the Caribbean as by its relation with Ireland.

Furthermore, the simple causal links drawn by Williams should be modified. Richardson (1987) points out that the slave trade depended on the demand for sugar, which itself was a function of economic changes in the sugar-consuming economies in Western Europe. Moreover, Ralph Davis's estimates show that in 1784-86 the West Indies purchased 11.3 percent of British manufacturing exports, rising to 19 percent in 1804-06, but falling back to 15.1 percent in 1814-16 and 9.1 percent in 1834-36. These numbers are not insubstantial, but they do not prove that the West Indies were more than just another market. Above all, however, the West Indies and slavery were important to Britain as a source of products that could not be produced locally. In the absence of

[76] As has been often noted, it is not quite clear whether Williams referred to the slave trade alone or to the more extensive triangular trade.

[77] In a classic paper, Engerman (1972) demonstrated that the quantitative effects of the slave trade on the British Industrial Revolution were negligibly small. He computed (p. 440) that total profits from the slave trade in 1770 amounted to at most £342,000 (an alternative estimate has the number as low as £44,000). Total GNP in 1770 can be roughly estimated at about £166 million. (computed by applying Crafts's revised growth rates to Dean and Cole's estimate of GNP at £232 million in 1801). Gross capital formation was between 6 and 7 percent of GNP and thus came to about £11 million. Even on the most favorable assumptions, then, the profits of the slave trade, had they all been invested in Britain, would have contributed no more than 3 percent of capital formation in 1770.

West Indian slavery, Britain would have had to drink bitter tea, but it still would have had an Industrial Revolution, if perhaps at a marginally slower pace.[78]

If an element of slavery and the Atlantic trade were of essential importance to the Industrial Revolution, it was not due to the West Indies but to slavery in North America. Before 1780 most of the raw cotton came to Britain from the West Indies, but clearly their potential to supply it was limited, and after 1790 the industry depended increasingly on the southern states of the United States. Simply put, without U.S. slave labor it is hard to see how the elastic supply of raw cotton would have been secured. Certain processes in the cotton industry could be mechanized, including some concerned with the production of the raw material (using, for example, the cotton gin). But the planting and picking of cotton in the fields of the southern United States remained a manual process, and as the demand for cotton increased, U.S. slave plantations rapidly switched to cotton. Without U.S. slavery, the British cotton industry would have run into a severe bottleneck. It is here and not in the consequences of eighteenth-century triangular trade that slavery truly "mattered" for the Industrial Revolution.

g. Science and Technology

The notion that Britain was the first to undergo an Industrial Revolution because somehow British technological success was due to Britain's having more "advanced" science is unsupportable. The premise itself is in dispute (Kuhn, 1977, p. 43), but even if it were true, the technology developed during the British Industrial Revolution owed little to scientific knowledge, as Mitch's chapter below stresses. The inventions that set the British changes in motion were largely the result of mechanical intuition and dexterity, the product of technically brilliant but basically empirical tinkerers, or "technical designers" (a term suggested by Hall, 1974, p. 148), such as John Wilkinson, Richard Arkwright, John Smeaton, Richard Trevithick, and Robert Stephenson. In a few cases, such as Claude Berthollet's chlorine bleaching and Humphry Davy's safety lamp, inventions were made by scientists of note, but that correlation does not prove that science itself was of great importance. Leading scientists were not wholly specialized at this time and dabbled in technology, just as Galileo, Huygens, Hooke, and Leibniz had a century earlier.[79] Unlike the technologies that developed in Europe and the United States in the second half of the nineteenth century, science, by conventional wisdom, had little direct guidance to offer to the Industrial Revolution (Hall, 1974, p. 151). Gillispie (1957) points out that the majority of scientific endeavours of the time concerned subjects of limited technological use: Astronomy, botany, crystallography and early exploration of magnetism, refraction of light, combustion. Eventually, of course, many of those discoveries found economic applications, but these took place, with few exceptions, after 1830.

If science played a role in the Industrial Revolution, it was neither through the "pure" foundation of technology on scientific understanding nor through the role of scientists in invention but rather through the spillovers from the scientific endeavor. We may distinguish between three closely interrelated phenomena: scientific method, scientific mentality, and scientific culture. The penetration of scientific *method* into technological research meant accurate measurement, controlled experiment, insistence on reproducibility, and systematic

---

[78] In a recent manuscript, Pomeranz (1998, ch. 4) re-assesses the Williams thesis and, much like Morgan (1996), concludes that while there may be something to the argument that profits from the West Indies ended up paying for some portion of the Industrial Revolution, the exact magnitude of this effect is hard to determine but is unlikely to be large. Pomeranz, however, also makes an important observation in that the main effect of the trade with America was the saving of land (ibid., ch. 5). The importation of sugar into Britain alone, he computes, saved by 1800 somewhere between 1.3 and 1.9 million acres just in terms of its calorific value. Adding to that the equivalent of other products of Britain's “ghost acreage” overseas, he computes the total equivalent of its colonial trade to be 3-4 million acres, adding around 20 percent to Britain 17 million arable acres. In this way, he maintains, Britain was able to avert the ecological consequences of its population growth. While such computations illustrate neatly the gains from trade between Britain and its colonies, they shed little light on the question why the Industrial Revolution occurred where and when it did.

[79] The two leading Newtonians of the early eighteenth century, the Dutchman Willem Jacob s'Gravesande and the Englishman (of French descent) Jean Desaguliers, were both active in introducing and improving Newcomen engines in continental Europe (Jacob, 1988, p. 130).

reporting of methods and materials using a shared vocabulary and standards. Even more important, perhaps, was scientific *mentality*, which taught engineers a rational faith in the orderliness and predictability of natural phenomena -- even if the actual laws underlying chemistry and physics were not fully understood (Parker, 1984, pp. 27-28). The growing belief in the rationality of nature, the archtypical enlightenment belief, led to a growing use of mathematics in pure science as well as in engineering and technology. This scientific mentality also taught engineers the "method of detail," analyzing technical problems logically by breaking them into components that could be more easily analyzed separately than as part of a whole (Pacey, 1975, p. 137). Engineers such as Thomas Telford, John Smeaton, and John Rennie moved effortlessly between experimental science and practical applications. George Stephenson, a remarkable example of this ability himself, wrote of the great Smeaton as having a "truly Baconian mind" -- a description that fits an entire class of British engineers active between 1760 and 1830. Scientific mentality also implied an open mind, a willingness to abandon conventional doctrine when confronted with new evidence, and a growing persuasion that no natural phenomenon was beyond systematic investigation. Finally, scientific *culture*, the culmination of Baconian ideology, placed applied science at the service of commercial and manufacturing interests (Jacob, 1988, 1997). As we have seen, lectures on scientific subjects drew eager audiences who met at provincial scientific society meeting places such as the famous Birmingham Lunar Society, coffeehouses, and masonic lodges to watch experimental demonstrations illustrating the application of scientific principles to pumps, pulleys and pendulums. Yet, as Robert Schofield (1972) has argued, these meetings were secondary to the networking and informal exchange of technical information between members. Scientific culture reinforced the entrepreneurial interests of science's audience by demonstrating how applied mechanics could save costs and enhance efficiency and thus profits. Much of this "provincial scientific culture," as Inkster (1991, p. 43) has called it, was private, meritocratic, non-elitist and thus in some ways in conflict with the social establishment. As Thackray (1974) has shown for Manchester, the interest in science was a means for the upstart commercial and manufacturing class to assert and legitimize itself. Because science was a natural and not a moral discourse, it provided a neutral common ground for otherwise hostile subgroups of the urban elite to communicate and express a “cultural solidarity and social cohesion” to set them apart from both the working class and the landed elite (p. 693). All in all, it was one of the most obvious ways in which “culture” affected technology and, in the long run, economic progress.

Eighteenth century science tried to provide implicit theoretical underpinnings to what empirically minded technicians did, even if the complete scientific base had not been fully worked out. Thus the steam engine depended on the understanding of atmospheric pressure, discovered by Continental scientists such as Evangelista Torricelli and Otto von Guericke, which somehow must have filtered down to Newcomen despite the fact that his world was the local blacksmith's rather than the cosmopolitan academic scientist's. Chlorine bleaching depended on the discovery of chlorine by the Swedish chemist Carl Wilhelm Scheele in 1774. Phlogiston theory, the ruling physical paradigm of the eighteenth century, was eventually rejected in favor of the new chemistry of Lavoisier; but some of its insights (e.g., the Swede Tobern Bergman's contributions to metallurgy) were valuable even if their scientific basis was flawed and their terminology quaint to modern readers. Some of the insights were valuable. Cardwell (1972, pp. 41-43) has shown that the idea of a measurable quantity of "work" or "energy" derives directly from Galileo's work on mechanics. The advances in water in the eighteenth century depended on a scientific base of hydraulics (Reynolds, 1983). Often, of course, bogus science produced bogus results, as in Jethro Tull's insistence that air was the best fertilizer. In the "development" stage of the basic inventions, in which engineers and technicians on the shopfloor improved, modified, and debugged the revolutionary insights of inventors such as Cort, Cartwright, and Roberts to turn them into successful business propositions, pure science played only a modest role.

Beyond these direct links, science and the scientific community created the cultural and intellectual background for the tinkerers, the mechanics, and the engineers who made the Industrial Revolution (Jacob, 1988, chap. 5; 1997, ch. 5). The scientific revolution of the seventeenth century taught a new approach to the study of nature, a mechanical philosophy in which natural phenomena were studied as independent units, increasingly separated from religious considerations. Without immediately abandoning the belief in a creator, it became increasingly possible to analyze nature without theology or magic. Because technology in its deepest essence involves the manipulation of nature and the physical environment, the metaphysical assumptions under which people operate are ultimately of crucial importance. Science in the seventeenth century became increasingly permeated by the Baconian notion of material progress and constant improvement, attained by the accumulation of knowledge. Although such relations are impossible to quantify, it stands to reason that in that regard science laid the intellectual foundations of the Industrial Revolution by providing the tacit and implicit assumptions on which technological creativity depended.

British science and scientists occupied a different position in society than elsewhere. As Thomas Kuhn states, the old cliché that British science was pragmatic and applied whereas French science was abstract, deductive, and formal seems to have survived the test of time (1977, p. 137; see also Inkster, 1991, p. 42). The origins of this phenomenon may be traced to an intellectual bifurcation of the seventeenth century, when British science came under the influence of Bacon whereas in France more Cartesian ideals triumphed. Bacon advocated that the purpose of science was to raise comforts and living standards, whereas the French traditions followed more lofty objectives. Bacon's science was empirical, experimental, and pragmatic whereas French science was theoretical and abstract. Such generalizations are inevitably hazardous, but water power provides at least one persuasive example. In Britain research on water power was conducted by practical engineers, such as Smeaton, John Banks, John Rennie, and William Fairbairn, in search of a better water mill. On the Continent work on water power was largely theoretical and carried out by mathematicians, such as Antoine Parent, Johann Euler, and Jean Charles Borda (Reynolds, 1983, pp. 196-265).

Yet the roots of the divergence between British and Continental science go beyond that. The Cartesian traditions in eighteenth-century France regarded the function of science to be to support the authoritarian state as the source of all order. In Britain, as Margaret Jacob (1988, p. 93; 1997, p. 113) has argued, the scientists in the 1660s and 1670s forged an alliance with the landed and commercial interests. After these interests triumphed politically in 1688, scientists in eighteenth-century Britain were on the whole part of the economic establishment, not of the opposition. They regarded it as a natural state of the world to cooperate with engineers and manufacturers to solve pragmatic technical problems.[80] The interactions between them, as we have seen, were institutionalized in the various scientific and philosophical societies that provided the meeting places, and informal contacts further strengthened these ties. Even some members of the landowning elite displayed a strong interest in technology, in part for economic reasons but also out of sheer curiosity. The Earl of Dundonald, the Viscount of Dudley and Ward, the Earl of Balcarres, and others were fascinated by the new technologies. There was a growing communication between scientists, engineers, and businessmen, and they engaged in a common effort to recognize technical problems and solve them. From the early eighteenth century on, scientists in Britain gave popular lectures on mechanical and technical issues, which were widely attended by audiences from the commercial and artisanal classes. The most famous of these lecturers was John T. Desaguliers, a noted physicist who made considerable contributions to the Newcomen steam engine and water power (Musson and Robinson, 1969, pp. 37-45; Reynolds, 1983, pp. 215-217, 280).

The state and official institutions in Britain had relatively little to do with these developments, the Board of Longitude being the most notable exception. The generation and diffusion of scientific and technological knowledge in Britain occurred spontaneously, by and for private interests. In France, by contrast, scientists depended on economic and personal relations with the *political* establishment, fostering an elitist and statist approach to science, which was thus particularly concerned with the engineering and technical needs of the state and above all with military needs. The French state subsidized and managed scientific enterprises, whereas in Britain the same role was carried out by the private sector (Gillispie, 1980, chap. 5). The counterparts to the British provincial societies in disseminating technical knowledge were the *grandes écoles*, which trained technicians and engineers. The first of these was the *école des ponts et chaussées*, founded in 1744, followed by the *école de dessin* in 1767 and the *école des mines* in 1783. After the Revolution, these were followed by the *école polytechnique* (1794) and the *école des arts et métiers* (1804) (Artz, 1966). All these institutions were run and funded by the government.

---

[80] Perhaps the best example is the problem of measuring longitude at sea. In 1714 the Longitude Act promised huge financial awards to any individual who could devise a method or tool to measure longitude accurately. The commercial interests here were quite transparent, and applied science -- in the person of watchmaker John Harrison -- proved up to task (Sobel, 1995).

In other countries, such as the Austrian Netherlands, the German states, and Russia, the direct intervention of the state was even more noticeable. Science and engineering were creatures of the state, meant first and foremost to serve the military and administrative organs of the government. In Britain, private interests dominated.

The difficulty in linking science and technology in this period is highlighted by one of the few quantitative measures of scientific output --periodicals. Although the value of a periodical is of course proportional to its subject matter, the quality of research, and the scope of its circulation, it is striking that the vast majority of scientific journals published in the eighteenth century appeared not in England or France but in Germany. Kronick shows that over 61 percent of all "substantive serials" appeared in Germany, with France and England accounting for 10.7 percent and 6.9 percent, respectively. The actual gap was smaller, because German scientific journals were comparatively short lived, but correcting for this does not alter the picture (Kronick, 1962, pp. 88-89). Similar gaps, although not as large, hold for the proceedings of scientific societies. The only category in which England led, perhaps significantly, was "translations and abridgements" (pp. 114-115).

On the basis of this background it is easier to understand the dispute between those like Mathias (1979, pp 45-87) and Hall (1974; see also Hall and Hall, [1964] 1988, p. 219), who deny science any serious role in the Industrial Revolution, and those like Musson and Robinson (1969) and more recently Jacob (1988, 1997), who try to restore science's role in explaining Britain's uniqueness. David Landes (1969, p. 104) and others have reversed the causal connection and maintaine that science owed more to technology that the other way around. The conventional argument that scientific knowledge was unimportant simply because much of what scientists knew was irrelevant to engineers and industrialists can no longer be maintained. Yet Jacob (1988, p. 181) may have gone too far in the other direction when she suggests that the Industrial Revolution occurred in Britain and not in France and the Netherlands because the lack of scientific knowledge on the Continent was such that there "many of the very men who had access to capital, cheap labor, water, and even steam power could not have industrialized had they wanted to: they simply could not have understood the mechanical principles necessary to implement a sophisticated assault on the hand manufacturing process." Certainly there was nothing in the inventions made between 1760 and 1830 that exploited a store of knowledge accessible only to the British. The physics and chemistry of the time were primitive, and the deeper theoretical principles behind such breakthroughs as the steam engine and soda making were not understood by anyone. France could and did generate highly sophisticated innovations, including the mechanical toys of Vaucanson, the Jacquard loom, the continuous paper machine, as well as the chemistry of Lavoisier and Berthollet. The difference, if there was any, was of degree, of emphasis, and above all of the depth with which technologically valuable knowledge had penetrated into the productive layers of society. More recently, Jacob (1997, pp. 132-33) has argued sensibly that some mechanical knowledge had to be part of one's mental world before mechanical devices could be invented and exploited, and that differences in scientific education across Europe went a long way toward explaining national differences in industrial progress. Yet the *average* level of scientific knowledge may not have differed all that much between England and France. What differed was its distribution, and its impact on the mundane needs of the “useful arts.”

What accounts for the differences in the intensity of interaction between persons with knowledge and persons of business? Every civilized society contains individuals who are highly educated and think for themselves, and individuals who are skilled and produce goods and services that add up to income and consumption. Technologically creative societies are those in which these two classes mingle socially, communicate with each other, and are interested in similar issues. In Britain the bridge between natural philosophers and engineers was broader and easier to cross than in other countries, and more than anywhere else, Britain could count on able people who could effortlessly move between the world of abstraction, symbol, equation, blueprint, and diagram and the world of the lever, the pulley, the cylinder, and the spindle. Information also travelled easier in eighteenth century Britain than in France, thanks to better passenger travel and mail services (Szostak, 1991). Yet such bridges existed elsewhere (as one glance at Diderot's *encyclopédie* will demonstrate), and Britain's advantage here was as partial as it was temporary.

The Inputs: Labor and Capital

Economic growth can be decomposed into increases in the quantities of inputs and changes in the way inputs are utilized. Increases in output per worker consist of changes in productivity and the accumulation of factors other than labor. Separating the two is in practice quite difficult. A related question, equally controversial and studded with theoretical pitfalls, is the effect of the initial endowments of factors of production on the rate at which the modern sector grows. A satisfactory model would allow us to approach the question, Why was Britain first? from a different angle. The Industrial Revolution involved massive accumulation of capital and profound reallocation of labor. How did factor markets carry out these functions? How crucial was the supply of factors to the Industrial Revolution? Where did the inputs come from, and how did market mechanisms channel them to where they were needed?

The operations of factor markets in Britain during the Industrial Revolution have been examined by Williamson (1987a; 1990a, chap. 7). In this work, Williamson poses the question starkly: How much did the imperfection of labor and capital markets cost the British economy? Questions such as, Were markets perfect? or Did they fail? are somewhat ill posed; factor markets are far from perfect even today, and "failure" is obviously in the eye of the beholder. Williamson's approach is to compare the actual operation of factor markets with an ideal neoclassical world in which competition is perfect, factors flow effortlessly between regions and sectors and the allocation of resources follows the theoretical rules devised by economists. The latter, purely imaginary world is, obviously, more efficient, but theory gives us no guide as to the *size* of the difference. Williamson reasons that if the forgone output due to factor market imperfections was very large, it could conceivably have slowed down industrialization and growth. Working with a multisectoral general equilibrium model, he poses counterfactual questions: How much faster would GNP have grown and industrialization have proceeded if factors had been perfectly mobile? He concludes that these gaps were indeed significant. The labor market imperfection alone was responsible for a 3.3 percent loss of GNP compared with the ideal world, and the capital market for an 8.2 percent loss.[81] More important, manufacturing output would have increased over 60 percent if the capital market imperfection had been eliminated, and manufacturing profits by 114 percent. It should be noted that Williamson's imperfections are intersectoral only; his computations still assume that capital and labor can reallocate themselves effortlessly *within* the sectors. In that sense they represent an understatement of the true values.

Apart from the question of how efficient factor markets were, the roles of labor and capital have been the subject of an interesting and important literature.

a. Labor

There are two competing and apparently incompatible views of the role of labor in the Industrial Revolution. One of them sees labor as a scarce resource, in fact as *the* scarce resource, and therefore the Industrial Revolution had a better chance of succeeding in areas in which it was abundant and cheap. The other regards technology as responding to labor scarcity and thus implies that *scarce* labor was an advantage in the industrialization race.

The first model is based on a number of assumptions that should be spelled out.[82] Because the model is not strictly speaking a growth model (it has few implications for the overall growth rate of the economy) and deals more with the composition and technological practices of some sectors, I termed it a "growing-up" model (Mokyr, 1976b). The assumptions are as follows:

1. Capital goods "embodied" the new technology. Then, as now, that assumption seems almost too obvious to justify. Steam engines, mule jennies, blast furnaces, paper mills, chaff cutters, and threshers are all examples of a new technology requiring a large capital expenditure. One cannot have the new technology without

---

[81] The combined loss does not equal the sum of these losses because in the model one market can adjust to compensate for imperfections in the other. Indeed, in Williamson's story eliminating both gaps would result in a lower gain in aggregate output than eliminating the capital market gap alone, an anomalous result he ascribes to nonlinearities in the solution of the model.

[82] Some of the following material is adapted from Mokyr, 1991b.

making an investment in the equipment that embodied it. Above all, there were factories that had to be built, maintained, heated, lighted, and guarded. The modern sector was physically located, by and large, in large buildings. And in contrast with France and Belgium, in Britain there were no more monasteries to confiscate and convert. This is not to deny the importance of disembodied technological change. It implies, however, that a lack of fixed capital could have retarded the transformation, as I shall argue later. The reverse does not hold: An abundant supply of capital did not guarantee the adoption of technological changes and the emergence of factories, because the owners of the capital could not be relied upon to lend it to aspiring factory owners. What mattered was venture capital, not aggregate savings.

2. The rate of accumulation depended crucially on the rate of profit. In the simplest model, in which factory owners could not borrow and depended on retained profits to finance new investment, this conclusion is trivial. In models with financial institutions, however, this relation is not appreciably weakened as long as the past performance of the firm is used as an indicator of its future profitability.

3. Wages were the main cost to the firm. If labor productivity is primarily determined by technological parameters and the prices of output are given, the rate of wages is inversely related to the rate of profit through the factor price frontier. In other words, because the productivity of labor depended on the technology in use, assumed to be accessible to all economies, the main reason why profit rates differed across economies was different wage levels reflecting differences in economic structures or factor endowments on the eve of the Industrial Revolution.

4. Technological change was more or less independent of factor prices. This would be the case if there was little choice in the range of techniques; i.e., the "best practice" techniques at the onset of the Industrial Revolution were the most efficient for any realistic set of factor prices.

5. Goods were internationally mobile, but labor and capital were not. It is assumed that labor was mobile only *within* a region but could not migrate across economies. Neither of these assumptions exactly conforms to the historical experience; they are made only to simplify the story. Hence, if there were no important differences in the propensities of capitalists to reinvest profits in their firms, the model predicts that areas that for some reason started off with low wages would, all other things being equal, undergo an Industrial Revolution at a faster rate.

The growing-up model is different from the standard growth models in that it is a disequilibrium model. Its dynamics depend on the coexistence and interaction of the "old" and the "new" technologies. It applies to the European as well as to non-European contexts (Pomeranz, 1998). The traditional sector, which produces the same good (or a close substitute) as the factories, can continue its existence for a long time after the process has started, because the modern sector is still too small to supplant it altogether. As long as the two sectors coexist, the modern sector earns a "quasi-rent," a disequilibrium payment that will eventually disappear when the manual industries have disappeared. Through continuous reinvestment, this rent in its turn provides the fuel for further growth of the modern sector. This model suggests that high-wage economies would have lower profits, lower rates of accumulation, and thus a slower and later Industrial Revolution. The model also predicts that wages in the modern sector would grow slowly if at all as long as the traditional sector remained a large employer. In this sense, the model is comparable to the labor surplus models of Lewis and Fei-Ranis popular in the 1970s. In contrast to those models, however, the "growing-up" model does not have to make any *deus ex machina* assumptions about the wage rate. The modern sector is small enough relative to the rest of the economy to take the wage parametrically (that is, the sector can hire workers at a wage rate that is unaffected by the number of workers it employs) and hence the lower the wage set in the traditional economy, the faster the modern sector could grow.[83]

[83] The logic of the model has since been adopted by other writers interested in other regions. In his work on the cotton industry in the South, Gavin Wright (1987) explicitly points to the South's emerging as a "low-wage region in a high-wage economy" as the main reason for the South's success in establishing a successful textile industry after 1880 (pp. 76, 124). Much in Wright's analysis of the postbellum southern industry has analogues in the growing-up model, especially his assumptions about labor and capital markets.

The second approach to the role of labor in the Industrial Revolution, most closely associated with the work of H. J. Habakkuk (1962), maintains that inventive activity in the nineteenth century was mostly labor saving and that scarce labor thus stimulated waves of technical change. This approach is based on a somewhat peculiar view of technological change, namely, that innovation was a process of choice between more or less equivalent alternatives, similar to the choice made by a firm facing an isoquant. Although Habakkuk was primarily concerned with the period after 1830, his approach extends naturally to the British Industrial Revolution. High wages and labor-supply constraints in Britain, in this view, stimulated the demand for labor-saving technological change (Landes, 1969, pp. 57-60). Yet the application of the model, at second glance, is fraught with difficulties. To start with, it is far from obvious that technological change during the Industrial Revolution was, on balance, more labor saving than capital saving: Von Tunzelmann (1981, p. 165) believes that, on balance, it was about neutral. MacLeod, examining the declared motives of eighteenth-century English patentees, found that only 3.7 percent of them stated that "labor saving" was the main purpose of the invention. Further, it always makes good sense to "search" for labor-saving innovations, even in low-wage economies, because labor always costs something, and thus innovations that reduce labor inputs increase profits. This is especially the case if, as was likely true in low-wage areas, production was highly labor intensive. In addition, as David (1975) has pointed out, the Habakkuk view implies that technological change is "localized" (that is, occurs in close proximity to the techniques actually used rather than over the entire range of feasible techniques). For the Habakkuk view to prevail, such localized technological change has to be stronger in the capital-intensive range of techniques than in the labor-intensive range. In that case a high-wage economy will naturally have chosen a less labor-intensive technique and will experience faster technological progress as the unintended by-product of this choice. Finally, although British wages were higher than on the Continent, some scholars (e.g., Flinn, 1966, p. 31) have insisted that the growth of population met the increased demand for labor and that there is no evidence for any labor scarcity.

Economists have examined the assumptions on which the two alternative theses are based and have made them explicit.[84] The seemingly obvious test of the low-wage hypothesis is that the areas of Britain that industrialized earliest should, at the outset, have had lower wages. The relevant variable here is *nominal* wages, because we are interested in the cost of labor, not in the standard of living. In this regard, at least, the hypothesis seems confirmed. The areas of Britain that industrialized first, the northwest counties of Lancashire and the northern midlands, had lower wages than the South in the middle of the eighteenth century (Hunt, 1986). During the Industrial Revolution this relation was reversed, so that by 1867 the industrial areas had higher wages. Yet although this pattern is repeated in a few other instances, such as the Low Countries (Mokyr, 1976a), it is far from universal. Ireland, by all accounts, had low wages but did not industrialize. Britain itself had higher wages than most of the European continent.

It must be, then, that the *ceteris paribus* clause in this model did not always hold. For instance, it is important to ask *why* labor was cheaper in one place than in another. If it was purely a matter of opportunity cost, as the growing-up model assumes, the implication that capital accumulation is faster follows. But if labor was cheaper in one place because it was less productive, the model encounters a difficulty. If wages were low because labor quality and thus productivity were low, the advantages of cheap labor vanish. Contemporary authors were aware of this. Arthur Young, writing in the late 1780s, notes that "labour is generally *in reality* the cheapest where it is nominally the dearest" (Young, 1790, p. 311).[85] In a paper dealing with a somewhat later period, Gregory Clark (1987a) shows the strong correlation of labor productivity with nominal wages, even using

---

[84] The literature stimulated by Habakkuk's pathbreaking book is quite extensive. See, for example, Landes (1965); Rosenberg (1963, 1967); Saul (1970); Temin (1973). Most of the debate is carried out in the context of Anglo-American differences, with Britain, interestingly enough, considered the *low-wage* economy (though in the period of the Industrial Revolution it would, relative to the rest of Europe, be the high-wage economy). A comparison between Britain and the Continent during the Industrial Revolution would be worthwhile, but so far this has not been attempted seriously.

[85] For a survey of contemporary thinking about the "cheap labor is dear labor" issue, see Coats (1958).

the same technology and capital intensity. Clark shows that the high labor cost in the Atlantic economies (always excluding Ireland) was essentially offset by the higher productivity of workers in high-wage countries. Clark concludes that "real labor costs turn out to be as high as those in Britain in most of the other countries except for the very low wage competitors in Asia. The per worker wage rate tells us very little about the true cost of labor" (p. 11).

Labor could vary in its productivity for a variety of reasons. Differences in education seem to have made relatively little difference in productivity, as Mitch's chapter below points out. Another interpretation emphasizes diet: Low-wage workers could not buy enough food, and their malnourishment caused their work to be of low quality. Poorly paid workers could be poorly fed workers. The connection between caloric intake and energy output of workers is well known. Workers on an insufficient diet do not necessarily get sick or die, their entire metabolism simply slows down, to the detriment of their productivity (Scrimshaw, 1983).[86] The dietary model is attractive, because the so-called efficiency-wage model seems quite promising in explaining the failure of premodern, poor societies to develop. Unfortunately, the evidence produced thus far to support this promising idea is ambiguous.[87] Although recent scholarship has concluded that French workers were, in all likelihood, worse fed than British workers (Fogel, 1989, 1991), the same is not true for the Irish, whose potato diets assured them of a plentiful if somewhat monotonous fare (Mokyr, 1983).

Productivity, however, depended on more than nutrition. Adam Smith thought that "the wages of labour are the encouragement of industry, which like every other quality, improves in proportion to the encouragement it receives. A plentiful subsistence increases the bodily strength of the laborer . . . where wages are high, accordingly, we shall always find the workmen more active, diligent, and expeditious, than where they are low." (Smith, [1776] 1976, p. 91). What Smith seems to be describing, however, is an upward sloping supply curve of labor, which makes people work *more* if the wage is higher. The question is, however, what makes people work *better* or harder per unit of time?

Recent thinking about the efficiency-wage hypothesis has shown that labor productivity can depend on the real wage paid to workers in a variety of ways. A simple model of this type is the shirking model, in which it is expensive to monitor the effort the worker puts in. High wages are a mechanism by which the employer extracts more effort from the worker, because a worker caught shirking risks being fired and losing his or her high-paying job. High wages could also increase productivity through reduced turnover. Another model derives a correlation between productivity and wages through an "adverse selection" mechanism: the worst-quality workers agree to work for less (see Akerlof and Yellen, 1986; Weiss, 1990).[88]

Differences in productivity in the early stages of the Industrial Revolution were also likely to arise from differences in workers' attitudes. Concentrating large numbers of workers (of both sexes) in one room and subjecting them to discipline, regularity, and the increasing monotony of the more advanced technique were some of the most difficult problems encountered by early factory masters (Thompson, 1967). Cheap labor was no advantage unless it could be effectively transplanted from the traditional to the modern sector. Sidney Pollard (1965, chap. 5) has pointed to the central paradox of the labor-supply question during the Industrial Revolution: "The lack of employment opportunities . . . existing simultaneously with a labor shortage is in part explained by the fact that the worker was averse to taking up the *type* of employment being offered, and the employer was unwilling to tolerate the habits of work which the men seeking work desired" (p. 196).

How a rural, mostly self-employed labor force was enticed to work in mostly urban mills is one of the most interesting questions in the debate on the Industrial Revolution, and yet it has not received much attention

---

[86] For examples, see Allen (1992b); Freudenberger and Cummins (1976); O Gráda (1992); and Scrimshaw (1983).

[87] The inadequacy of British diets both before and during the Industrial Revolution has been recently documented by Shammas (1990, pp. 134-148). For a dissenting view that maintains that eighteenth-century diets were by and large sufficient, see Riley (1991).

[88] A recent and pioneering attempt to apply this class of models to the Lancashire cotton industry in the first half of the nineteenth century is made by Huberman (1992).

in the literature produced by economists. One answer given, ironically, by the social historian Perkin is purely economic: "By and large, it was the prospect of higher wages which was the most effective means of overcoming the natural dislike for the monotony and quasi-imprisonment of the factory" (Perkin, 1969, p. 130). Pollard (1965) and Thompson (1967) suggest a variety of alternative ways in which the factory owners educated their workers in their own image, trying to imbue them with an ethic that made them more docile and diligent. Punctuality, respect for hierarchy, frugality, and temperance were the qualities that the value system tried to convey onto the younger generation. The factory owners used a combination of approaches; they relied first and foremost on semi-compulsory apprenticed child labor from workhouses ("pauper apprentices") and on women driven out of their cottage industries by the rapid mechanization of spinning. Gradually, they created a more balanced labor force by a combination of higher pay and social control. An example is provided by the research of Huberman (1986; 1991; 1992; 1996). Huberman points out that although in the pre-1800 period the labor market in Lancashire worked in the classical fashion, with flexible wages equating supply and demand, employers soon found that they needed more than a labor force that was available. They needed a labor force that was loyal, reliable, and motivated. To insure this they paid wages that soon became institutionalized as "fair wages" and lost their flexibility. The emergence of such wage rigidity in some industries meant that when demand fluctuated, the adjustments would take place through quantity adjustments: layoffs and short-time became commonplace.

Aside from the question of the productivity of labor, the wages the factory masters had to pay were determined by the other forms of employment open to the workers.[89] The opportunity cost of labor was determined by its productivity in the traditional sector, which still dominated the economy. Before 1850 the modern sector was still relatively small and thus close to a price taker in the labor market. But there was more to the traditional sector than agriculture.[90] At different times the domestic weavers, spinners, nailers, frame knitters, and cutlers, whether they were in the putting-out system or working on their own account, found their economic position threatened as the Industrial Revolution progressed. As the factories gradually expanded, they drove down the price of substitutes and thus the incomes earned by the outworkers and independent artisans in the traditional sector. Slowly at first, but with increasing force, domestic industry was ineluctably transformed by the Industrial Revolution. The modern sector, in a sense, created its own labor force.

Ultimately, then, domestic industry was doomed, but during the long transition its relation with the modern sector was complex (Ogilvie and Cerman, 1996). In many industries, mechanized factory production and manual home production were complementary, and although the type of industrial commodities produced in domestic industry changed substantially, the outwork system showed a remarkable tenacity in its struggle with the factory system. The mechanization of spinning led to a short-lived boom in domestic weaving, and some domestic industries, like tailoring, frame knitting, nail making, and boot and shoe production, remained domestic until well into the second half of the nineteenth century (Bythell, 1978). In the woolen industry, Hudson (1983, pp. 135-136) notes a symbiosis between the company mills and the workshops attached to domestic clothiers' homes. Mass production needed special-purpose machinery that could not itself be mass-produced. Rising incomes maintained an upmarket demand for custom-made, high-quality products, such as handmade clocks, fine linen, and custom-made furniture (Sabel and Zeitlin, 1985). The wage rate in these "sweated trades" was often very low. Since domestic industry was open to anybody, it set the lower bound on the opportunity cost of labor.

To be sure, the wage rate in the modern sector was higher and rose faster than that in the traditional sector. Still, the wages earned were not entirely independent of each other, unless the labor market was subject to extreme segmentation. Thus the growing modern sector produced its own labor force, and although real wages

---

[89] The exact alternative is not clearly defined, which makes the notion of opportunity costs, so beloved by economists, somewhat tricky. By 1815, for instance, emigration has to be considered as a possible factor in setting a floor to the real wage. In Ireland this lower bound was reached by more people than in Britain, and thus Irish migration already became quite substantial before 1850.

[90] This point is still not fully appreciated by many economic historians. Thus O'Brien (1996) interprets differences in long-term economic trends between France and Britain in terms of the ability of the agricultural sector to release labor.

ultimately could not be kept down, the slowness of their rise in spite of rapidly increasing labor productivity has to be seen as part of the interaction of the modern and the traditional sectors.[91] This sheds an important light on the role of cottage industry prior to the Industrial Revolution. The preexistence of cottage industries was neither a necessary nor a sufficient condition for the modernization of industry (Coleman, 1983). But as Jones (1968) and others note, cottage industries catering to distant markets tended to arise in areas where agriculture paid low wages. These were not necessarily areas in which agriculture was backward and poor. In the English Midlands the heavy soils were not suitable to the new husbandry based on mixed farming and stall-fed livestock. This left these regions at a comparative disadvantage in agricultural production, and they increasingly specialized in manufactured goods. In other areas cottage industries emerged because high population to land ratios reduced average farm size. Much of this specialization crystallized, as was argued in a seminal paper by Jones (1968) and as has recently been confirmed by Kussmaul (1990), in the second half of the seventeenth century. This specialization provided the historical background to the supplies of labor that ended up in the factories a century later.

Although the transition from domestic industry to modern industry was at times difficult and varied from region to region, the conclusion that the former was a positive factor in the establishment of the latter has been widely accepted.[92] A number of factors have been proposed as possible explanations of this nexus, including the supply of entrepreneurship by the domestic system, the preexistence of skills, a material culture more directed at the market and technological bottlenecks within the domestic sector that may have led to further innovations through their function as "focusing devices." Some of these, like the flying shuttle, increased the productivity of domestic workers. Others, like the power loom, were feasible only in a factory setting. Here, too, more detailed research is needed. Yet it is clear that the role of domestic industry in supplying a more abundant and elastically supplied labor force, especially from women and children, should become an essential part of this research program.

Above all, it is misleading to view the Industrial Revolution solely as the transition of labor from rural and agricultural occupations to urban and industrial occupations. The critical event was not the creation of an industrial labor force as such but its transformation. In the domestic system workers toiled at their homes, but they were usually only part-time industrial workers, cultivating small plots and hiring themselves out as seasonal wage workers during harvest time. In the modern sector the existence of a large fixed investment implied that part-time operation was uneconomical. The factory worker lost his or her freedom to allocate time between labor and leisure as he or she wished: either the worker wholly submitted to the requirements of the employers and worked the days and hours prescribed by the mill owner or he or she did not work. Although cottage industry in various forms supplied a portion of the labor force needed by the Industrial Revolution (Bythell, 1969, pp. 257-263; Redford, [1926] 1964, p. 41), there were workers, especially in rural areas, who hesitated to make the great leap. Only their sons and daughters realized the hopelessness of the situation and moved (Lyons, 1989; Redford, [1926] 1964, p. 186). Women and children constituted an essential part of the industrial labor force (Berg, 1994).[93] Goldstone (1996) has recently argued that the supply of cheap female labor aged between puberty and their (relatively late) marriage age provided Britain with a strategic advantage in terms of labor supply, especially in cotton mills. Precisely because young women could be paid very low wages (given their low opportunity costs), Goldstone notes, they often turned out to be a cheaper source of labor than adult men. Berg and Hudson (1992, p. 36) also point out that domestic industries released large reserves of cheap and skilled child

---

[91] See Mokyr (1976b) for an algebraic representation of this interaction and some further implications.

[92] For some reflections, see Clarkson (1985, pp. 28-38); Kriedte (1981, especially pp. 152-154); Mokyr (1976b, pp. 377-379). A recent summary is provided by Hudson (1996).

[93] For a recent summary of existing literature on women in the labor force, see Honeyman and Goodman (1991), Bythell (1993), Berg (1994), and Horrell and Humphries (1995a). Recent work on child labor includes Nardinelli (1990), Horrell and Humphries (1995b), and Tuttle (1997).

and female labor leading to high proportions of women and children in the mills. Women fell outside the standard craft apprenticeship system, meaning that few of them could consider themselves skilled artisans, but obviously they were a cheaper source of labor. Moreover, factories needed dexterity, docility, and discipline, and women and children provided these disproportionately before 1850.[94] Children and women's work cushioned the disruptive effects that technological change had on the earnings and employment of married men and allowed the losing economic groups to adjust. Above all, however, youngsters had more malleable skills and personalities and could be conditioned. In 1835 Andrew Ure argued, no doubt with some exaggeration, that "even in the present day...it is found to be nearly impossible to convert persons past the age of puberty, whether drawn from rural or handicraft occupations, into useful factory hands."

What about immigration? In Ireland, where the collapse of domestic industry in the 1830s was swift and brutal, migration of workers to England and Scotland was widespread (Collins, 1981), and these immigrants were an important supplement to the British labor force during the Industrial Revolution (Redford, [1926] 1964, pp. 132-164). As Pollard (1978, p. 113) puts it, "[Irish emigrants] were, in many aspects, the mobile shock troops of the Industrial Revolution, whose role consisted in allowing the key areas to grow without distorting the labor market unduly." Recently, Williamson (1986, 1990a) has questioned the importance of the Irish workers to British industrialization. His calculations assume that the Irish formed an unskilled labor force and that agriculture was more unskilled-labor-intensive than manufacturing. Consequently, he finds that the main impact of Irish immigration was on agricultural output. Although most Irish ended up in rural areas, Williamson points out that their arrival slowed down the migration of British rural workers from the countryside to the cities. It is possible to argue that further disaggregation could overturn this conclusion in some industries. The Irish tended to concentrate in certain sectors and industries, such as mining, construction, and transportation, and in these industries their labor may well have contributed more than Williamson's aggregate computations suggest.[95] On the whole, however, there is little reason to doubt Williamson's conclusion, simply because the number of Irish in Britain, though considerable, was simply not large enough to make a decisive impact on Britain's economy. In 1841 it is estimated that there were 830,000 "effective Irish" in Britain, of whom 415,000 were Irish born and the rest descendants of Irish emigrants. If we assume that all the emigrants and half of the others were in the labor force, the Irish would have added 620,000 workers, which out of a total occupied labor force of about 6.8 million would have amounted to about 9 percent; not a trivial addition, but not large enough to change the parameters dramatically.

Besides the question of the reallocation of labor from the traditional to the modern sector, there are many other loose ends to consider in the area of labor supply during the Industrial Revolution. One question is what happened to participation rates. We know little about these rates for the eighteenth century, and scholars have used population growth rates as a proxy for labor force growth rates. After 1801 the census provides figures for total occupied population that allow us to compute some very approximate participation rates. For what it is worth, the participation rate shows an initial decline from 1801 to 1831 and then rises until 1851 (Deane and Cole, 1969, pp. 8, 143). These changes are small and reflect primarily the changing age structure and measurement error.[96] The concept of a participation rate is in any case something of an anachronism, because it requires a worker to be able to declare himself or herself as either being in the labor force or not. In a society in which a

---

[94] Berg and Hudson also argue that these age and gender differentials influenced innovation and were influenced by it, but persuasive evidence for this interaction is thus far lacking.

[95] Williamson (1990a, p. 160) points out in a long footnote that if regional and industrial labor markets had been highly segmented, the Irish emigrants might have had a larger impact than his estimates imply, because they entered through urban gates and thus at first, at least, would have depressed industrial wages more, thus raising profits and stimulating capital accumulation. Irish immigrants were highly concentrated in a small number of specific urban occupations (Lees, 1979).

[96] Occupied population as a percentage of total population went from 44.86 percent in 1801 to 43.90 percent in 1831 and then rose to 45.28 percent in 1841 and 46.46 percent in 1851.

large if declining percentage of the labor force was economically active in households (farms or workshops), this is not an unequivocal measure even if we had better data. It is thought that the Industrial Revolution mobilized a large part of its labor force by turning part-time workers into full-time workers and transferring workers from "disguised unemployment" to regular work (Pollard, 1978). What is clear is that in many of the more dynamic industries of the Industrial Revolution, including cotton, female and child labor predominated quantitatively in the early stages of the Industrial Revolution. What we do not know with enough precision is how many of these workers were drawn into the labor force altogether and how many were already active either in domestic manufacturing or agriculture. At some point after 1815 this reliance on non-adult-male labor began to decline, and by 1860 it was significantly lower than at the start of the century (Horrell and Humphries, 1995a, 1995b). As the cottage industries had gone into decline, participation rates, however defined, must have been significantly lower already by the middle of the nineteenth century than at their peak during the heyday of the Industrial Revolution.

On the whole, both cottage industries and factories practiced a division of labor between the genders. In the cottage industries women performed mostly low-skill jobs, left most of the skilled work to men, and were excluded from apprenticeship. Arguments that view the division of labor between genders as the outcome of attempts by men to maintain a social status in the family and the community have frequently been made, but hard evidence that would discriminate between this hypothesis and alternative ways of explaining the data is lacking so far. In a few new occupations, such as mule tending, women were excluded. Some technologies may have been especially designed to use female labor, and the evidence from the Birmingham toy trade suggests that women could even operate relatively heavy machinery (Berg, 1994, pp. 144- 156).

Changes in the amount of labor performed per worker were possibly of greater importance to the labor supply than changes in participation rates. It is also a variable for which aggregate information is the hardest to come by. Labor input per worker could increase by lengthening the laboring day and the number of days worked and by reducing involuntary unemployment. Did workers in 1830 work more than in 1760? This view is certainly part of the conventional wisdom. Pollard (1978, p. 162) has no doubt that this is the main explanation for the rise of family income before 1850. Jones (1974, pp. 116-117) and Freudenberger (1974, pp. 307-320) are equally certain that workers toiled longer hours during the Industrial Revolution.

De Vries's idea of an "industrious revolution," presented earlier in this chapter, also implies an increase in labor input per worker and less leisure. This account sounds plausible enough, but can it be sustained by evidence? Unfortunately, we do not know with any precision how many hours were worked in Britain before the Industrial Revolution in either agricultural or nonagricultural occupations. Some progress has been made in recent years, however. In a recent important paper, Voth (1998) has tapped a new source of data to examine what happened to the length of the work-year in England in the latter part of the eighteenth century. His finding from analyzing the Old Bailey court records is that a considerable increase occurred in the number of hours worked in London between 1760 and 1800, perhaps as much as a third. The reason for this increased diligence is in part a decline in real wages, which spurred workers to work harder to maintain their living standards and buy the new consumer goods that were appearing on the market at this time. If this finding is confirmed, an increase in labor hours per worker would explain what little growth of output per capita there was. Yet London was not the same as England, and in a paper based on two types of rural workers (threshers and sawyers), Clark and Van Der Werff (1997) conclude that work effort had changed little in England since the late Middle Ages. The sources and methodologies used by these two papers are quite different, and the findings may not be totally irreconcilable. As of now, however, they clearly depict very different pictures of changes in the rate of partipication in formal labor markets. In the cottage industries the distinctions between work, leisure, and social life were not as sharply drawn as in our own time. Most accounts maintain that workers started the week slowly, then picked up steam as the weekend approached, often working very long days toward the end of the week (Hopkins, 1982, p. 61;

Thompson, 1967, p. 50). The decline of "St. Monday" (Reid, 1976) could therefore have been less of a net increase in the working week than a rearrangement to distribute the effort more evenly.[97]

McKendrick (1974, p. 163) derides the idea that longer hours explain higher incomes, labeling it a "prelapsarian myth of the golden past," and asserts that premodern labor was "grinding toil," as bad as factory labor but less remunerative. It is indeed easy to document many cases of long and hard hours in cottage industries; days of fourteen to sixteen hours were common (Rule, 1983, pp. 57-61). It is not clear, however, how common such long days were and to what extent they did not make up for the customary long weekend or for usually low wage rates. Much of our information here comes from nineteenth-century sources, which may be biased because economic conditions were deteriorating for cottage industry. If labor supply curves were downward sloping or backward bending, as is widely believed, the declining wage rates in domestic industry in the nineteenth century led to longer working days. Still, the idyllic picture drawn by some (Medick, 1981; Thompson, 1967) of working conditions in domestic industry in the eighteenth century is probably unrepresentative of premodern labor conditions. The most recent attempt to answer the question is provided by Gregory Clark (1994), who concludes from the fact that weekly earnings rose faster than piece rates that workers in factories indeed worked longer.

One reason the comparison of factory and domestic work may yield misleading conclusions is that the representative industry discussed for the nineteenth century is often the textile industry, and especially cotton spinning. The laboring days of workers in the cotton mills before the mid-1840s were long, even by the standards of the time. The labor day was extended by as much as two hours and the number of working days per week was set at six, resulting in working weeks of seventy-six hours, compared to about sixty hours in most other industries. Official holidays were few, and unofficial leaves had to be made up with overtime (Bienefeld, 1972, pp. 30, 49). In mines, too, labor hours were increased during the Industrial Revolution. These extensions were, however, far from universal. A study of Birmingham and the Black Country has found no evidence of longer working hours, and the traditional workday of twelve hours including meals remained the most common practice (Hopkins, 1982). Only a small proportion of the labor force was actually employed in satanic mills or mines by 1840; most British workers were still employed in agriculture, domestic service, construction, and small workshops where work habits changed little.

Another possible source of labor was the reduction of involuntary unemployment. On the one hand, the amplitude of business fluctuations gradually increased after 1760, and as slumps became more severe, short-time and layoffs became more common. On the other hand, improved transportation and communication allowed a more efficient organization of the economy, thus reducing the problem of seasonal unemployment. The notion of large reserves of unemployed workers awaiting a rise in labor demand is much in dispute, although O'Brien and Engerman (1991) and others rely on contemporary opinion that Keynesian unemployment was a serious problem in the eighteenth century.[98] The evidence, however, is not wholly persuasive (Blaug, 1968, p. 15). Of similar interest is the question to what extent modernization reduced the multitudes of unemployables: vagrants, beggars, prostitutes, and other persons on the fringes of society. A glance at Henry Mayhew's description of London in the late 1840s suffices to warn us that the Industrial Revolution did not eliminate these people and possibly caused an increase in their proportion of the British population during the period.

b. Capital

---

[97] Voth's data (1998) show that in the 1750s Monday was a typical day off, not much different from Sunday, whereas by 1800 it had become a day that was not statistically different from other weekdays. This finding suggests that the decline of St. Monday was firmly located in the second half of the eighteenth century.

[98] Keynes himself, in a famous statement, expressed the view that the writings of the mercantilists suggest that "there has been a chronic tendency throughout human history for the propensity to save to be stronger than the propensity to invest. The weakness for the inducement to invest has been at all times the key to the economic problem" (Keynes, 1936, p. 108).

The role of capital is not less controversial than that of labor. Recent work has concentrated on three issues. The first is the question of how capital markets worked during the Industrial Revolution and what effect they had on the process of technological change and accumulation. The other two issues have been raised primarily by economists, namely, the speed at which capital accumulated and the changes in its composition (circulating vs. fixed). On the issue of how capital markets worked, Larry Neal (1990) has recently pointed out that in the eighteenth century there was in fact an international capital market that funneled funds between different countries and that was clearly integrated, except when disrupted by war. There are also signs that British internal markets improved their operations during the Industrial Revolution: Buchinsky and Polak (1993) find that after 1770 there was a growing correlation between London interest rates and Yorkshire property transactions, though they find no sign of integration before that. Hoppit (1986) has reached a similar conclusion based on bankruptcy statistics. Although it would be premature to speak of a well-integrated capital market by 1800, clearly the capital market was becoming larger, more efficient, and more "modern" during the years of the Industrial Revolution.

Whereas the role of capital markets in the British economy as a whole is indisputable, their importance to the Industrial Revolution, properly speaking, is more difficult to assess. The biggest borrowers in Europe in this period were governments that needed to finance deficits. The demand for credit also came from merchants with bills to be discounted, entrepreneurs active in canal and road construction, landowners in need of funds for the purpose of enclosure and other improvements, and construction interests. There was some inevitable overlap between these borrowers and what we would consider the "modern sector," but it was relatively small. Moreover, the smallness of the modern sector relative to the entire British economy meant that its demand for loanable funds did not loom large relative to the needs of the economy. Dealing with the supply of savings on an aggregate level, however, is even more misleading than an aggregate analysis of labor markets. Such an analysis assumes the existence of *a* capital market that allocated funds to all competing users, presumably on the basis of an expected rate of return and riskiness. Certain developments, especially the growth of transport networks, would have been slowed down considerably, and possibly aborted, had it not been for capital markets. As far as the manufacturing sector is concerned, however, matters are quite complex.

How did the Financial Revolution, which preceded the Industrial Revolution, affect it? The standard view of the interaction of "the two revolutions" has been that they had very little to do with each other. Postan argued in 1935 that "within industry almost every enterprise was restricted to its own supplies. The Industrial Revolution got under way while capital was not yet capable of moving between 'alternative employments'" (1935, p. 74). This view is now recognized as too simple: Financial markets were far more complex and subtle and their impact more pervasive than the earlier writers assumed. Yet there is little evidence that these financial markets were instrumental in helping modern industry more than vice versa.

Regarding the supply of capital, the most thorough work has been carried out by François Crouzet (1965, 1972, 1985b), complemented for the later period by Cottrell (1980). This work demonstrates that the capital needs of the modern sector during the Industrial Revolution were met from three sources. First were the internal sources in which the investor borrowed, so to speak, from himself using his private wealth (or that of his family) for start-up and plowing his profits back into the firm. Second, there were informal, or "personal," capital markets in which borrowers turned to friends, relatives, or partners for funds. Third, there was the formal capital market in which the borrower and the lender did not meet and in which attorneys, brokers, and eventually financial institutions (banks, insurance companies, stock markets) fulfilled their classic functions of intermediating between lenders and borrowers, concentrating information, and diversifying portfolios. The questions we must ask are, how important were these three forms of finance in the Industrial Revolution? and how can we explain this complex and seemingly inefficient mechanism? Students of the Industrial Revolution agree that most industrial fixed capital originated from internal finance. Crouzet (1965) concludes that "the capital which made possible the creation of large scale 'factory' industries came . . . mainly from industry itself . . . the simple answer to this question how industrial expansion was financed is the overwhelming predominance of self-finance" (pp.

172, 188). In a later paper he qualified this conclusion somewhat but insisted that it remained "broadly valid" (Crouzet, 1972, p. 44; Crouzet, 1985b, pp. 147-148).[99]

In the early stages of the Industrial Revolution, the fixed-cost requirements to set up a minimum-sized firm were modest and could be financed from profits accumulated at the artisan level (Crouzet, 1965, p. 165; Pollard, 1964). Plow-back then provided a regular, almost automatic mechanism by which profits augmented the capital stock. As technology became more sophisticated after 1830, the initial capital outlays increased, and it became increasingly difficult to rely on internal finance to start a business. For railroads this was of course out of the question. For existing industrial firms, retained profits usually remained central to the accumulation of capital. Even in a world in which firms relied exclusively on retained earnings, an intersectoral capital market could function. Individuals who made their fortunes in commerce, real estate, or the slave trade could use these funds to diversify into manufacturing. There were examples of merchant princes entering modern manufacturing, such as the case of Kirkman Finlay, an overseas merchant who entered cotton spinning between 1798 and 1806, and the Wilson brothers who established the Wilsontown ironworks. On the whole, however, these cases were exceptional (Crouzet, 1985a, pp. 99-100).

The second source of funds, the informal capital market, can easily be illustrated with examples, but it is not known how important this form of finance was relative to other sources. Postan (1935) argues that capital was still a very personal thing, which most people wanted to keep under control. If one lent it out, it was only to an intimate acquaintance or to the government. Even partnerships, which were frequently resorted to in order to raise capital while avoiding the costly process of forming a joint-stock company, were usually closely tied to family firms. The taking in of strangers as sleeping partners merely for the sake of getting access to their wealth was relatively rare at first (Heaton, 1937, p. 89). This caution slowly dissipated during the Industrial Revolution, but active partners often bought out the others, and the advantages of partnership were as much in the division of labor as in the opportunity to raise credit. Many of the most famous characters in the Industrial Revolution had to resort to personal connections to mobilize funds. Richard Arkwright got his first loan from a politician friend, and James Watt borrowed funds from, among others, his friend and mentor, Dr. Joseph Black. Although the phenomenon was thus widespread (Crouzet, l965, p. 184; Mathias, 1969, pp. 150, 162-163), personal loans are as much of interest as a symptom of how the system operated as for the fact that they were a major channel through which funds were mobilized. Crouzet points out how exclusive and selective these personalized credit markets were: To have access to these informal networks one needed to be a member of them and be "known and well thought of in the local community" (1985a, p. 96). The market for capital can thus be seen once again to have depended on the market for information.

As the modern sector grew, intrasectoral flows of funds between firms became more important, especially flows occurring within the same industry. Insofar as these mechanisms only reallocated funds among different industries in the modern sector, the upper bound that the rate of profit imposed on the rate of growth of that sector did not disappear. Instead of constraining the individual firm, the supply of funds now constrained the modern sector as a whole. Although there were important exceptions, by and large the modern sector pulled itself up by its own bootstraps.

The third mechanism for obtaining capital, the formal credit market, operated primarily through merchants, wholesalers, and country banks.[100] The consensus on the role of the banks is that, with some exceptions, they rarely figured in the financing of long-term investment. Their importance was mainly in

---

[99] For similar statements, see, for example, Mathias (1969, p. 149) and Cameron (1967, p. 39). Cameron goes so far as to assert that "the rate of growth of capital is therefore a general guide to the rate of profit," though he concedes that alternative investment opportunities for the factory master could upset that correlation.

[100] Joint-stock companies were exceedingly rare, in part due to the Bubble Act that mandated they could only be incorporated through Parliament, but also because promoters defrauded their stockholders, and managers usually mismanaged the companies (Pollard, [1965] 1968, p. 25; Robb, 1992). There were a few exceptions, such as the British Cast Plate Glass Co., established in 1773, which imitated the French Royal St. Gobain manufactory, although it remained a private company.

satisfying the need for working capital, primarily by discounting short-term bills and providing overdrafts (Flinn, 1966, p. 53; Pressnell, 1956, p. 326). Pollard has made a case for the reexamination of the importance of the banks on these grounds. Given that banks provided much short-term credit, firms short of capital could use all their internal funds on fixed investment (Crouzet, 1965, p. 193; Pollard, 1964, p. 155). Pollard, however, assumes that fixed capital grew at a rate much lower than implied by Feinstein's figures. His own earlier estimates imply a rate of growth of fixed capital of 2.4 percent per annum between 1770 and 1815, whereas Feinstein's fixed capital estimates grew at 4.2 percent per annum in the same period (Feinstein, 1978, p. 74). In manufacturing and trade the discrepancy is larger; according to Feinstein, gross fixed capital formation grew between 1770 and 1815 at 6 percent per annum, as opposed to Pollard's 3.4 percent (Feinstein, 1978, p. 74; see also Feinstein and Pollard, 1988). Thus financial constraints on capital accumulation may have been more stringent than Pollard originally presumed because he underestimated the needs. Moreover, substituting fixed capital for circulating capital may have been less simple than he thought, because as industrial output increased, the demand for circulating capital grew as well. Feinstein shows that between 176O and 1830 fixed capital in industry and commerce increased from 5 percent of domestic reproducible capital to 18 percent, whereas circulating capital in industry and commerce increased from 6 percent to 7 percent in the same period. Was the activity of banks enough to finance an increase of 164 percent in working capital over seventy years? Cottrell (1980, p. 33) concludes cautiously that there are indications that industrial growth before 1870 may have been blunted by shortages of circulating capital. Honeyman (1983, pp. 167-168) maintains that small businessmen found banks unreliable, and that even for circulating capital, kinship and friendship groups were preferred. The difficulty in obtaining funds led to the selective weeding out of the industry of entrepreneurs of humble origins who did not have access to these informal sources of funding and thus failed to survive crises during which working capital was hard to obtain. From a different point of view, Cottrell speculates that short-lived firms had better access to formal capital markets than firms that survived. The sharp fluctuations in the financial sector dragged into bankruptcy many industrial firms, and this effect may result in an underestimate of the importance of the plow-back of profit as a source of investment, because the firms that left records would tend to be *less* dependent on external finance (Cottrell, 1980, pp. 35, 253-255). Yet it remains to be seen whether enough evidence can be produced to jeopardize the widely held belief in the predominance of internal financing in this period.

Thus capital scarcity and biases in the capital markets slowed down the rate of accumulation and the speed of industrial growth. The reliance on plowed-back profits for investment clearly meant a slower growth rate compared to a world in which borrowers could access savings regardless of its source. In spite of these qualifications, it is still true that if credit markets had not existed at all, the accumulation of fixed capital would have been somewhat slower, though the rehabilitation of the banking system does not go far enough to allot it a truly strategic role in the Industrial Revolution.[101]

To what extent can economic theory explain the picture of the plow-back of retained profits and self-finance? The limited willingness of commercial banks to finance long-run projects is understandable. Banks needed their assets in liquid form to be able to pay depositors on demand since there was no lender of last resort.[102] This constraint was a result of the nature of commercial banks. Investment banks and other forms of financial intermediaries did not have to maintain liquid portfolios. Why such institutions were relatively unimportant in Britain (compared to Continental countries) is still an unanswered question. Yet the reliance on internal finance during the Industrial Revolution is not surprising. Firms tend to prefer internal over external

---

[101] See Cameron (1967) and Crouzet (1972). It is possible that further work on the asset composition of British banks may revise this conclusion for the period after 1844, which might explain Good's (1973) finding that the ratio of banking assets to GNP was relatively high in Britain compared to later industrializers (see also Collins, 1983).

[102] The necessity for banks to preserve liquidity was made into a virtue by the so-called real-bills doctrine, which stipulated that if banks confined themselves to short-term, self-liquidating loans (such as discounting commercial bills), the price level would remain stable. Regardless of whether there was any merit in this theory in the short run, in the long run it confined commercial banks to supplying, almost exclusively, circulating capital.

finance, even though economic theory suggests that the reliance on retained profits is inefficient. In the post-World War II United States, too, firms have obtained over 70 percent of their finance from internal funds (MacKie-Mason, 1990).[103] The use of internal funds during the Industrial Revolution is thus not a historical anomaly.

Economic theory has in recent years provided substantial insights into the reason for the persistent "imperfection" of capital markets. In earlier theoretical work such as Hicks (1946) and Scitovsky (1971), firms were perceived to face upward-sloping supply curves of loanable funds, which would be consistent with internal financing. These models were not pursued, however, and their microeconomic foundations were never quite made clear. More recently, though, with developments in the economics of information, our understanding of the economic processes involved has improved.

For instance, Mayshar (1983b) argues that it is not risk per se that causes real-world capital markets to deviate from the theoretical constructs but divergences of opinions among potential lenders with respect to the rate of return. Such divergences would of course gradually disappear in a stationary world in which no new information was created. But in a world of rapid technological change, shifting demand patterns, and a changing political environment, divergences were not only possible but in fact inevitable. Thus rapidly changing conditions during the Industrial Revolution effectively precluded the efficient operation of capital markets. Mayshar pictures savers as forming concentric circles around the entrepreneur, with his own funds in the center and next those of the people closest to him (friends and relatives), who were the sources most likely to lend to him. The farther one gets from the center, the more the expectations tend to diverge from the entrepreneur and the higher the rate of interest that he has to pay. Similarly, Stiglitz and Weiss (1981) show how the informational asymmetry between lender and borrower can lead to adverse selection in which a rise in the interest rate causes the borrowers with the safest projects to drop out of the market. This means that interest rates will generally not clear the credit market and credit rationing may be quite general. Under credit rationing, many entrepreneurs found themselves rationed out of the market and hence had no choice but to rely on self-finance. Whether potential borrowers preferred to rely on their own resources or whether they were rationed out of the credit market, the experience of the capital market during the Industrial Revolution clearly shows the applicability of these models.

The assumption of asymmetric information seems especially apposite. Because much of the technology was new, the information gap between entrepreneur and saver or banker was even greater than in our own time. A banker in 1790 would have much less information about the economic potential of a mule or a modern calico printer than he would about the quality of an investment in, say, a flour mill or a fence around enclosed land. Many firms, as well as their technologies, were new and had no reputations of creditworthiness. Young, growing firms tend to be the most severely credit-rationed. Consequently, some of them ended up establishing their own banks (Crouzet, 1985a, p. 19).

On the questions of the size and composition of the capital stock, our knowledge has been increased by Feinstein (1978) and Feinstein and Pollard (1988), who, with their collaborators, have created a data base to investigate the quantitative aspects of capital formation in this period. Feinstein's data permit us to test two hypotheses that have dominated the literature on capital in the Industrial Revolution. One hypothesis is the Lewis-Rostow claim that the investment ratio doubled during the Industrial Revolution. The other is the Hicks--Ranis-Fei view that the truly fundamental change was the shift from predominantly circulating to fixed capital. Both hypotheses have been criticized vigorously, and we are now in a position to assess these criticisms.[104] Feinstein's data imply that the dismissal of Rostow's hypothesis was premature. The ratio of total gross investment as a proportion of GDP rose from 8 percent in 1761-1770 to 14 percent in 1791-1800, and after a temporary setback in 1801-1811 returned to 14 percent for the half-century after 1811 (Feinstein, 1978, p. 91).

[103] Calomiris and Hubbard (1991), studying the years 1936-1937 found that firms were in fact willing to pay a substantial tax on their invested retained earning rather than go outside for funding.

[104] The Rostow hypothesis was criticized, among others, by Habakkuk and Deane (1962). For a critique of the importance of fixed capital in the Industrial Revolution, see Pollard (1964).

More recently, Crafts (1983a) has revised Feinstein's estimates, criticizing in particular the price deflators that Feinstein used. Crafts's figures still show a doubling of the investment ratio from 5.7 percent in 1760 to 11.7 percent in 1830; this reproduces the Lewis-Rostow prediction of its doubling with dead accuracy, though somewhat more gradually than Rostow thought, which is hardly surprising in view of the highly aggregative nature of this ratio.

As to the Hicks-Ranis-Fei hypothesis, fixed capital rose from 30 percent of national wealth to 50 percent between 1760 and 1860, while the corresponding ratio of circulating capital declined mildly from 11 percent to below 10 percent. In industry and commerce the ratio of total circulating to total fixed capital fell from 1.2 in 1760 to .39 in 1830 and .30 in 1860 (Feinstein, 1978, p. 88). The absolute amount of circulating capital increased as well during the Industrial Revolution, but its growth was dwarfed by the rise in fixed capital. In this sense, then, the Hicks-Ranis-Fei view is corroborated. The economic reasons for the change in the composition of capital are rather obvious. Improved transportation, communications, and distribution reduced the need to hold large inventories of raw materials, fuel, and finished products. There are well-understood economies of scale in the holding of inventories and cash, so that it is clear that larger firms needed less circulating capital per unit of output than domestic industry. This may have been partially offset by the requirements of new inputs, such as fuel and spare parts. A second factor in the relative decline of circulating capital is the decline of output prices due to productivity growth, which reduced the value of goods in progress and raw materials relative to that of buildings and equipment.

The importance of capital in the Industrial Revolution was not identical to the importance of the Industrial Revolution in capital formation. In current prices, in the early days of the Industrial Revolution (1761-1770), manufacturing and mining accounted for only 12.5 percent of gross domestic fixed capital formation. Although the annual investment in industry increased almost 15-fold between 1760 and 1830, the share of mining and manufacturing in 1831-1840 was only 21.1 percent (Feinstein, 1988a, p. 429). Yet without capital the modern sector would not have been able to grow. The unit setup costs of firms was rising steadily, and the number of firms, as the industry expanded, was growing rapidly. Consequently, in iron, cotton, steam, and transport gross capital formation increased by huge factors, and the stock of capital mushroomed to unprecedented levels. In mining, for example, gross capital formation in 1830 was 15.6 higher than in 1760 (Pollard, 1988, p. 63). In cotton the *stock* of capital in 1788 was only 12 percent of its level in 1833 (Chapman and Butt, 1988, pp. 124-125). All the same, fixed capital in cotton in 1833 was only 1.5 percent of the national stock of reproducible fixed assets. The smallness of the share of the modern industries in the economy is in and of itself not sufficient to show, however, that they were not capital constrained.

Oddly enough, the total factor productivity estimates seemingly imply that capital formation was a comparatively minor factor in the macroeconomics of the Industrial Revolution. The most recent figures produced by Crafts and Harley (1992) suggest that capital accounted for about half of the aggregate growth of the economy between 1760 and 1830. Because capital grew at about the same rate as output and only slightly faster than labor, however, it contributed little to growth proper. In the period 1760-1800 the rise in the capital to labor ratio and in total productivity each accounted for half the rise in per capita income; after 1800 the contributions fall to 30 percent for capital and 70 percent for productivity. Feinstein, who was the first to notice this, rejects this interpretation and points to the importance of capital as the "carrier" of technical progress. Insofar as capital and technological progress were complementary, the arithmetic of total factor productivity estimates are misleading, since these computations assume that the contributions of capital and productivity are additive and independent. A more accurate estimation would try to take into account the interaction between the two.

How important to the course of the Industrial Revolution were the failings of the capital market? Crouzet has concluded that "the eighteenth century capital market seems, to twentieth century eyes, badly organized, but the creators of modern industry do not seem to have suffered too much from its imperfection. . . . English industry, compared with that of the Continent, seems to have overflowed with capital" (Crouzet, 1965, pp. 187-188). This conclusion may be ripe for some reexamination. First, while the comparison with the Continent is probably accurate on the whole, there were important exceptions (Mokyr, 1975). On the Continent, too, self-finance was the norm, and it is not quite clear whether Britain was much better supplied with capital than, say, Belgium. Moreover, it seems inescapable that the Industrial Revolution in Britain would have occurred faster and more efficiently if financial constraints had been less stringent. Given that the modern sector as a whole was at first rather small compared with the rest of the economy, the capital market's imperfection meant that from the outset the rate of profit set a ceiling on the rate of accumulation. The existence of *some* capital markets does not necessarily refute this argument. If these markets channeled savings from one firm to another in the modern sector, the constraint on the sector as a whole remained in force, and fixed capital had to grow by pulling itself up by the bootstraps. Postan put it well in his classic article: "By the beginning of the eighteenth century there were enough rich people in the country to finance an economic effort far in excess of the modest activities of the leaders of the Industrial Revolution. . . . What was inadequate was not the quantity of stored-up wealth but its behavior. The reservoirs of savings were full enough, but conduits to connect them to the wheels of industry were few and meager" (Postan, 1935, p.71).[105]

5. The Factory and the Modern Industrial Firm

The creation of the workplace, in which many workers were assembled together under one roof to jointly produce an output and were subject to discipline and coordination, has become one of the symbols of the Industrial Revolution. To some extent this is a myth: Some large factories did exist before 1750. The great silk mills in Derby and Stockport, the ironworks of Ambrose Crowley in Newcastle, and metalworks of John Taylor and Matthew Boulton employed many hundreds of workers before 1770. Yet such large plants were rare. Large capitalist enterprises were far more common, but they typically left most of the work to be carried out in workers' homes, and only a few stages of the product were completed in centralized sites. In wool, for example, a large employer like Samuel Hill in Yorkshire in the 1740s employed 1,500 workers, mostly in putting-out.

Part of the story of the Industrial Revolution is that these employees were brought to work in centralized plants, thus changing the nature of work and with it the basic functioning of the family and the household. Increasingly, households became specialized units designed for consumption only, whereas production was carried out in a firm, geographically divorced from the home and often subject to different rules and hierarchies. Why did this happen? Some economists, such as Oliver Williamson (1980), declare that by saving on transactions costs, factories were simply more efficient than cottage industries (whether putting-out or independent producers), and thus their rise was inexorable. Such a simplistic approach cannot possibly do justice to the historical reality (S.R.H. Jones, 1982; Szostak, 1989). After all, the domestic system survived for many centuries, and its demise was drawn out over a very long period. Its advantages were many: It kept families geographically intact, it was flexible and more adaptable to fluctuations in demand and supply, and it left the workers free to choose any point on the leisure-income trade-off rather than forcing them into rigid work schedules and the discipline of the factories. Geographical centralization of production under one roof and the imposition of factory discipline did not always go hand in hand and need to be explained separately. If the Industrial Revolution, as it did, replaced a predominant domestic organization of manufacturing by one that was largely concentrated in specialized workplaces away from homes, it stands to reason that something changed in the economy that accentuated the advantages of centralized work places relative to the advantages of domestic production.

The most obvious candidate for the cause of such a shift is that the new technologies changed the optimal scale of the producing unit and introduced increasing returns where once there were constant returns. Some equipment could not be made in small models that fit into the living rooms of workers' cottages and thus required large plants: iron puddling furnaces and rollers, steam and water engines, silk-throwing mills, chemical

---

[105] Crouzet's statement that the early factory masters "did not suffer" seems oddly incompatible with his own evidence. Two paragraphs below this statement, Crouzet cites the cases of two highly successful firms, the Walker brothers and McConnel and Kennedy, who paid themselves miserably low salaries in order to maximize the income available for plowing back (Crouzet, 1965, pp. 188-189). Some of the most famous inventors and entrepreneurs (Cartwright and John Roebuck immediately come to mind) foundered for lack of working capital, and Richard Arkwright's success is often attributed not to his technical skills but to his virtuoso ability to remain afloat in the treacherous currents of finance during the early stages of the Industrial Revolution.

and gas works -- all required relatively large production units. Heating, lighting, power supply, security, equipment maintenance, storage facilities, finance, and marketing were all activities in which scale economies were obviously the result of technical considerations. Long ago Usher wrote that "machinery made the factory a successful and general form of organization. . . . Its introduction ultimately forced the workman to accept the discipline of the factory" (Usher, 1920, p. 350). Landes (1986, p. 606) has recently restated this argument in unambiguous terms: "What made the factory successful in Britain was not the wish but the muscle: the machine and the engines. We do not have factories until these were available." Both would agree, of course, that factories without machinery were not only possible but actually existed; in the long run, however, their success depended on technology. Maxine Berg, who has argued forcefully for the viability of small-scale production until the 1830s and beyond, concludes that the transition to the factory system "proceeded at a much faster pace where it was combined with rapid power-using technological innovation" (1994, p. 207).

Others have rejected this position: Stephen Marglin (1974) set the tone, which was echoed by others as diverse as Berg (1980), Cohen (1981), and Szostak (1989, 1991). Their argument is that technological change was not necessary for the establishment of centralized workshops, which in fact preceded the great inventions of the last third of the eighteenth century. Berg (1994, p. 196-97), Hudson (1992, p. 28), and Szostak (1989, p. 345) point to industry after industry that established centralized workshops employing practically the same techniques as cottage industries: wool, pottery, metal trades, even handloom weaving and framework knitting. Rosenberg and Birdzell (1986, p. 186) feel that "more and more control would have devolved upon the factory master" had the steam engine and semi-automatic machinery never been invented. Marglin's own view is little more than a Marxist tale of woe according to which factories enabled employers to exercise more control over their workers and to squeeze more profits out of them. Technological progress in this interpretation tended to be a by-product of the intensification of social control. The fact that in many industries workshops preceded the emergence of new technologies does not prove, of course, that technological factors were unimportant in the development of the factory, only that they were not the *only* factors. The large workshop's occurrence may have preceded mechanization in many industries, but surely its ultimate triumph was a result of the growing advantage that new technologies bestowed on factories. Marglin's argument is further undermined by the fact that from the point of view of employer control, the distinction between factory and domestic workers is not as sharp as is usually supposed. Many of those workshops were not factories in the traditional sense of the word -- they imposed no discipline, observed no tight schedules or regulations, and paid workers by the piece. The employer hardly cared if the worker worked hard or not, if he or she arrived at work on time, took Mondays off, or drank on the job. These workshops were purely "rent and charges" kinds of places and thus were quite different from Marglin's oppressive and tightly controlled mills (Clark, 1994). On the other side of the equation, social control gradually invaded the domestic economy during the years of the Industrial Revolution. A series of acts passed between 1777 and 1790 permitted employers to enter the workers' premises to inspect their operations, ostensibly to curb embezzlement. Unwin (1924, p. 35) concludes that by this time "there was not much left of the independence of the small master, except the choice of hours."

All the same, while technology clearly played a role, it cannot account for the entire phenomenon. What needs to be explained is not why factories were superior to domestic industry, because they clearly were not under all circumstances and not in all products and processes. Manufacturing was sufficiently diverse and variable to let the degree of plant-level economies of scale vary all over the map, both over time and over a cross section at a specific moment. A more cogent complement to the technological determinism of Landes is provided by economists such as Millward (1981), Szostak (1991), Clark (1994), and Langlois (1995). Some of this reasoning derives from the economics of information. The organization of production by wage labor under any system depends on information that the employer can amass on the effort the worker puts in. Paying workers a piece rate -- uniformly practiced in putting-out industries -- solves this problem if the employer has no difficulty assessing the quality of the final product and if there are no cross effects between workers' productivities (so that the effort of one worker does not affect the output of another). In the domestic system, employers faced a double problem: Workers could increase their earnings by cutting corners on quality and finish, and the embezzlement of raw materials (which usually belonged to the capitalist) was a widespread complaint (Styles, 1983). The problem of embezzlement, like quality control, was one of information costs; measuring the precise quantities of yarn supplied to a weaver and comparing those with the final output was itself costly, and had to be correlated against normal losses of raw material during the process of production, which the employer did not observe directly. As the division of labor became tighter and the final products more complex, the decentralized division of labor practiced in the putting-out system became increasingly costly. Factories, too, usually paid piece rates, but the monitoring of quality was much easier because the employer could inspect the inputs and the production process as well as the output. Factories also reduced embezzlement and capital costs incurred by workers' negligence. In addition, in factories there was the option of paying workers a time rate, which would be necessary if the marginal product of labor was hard to assess or beyond the worker's control. The problem is, however, that embezzlement and quality control were age-old problems and it is far from clear what changed around 1750 to tip the balance of advantages gradually in the favor of factories.

Moreover, economists have increasingly realized that all systems in which one individual works for another -- that is, all capitalist systems -- are subject to an agency problem: The employer (or "principal") has to manipulate the incentive system to ensure that his worker ("agent") operates so as to maximize the profits of the enterprise. Factories solved the agency problem by imposing direct monitoring of labor by supervisory personnel overseeing the efforts put in by the workers, their use of raw materials, and the care with which they carried out their tasks. Thus the main advantage of factories in this view was that it permitted the employer to ascertain whether fluctuations in output were due to the worker's effort or to a circumstance beyond his or her control. The incentives set up by the factory system to solve the agency problem were largely negative: A negligent or dishonest worker could be fined, dismissed, or even punished physically (Pollard, [1965] 1968, p. 222).

Some specific examples of this general problem have been proposed as explanations for the rise of the factory system. Szostak (1989, 1991) argues, for example, that the employer used centralized workshops to produce standardized goods of more uniform quality, because more integrated markets and changes in distribution methods in the eighteenth century required these changes (see also Styles, 1992).Szostak links the rise of standardization to the growing integration of the British market for manufactures, which he attributes to improvements in transportation. Standardization and uniformity demanded a special kind of quality control, which required continuous supervision and thus factories. Alternatively, as new technology was embodied in more sophisticated and expensive capital goods, the employers became more concerned with the workers' treatment of these machines, because negligence and sabotage became increasingly costly to the firm. Factories may also have induced innovation directly. Some writers, beginning with Adam Smith, strongly believe that a finer division of labor leads to mechanization because the division of labor splits production up into simpler parts, and simple processes are easier to mechanize. Moreover, in the domestic system the entrepreneur rarely observed actual physical production as it was occurring. Once he actually observed the interaction of his labor, his equipment, and his materials, as happens in centralized workplaces, he was more likely to come up with ideas how to save all three than the absentee putting-out merchant manufacturer. A variation on this theme is provided by Langlois (1995) who argues that an increase in demand led to an increase in volume, which made standardization possible. Langlois notes that a sufficiently large number of products makes it worthwhile to invest into a standardized way to mass-produce using jigs or dies which represent fixed costs. With higher fixed costs, argues Langlois, workers' shirking becomes more costly and it becomes more imperative for the firm to monitor the workers' efforts and thus process supervision eventually becomes economical.[106] While ingenious, this explanation might be viewed to depend on an autonomous increase in the demand for manufactures after 1750 as a deus ex machina, which, as we have seen, is not devoid of difficulties. Perhaps a combination of growing population and improvement in transportation à la Szostak will provide some of the primum movens here. It is more likely, however, that such an increase is indeed not independent of changes in technology. After all, major changes in technology in one industry will be perceived as shifts *of* the demand curves in a complementary

---

[106] As Langlois realizes, his theory only holds if the worker's marginal product is costly to measure; as long as marginal products are cheap to measure, fixed costs will not lead to factory discipline. Properly speaking, therefore, Langlois's theory belongs to the information-based theories of factory work.

industry: a sharp decline in the price of cotton yarn resulting from technological change in spinning will lead to an increase in the demand for weaving on one side and carding on the other.

A third explanation of the rise of the factory has to do with the division of labor and is logically independent of the technological and informational interpretations (though in reality the three were closely intertwined). Dividing labor into small tasks carried out by specialists has two advantages. The first, stressed by Adam Smith, assumes that all workers are the same at first but that the division of labor enhances productivity because specialized workers get better at what they do through learning and experience, because time is saved in moving work between workers rather than workers between different tasks, because of the simplification of tasks allowing more routinization, and because of the putative effect that the division of labor has on invention. Routine and repetitive work tends to be less skill-intensive, cheaper, and possibly more productive. The second advantage, emphasized by Babbage, assumes that workers differ inherently in their abilities and that the division of labor maximizes productive efficiency because workers can specialize in those tasks in which they have a comparative advantage. Specialization assures that workers are not asked to carry out tasks for which they are overqualified (which would be wasteful) or underqualified (leading to costly errors).

The advantages of the division of labor have been challenged by Marglin (1974), but when all is said and done his attack on one of the oldest and most widely believed tenets of economics has been beaten back without causing serious damage. Landes (1986) points out that Marglin fails altogether to deal with the Babbage argument and that his "evidence" for the falseness of Adam Smith's famous pinmaking example is based on a misreading of the literature. Experience and learning by doing are simple facts of life. Perhaps in a pin factory or an automobile assembly plant, the simplest jobs can be learned quickly and little more is learned after a few weeks, but in most skilled jobs, years of apprenticeship are required. Whether the difference between me and my dentist is due to innate abilities or to training, in neither case is it likely that productivity would be enhanced by us swapping jobs.

Yet the division of labor did not require factories. Domestic industries practiced it, and a large part of the function of the merchant entrepreneur was to shuttle goods in process from one cottage to another. In activities where technical factors made domestic production infeasible, such as fulling and calico printing, the manufacturer carried out the work in a "mill." Domestic industry did the rest. Decentralized specialization had advantages, but it also had costs, such as the transport costs of goods in process and the transactions costs of measuring and counting output at each stage.[107] As the division of labor became finer, the final products more complex, and the equipment more expensive, the costs of geographical dispersion rose, and firms switched from decentralized to centralized production.[108] The biggest advantage of rural domestic industries was their ability to switch labor back and forth from industrial to agricultural activities and thus exploit off-season labor. In effect, this means that outwork had access to cheaper labor than factories. It has been argued that the long survival of domestic producers in Britain, as opposed to the swift victory of the factory in the United States, was due to the differences in the seasonality of the demand for labor as British agriculture relied more on grains with its highly seasonal labor demand pattern (Sokoloff and Dollar, 1997). As the short-term mobility of labor increased with transport improvements, this advantage gradually diminished. The long-term decline in transportation costs tipped the balance in favor of the factory in other ways.[109] The rise of factories and changes in technology during

---

[107] S.R.H. Jones (1982, p. 126) minimizes the importance of transportation costs, but he fails to take into account that the geographical dispersion of work involved more costs than just the direct transport costs. Bad weather, for example, could totally disrupt the supply of raw materials and goods in process and thus wreak havoc on production and delivery schedules.

[108] A detailed summary of the advantages of the two systems can be found in Szostak (1989).

[109] Declining transport costs basically led to an increased division of labor, and while a rough division of labor was consistent with putting-out, as the division became finer the advantage moved toward factories. More integrated markets also led to a greater demand for standardized products and for quickly changing national fashions; here, too, factories had an advantage. See Szostak (1989, p. 348).

the Industrial Revolution can be seen as a prime example of co-evolution. Knowledge and business organization are both subject to autonomous innovation and selection processes, but also affect each other. Technological progress led to lower prices and better or new products, which increased demand and thus expanded the market; an increase in the extent of the market further led to a finer division of labor which increased productivity further and led to changes in industrial organization. This kind of positive feedback process serves as a good illustration how the Industrial Revolution can be seen as a self-reinforcing process.

All the same, the transition process took a long time and was far from monotonic. For many industries, factories did not mean the instant end of domestic industry but its temporary expansion, because when some activities were moved to mills, there was increased demand for the output of those production stages that remained for the time being in workers' homes. In some industries growth occurred through the expansion of the domestic industries.[110] Berg (1994, p. 274) points out that in the Birmingham metal trades, the industry's growth brought about a bifurcation in which large firms, some of which worked through factories, and domestic producers expanded at the expense of "substantial artisans." The final collapse of domestic industry did not come until the middle of the nineteenth century. In the long run, however, the triumph of the factory was as complete as it was inevitable.

The large industrial mills, emerging during the Industrial Revolution, created new management problems not hitherto encountered. Before the large factories there had been large firms, but these firms had been primarily commercial in purpose, and rarely operated facilities larger than warehouses or merchant ships. Agricultural estates, too, were often managed on a large scale but did not require the degree of coordination and direct control needed in manufacturing. The factories created a demand for a new skill hitherto largely confined to military commanders and sea captains, the need to organize, coordinate, and “run” substantial groups of people engaged in complex tasks in which the action of each individual affected those of all others and the nature of the outcome. They created new problems of labor and information management, including cost accounting, internal communication within the firm, coordination with other firms in the industry, negotiations with external suppliers and workers, and technical information management such as staying up-to-date on new industrial practices and other innovations.

Managerial ability was a form of human capital, and by all accounts it was not in generous supply in the British economy during the Industrial Revolution. There was not much separation between management and ownership: the entrepreneur usually was in charge. If anyone else was to be delegated any power, they would be in most cases be partners, a status awarded mostly to sons of partners or investors, although in a few cases technical expertise helped too.[111] But how did he carry out his day-to-day tasks of coordination and management? The managerial revolution, in which large corporate structures were managed according to reasonably well-understood principles of the flows of information and authority that could be taught and diffused was still many decades away, and the “visible hand” as Alfred Chandler termed it was still rather shaky in the early stages of modern manufacturing.[112] Much of it was improvised, learned by experience, stumbled into. Often, serious and

---

[110] An example is the career of Peter Stubs, a Lancashire filemaker, whose business was largely based on a network of outworkers run from the inn he kept in Warrington until he built his first workshops nearby in 1802.

[111] Some of the famous partnerships of the Industrial Revolution were based on close personal trust and complementarity between the two partners allowing a division of labor that often was the key to success. The most famous partnership was that of Matthew Boulton and James Watt, the classical wedding of business acumen and technical genius. At the Scottish Carron works, a somewhat similar symbiosis emerged between John Roebuck and Samuel Garbett. At the Etruria pottery works, Josiah Wedgwood's partner, Thomas Bentley, took care of sales, with Wedgwood dealing mostly with the production side.

[112] Alfred Chandler (1977, p. 70) has argued that early cotton mills in the United States were “run by merchants for merchants.” He shows that apart from a few exceptional firms, the first examples of modern management occur in railroads where the needs for control and coordination were especially acute and that the management techniques of the modern corporation originated in the railroad industry. Things in Britain were not much different and Berg (1994, p. 207) concludes that the Chandler thesis can provide no exclusive model of the development of eighteenth and nineteenth century industry.

costly managerial errors were made, especially due to primitive accounting.[113] Most managers, including Watt and Wedgwood, carried out their own correspondence and much of the clerical work. Indeed, in retrospect it is surprising that things worked out as well as they did. Some scholars have cited Josiah Wedgwood as an example of a successful and modern innovator of management techniques, though Hudson (1992) rightly points out that the great potter from Burslem was anything but a typical entrepreneur.

To some extent, the way that factory masters coped with management problems was through subcontracting. Pollard, in his classic and unique work on the subject points out that subcontracting, a remnant of the domestic system, survived into the factory age "if not as a method of management, at least as a method of evading management" (Pollard [1965] (1968), p. 52). By subcontracting, entrepreneurs could shift the risk around, make others responsible for their mistakes, and reduce overhead costs. But specialization and comparative advantage were important too. Master mechanics and builders came to the factories to install, maintain, and repair equipment on their own time, with their own tools, accompanied by their own paid assistants and carried out the job according to their own judgment and taste. Coal, cloth, and cotton yarn were produced using this system.[114] Technologically complex tasks were farmed out to mechanical or civil "consulting engineers" whose status as independent consultants crystallized in the last third of the nineteenth century. William Lazonick (1991, p. 140) has extended this thesis by pointing out that on the shop floor itself many of the functions that eventually were to be carried out by management through a hierarchical structure of foremen and supervisors were still carried out in the first half of the nineteenth century by a "labor aristocracy" of skilled, well-organized set of operatives. These operatives exerted a fair amount of independent discretion over both the laborers and the equipment under their control even if formally the entrepreneur owned the capital and employed the workers.

Subcontracting or "out-sourcing" as it is called today is neither inefficient nor a sign of "incomplete development" but was a rational result of specialization. In a world of costly and asymmetric information it often makes mores sense for a firm to hire an outsider to carry out a certain activity rather than do it by itself. All the same, there may have been cases in which subcontracting occurred largely because supervising a large number of people and activities was difficult in a age before large-scale modern management. Especially in coal-mining this may have been a problem and that sector was almost entirely run on a subcontracting basis. Subcontracting also relieved the firm from the need of computing complicated payrolls and by definition farmed out much of the labor supervision to lower levels.

The importance of the factory as a social institution can hardly be overestimated. The divorce between household and workplace imposed substantial costs on the industrial worker, from the psychic costs of having to witness family members supervised and monitored by others to the very real costs of the time spent on commuting (Smelser, 1959). The introduction of discipline and order into the lives of workers was another dramatic novelty. Until the Industrial Revolution discipline was largely a family matter. Industrial workers, whether they were independent artisans or part of a putting-out system, rarely encountered the phenomenon. Even on board merchant ships discipline could only be enforced by means of harsh penalties. The transition was not sharp; many factory owners hired whole families and used the family as a tool to enforce discipline.[115] Yet workers detested the mills and resisted discipline, and employers were often desperately looking for solutions

---

[113]A good example is the otherwise well-run firm of Boulton and Watt, where nobody had a clue as to which departments were earning or losing money, and the Scottish Carron iron company in which one manager estimated a profit of £10,500 when in fact £10,000 had been lost (Pollard [1965], 1968, p. 267). Overproduction and other errors of judgment occurred so often that one thoughtful economic historians sighs that they "can hardly fail to diminish any estimates of the commercial acumen of the cotton entrepreneurs (Payne, 1978, p. 189).

[114] In 1833, half the child workers employed in cotton spinning were employed by the mills, the rest worked for other operatives (ibid., p. 58).

[115] For instance, Robert Peel's factory in Bury employed 136 workers in 1802, of whom 85 belonged to 26 families (Smelser, 1959, p. 185).

to the stubborn problems of absenteeism, drunkenness, sloppiness, and unruliness. "The concept of industrial discipline was new, and called for as much innovation as the technical inventions of the age," writes Pollard ([1965] 1968, p. 217). Firms designed incentives to bring about the discipline, but they also preferred to hire women and children, who were believed to be more docile.

The advantages of introducing worker discipline were not identical to those realized by the factory system as such, as the two were not always coincidental. The gains of discipline have traditionally been regarded as the advantages of coordination. Factories required coordination between different activities of the laborers, as well as between labor and capital. Equipment such as steam engines, overhead costs such as heating, lighting, and fuel, and maintenance and supervisory personnel were fixed costs in the short run, and so if workers were absent or lazy there was costly waste involved. Above all, employers needed workers to be punctual.[116] Discipline was also necessary, however, to maintain quality standards, to avoid embezzlement, to prevent fights between workers, and to deliver goods in time. The equipment handled by workers was expensive, so that errors could be very costly for the capitalist. Industrial and mining accidents due to workers' negligence could be expensive and led to strictly enforced rules. Discipline, by regulating and equalizing the amount of time and effort supplied per worker, also saved on hiring costs (as workers were made more uniform) and reduced the variance of labor input and thus of output. Discipline, as a substitute for monitoring, saved on costs as it internalized into the worker's behavior the objective function of the firm. To be sure, it can be argued that some of the costs of the absence of discipline could be overcome by holding larger inventories and by adjusting hiring practices to absenteeism (Clark, 1994). But apparently such alternatives were expensive and the advantages of discipline were such that most of the famous entrepreneurs of the time, including Josiah Wedgwood, Richard Arkwright, Samuel Oldknow, and Matthew Boulton struggled with the problem. Clark's argument that discipline was a means to extract a greater effort from workers and could be viewed as advantageous to them if it raised their income is interesting but does not contradict the more technical advantages of discipline (Clark, 1994).[117]

Finally, it seems plausible that the "authority relations," to use Williamson's (1980) term, that came to dominate interactions between capitalists and employees in factories were instrumental in overcoming resistance to technological progress. In the extreme case, the employer not only controlled labor, inventories, and fixed capital but could also choose the technique of production by himself. Outworkers tended to be at the forefront of resistance to new technologies out of fear that laborsaving machinery would reduce the demand for their labor (Calhoun, 1982). Authority and discipline might have reduced, at least for a while, the ability of labor to resist technological progress. The factory, however, did not solve the problem of resistance altogether; unions eventually undermined the ability of the capitalist to exploit the most advanced techniques. Collective action by workers imposed an effective limit on the "authority" exercised by capitalists. Workers' associations tried to ban some new techniques altogether or tried to appropriate the entire productivity gains in terms of higher piece wages, thus destroying the incentive to innovate. On the other hand, such strikes often led to technological advances aimed specifically at crippling strikes (Bruland, 1982; Rosenberg, 1976, pp. 118-119).[118] On balance, it

---

[116] Employers reserved their harshest fines for latecomers, whereas the prize for good (probably docile) workers, not surprisingly, was a clock (Landes, 1983, p. 229).

[117] Clark's view of factory discipline is tantamount to an "Alcoholics Anonymous" view of workers in which they willingly commit themselves to a system that coerces them to work harder (and thus eventually earn more), which they would not do if they were left to themselves. As Langlois notes, this is tantamount to the teams of bargemen in pre-Revolutionary China who allegedly hired an overseer to whip them. In standard neoclassical models such behavior can only be understood if the utility function has a very unusual non-concave shape (so that local maximization does not lead to a global optimum) or if there were strong interdependencies between workers' productivities, so that each worker would only work hard if he or she knew that discipline would make other workers work just as hard. The latter explanation seems by far more plausible but oddly enough seems to be rejected by Clark.

[118] The most famous example of an invention triggered by a strike was that of the self-acting mule, invented in 1825 by Richard Roberts at the prompting of Manchester manufacturers plagued by a strike of mule operators.

is hard to know whether the decentralization of the putting-out industry, with its obvious potential of "divide and rule," was less conducive to technological change than factories -- yet this dimension has been altogether missed by scholars absorbed by static efficiency gains and transactions costs.

6. The Consequences: The Standard-of-Living Debate

The standard-of-living debate concerns what happened to living standards during the Industrial Revolution. It is one of the most lively yet most inconclusive debates in the entire Industrial Revolution literature. The discussion has been complicated in part because it became intertwined with political and ideological elements, the "optimist" school largely finding its supporters among the conservatives, the "pessimist" school mostly drawing upon socialist and left-leaning scholars. The philosophical question whether industrial society has been a positive development in human history reaches far beyond the relatively modest boundaries of economic history. What should have been a purely quantitative debate about numbers and deflators has divided scholars deeply on lines that correlate strongly with ideological positions. Those like E. P. Thompson and E. J. Hobsbawm, who have regarded industrial capitalism as enslaving and alienating, have tried to round off their position by arguing that it was also immiserizing. Those like T. S. Ashton and R. M. Hartwell, who are sympathetic to bourgeois capitalism and the achievements of free-market societies, have insisted that industrialism was liberating as well as enriching. Some of this ideological baggage seems to have been shed in the past decade, but scholarly opinion has remained divided.

Beyond that, however, there is a certain ambiguity regarding the terms on which the debate is being conducted. This ambiguity has been explained well by Hartwell and Engerman (1975) and further refined by Von Tunzelmann (1985). Three separate debates can be distinguished, which need to be kept logically separate:

1. The factual debate, which is concerned with what actually happened in Britain between 1760 and 1830 or 1850.

2. The counterfactual debate, which tries to identify the *net* effect of the Industrial Revolution on living standards. This question is logically equivalent to asking what would have happened to British living standards if everything had been the same in the period in question except for the technological changes of the Industrial Revolution.

3. The hyper-counterfactual question, which asks whether, given everything that happened, it would have been possible to follow a set of economic policies that would have made economic welfare more than it actually was.

The answer to the *second* question, whether without the Industrial Revolution living standards would have held up as much, is eloquently answered in a famous passage by T.S. Ashton in the closing paragraph of his little book (1948, p. 111): "There are to-day in the plains of India and China men and women, plague-ridden and hungry, living lives little better . . . than those of the cattle that toil with them . . . Such Asiatic standards, and such unmechanized horrors, are the lot of those who increase their numbers without passing through an Industrial Revolution." A simple calculation confirms Ashton's eloquence: If we take the weights computed by Crafts for labor, capital, and natural resources, we can compute the change in income per capita that would have occurred due to the growth of population and its pressure on other resources but without any productivity increase.

The counterfactual exercise is set up as follows: Assume that labor and resources changed at their actual historical rates and constrain productivity growth to zero. We have to make some assumptions about the counterfactual rate of capital accumulation. Three alternative assumptions will be employed: (1) the capital/labor ratio would have remained the same (requiring a savings ratio higher than the actual one), (2) the savings ratio would have remained at its historical level (that is, rising gradually), and (3) the savings ratio would have remained fixed because of the lack of suitable investment projects and the changing age structure of the population. Table 1.3 below presents the decline of income per capita implied.

TABLE 1.3 Counterfactual Decline in Income per Capita "Without" an Industrial Revolution (annual changes, in percentages)

Period	Assumption 1	Assumption 2[a]	Assumption 3[b]
1760-1800	0.045%	0.125%	0.185%
1800-1830	0.15%	0.41%	0.46%
Income in 1830 (1760=100)	93.9	84.1	80.9

a Assuming savings rates equal means for period.
b Assuming savings rates equal actual ones for first decade in the period.

SOURCES: Rates of change and shares from Crafts (1985a, p. 81). Savings rates from Feinstein (1981, p. 131). Cols. 2 and 3 required the estimation of capital/output ratio, computed from data in Mitchell (1988, p. 864) and Feinstein (1981, p. 136).

The calculations in Table 1.3 actually understate the hypothetical decline in living standards slightly, because they do not take into account the war-related shocks and the string of poor harvests that plagued Great Britain. All the same, they indicate that in the absence of an Industrial Revolution, a rising population -- as Malthus had predicted -- would have encountered declining living standards.

Yet the picture is more complex than that. The closest we can get to a controlled experiment of an economy that had a history similar to Britain's in terms of population growth and supply shocks, but without the Industrial Revolution, is Ireland's. Ashton used the example of Ireland as a warning against what could happen without industrialization, but clearly there are no such simple lessons to be learned from the Irish example. In fact, *average* living standards in prefamine Ireland did not decline much, even if there was a deterioration in the distribution of income (Mokyr and O Grada, 1988). The Great Famine, of course, was a hugely traumatic event that might well have been, if not averted, much mitigated had Ireland developed more of a modern sector. Had the potato blight not happened, however, our verdict regarding this example of a nonindustrializing country that experienced population growth might have been less harsh. Much of continental Europe also experienced population growth in this age, yet experienced neither an intensive rate of industrialization nor grievous famines. The best we can do is to conclude that Ireland may have been more *vulnerable* to accidental shocks because of the absence of an Industrial Revolution.

Turning to the *third* question, the hypercounterfactual one, modern research has clarified the issues and made an argument regarding the possibility that a more enlightened policy could have smoothened the pains of industrialization. Two of the most prominent cliometricians have made, from quite different points of view, arguments to the effect that "the thesis of the Hammonds that a suitably enlightened government could have brought about higher living standards is vindicated" (Von Tunzelmann, 1985, p. 221). Von Tunzelmann employs dynamic programming to show that it was possible for the British economy to have attained the final values of 1850 and yet have supported a higher consumption level. In the actual experience, in this view, industry tended to be too capital intensive in its early stages. Of course, such an optimal path could only be achieved by the deliberate interference of the government into the price system. Such an interference would, however, have had further ramifications that Von Tunzelmann does not explore. His important insight that things *could* have been better than they were does not necessarily support an argument that government interference would have moved the economy's path from the actual to the optimal. In a slightly different vein, Jeffrey Williamson (1990a) argues that Britain underinvested in its overhead capital, especially in urban areas. The rate of return on social overhead capital was very high, but Williamson argues that an unfair and inefficient tax system led to what he calls "public sector failure." As a consequence, Britain's standard of living was affected by an imbalance between private and public goods. Overhead projects such as sewage, water supply, fire protection, public health, and other "urban amenities" were undersupplied. Williamson's thesis is similar to John Kenneth Galbraith's analysis of the U.S. economy in his famous *The Affluent Society*.

The first of the three debates, the actual standard-of-living debate, is the main battlefield on which scholars have argued for decades. A summary of the debates and some of the best-known papers can be found in Taylor (1975). By the mid-1970s the debate had reached something of an impasse in which neither camp had scored an all-out victory and most other scholars turned elsewhere with their interests. In the 1980s, however, a number of important contributions were made by economists. The debate has bifurcated into one concerning purely economic indicators and a more inclusive set of biological indicators. The most important contributions to the economic evidence in the 1980s were made by Feinstein and Crafts, who examined aggregate consumption, and by Lindert and Williamson's work on real wages.

The message that these economists draw from their evidence was remarkably consistent. Their conclusion is that living standards remained more or less unchanged between 1760 and 1820 and then accelerated rapidly between 1820 and 1850, so that by the middle of the century living standards had improved considerably for a number of decades. Feinstein (1981, p. 136) estimates that consumption per head in 1841-1850 was 72 percent higher than in 1811-1820, and Crafts estimates the rate of growth of per capita consumption between 1821 and 1851 at a lower but still respectable 45 percent (1985a, p. 95). Lindert and Williamson estimate real wage growth between 1819 and 1851 at 80 percent for all "blue collar workers" and 116 percent for "all workers" (1985a, p. 187). Crafts has revised these estimates as well, tempering but not overturning the new optimist message.

Yet these economic indicators failed to sweep the field. Although it is reasonable to conclude that standards of living did not *decline* for extended periods of time during the Industrial Revolution, the optimist victory declared by Lindert and Williamson has turned out to be premature. For one thing, the optimists have essentially conceded the entire period before 1820, thus focusing the debate on the three-and-a-half decades between the Battle of Waterloo and 1850. Yet even for this period, ambiguities remained. The aggregate consumption data produced by Feinstein and refined by Crafts are residuals, the difference between highly speculative data of output and investment. By construction, they cannot account for changes in income distribution, and Feinstein warned that "the basic estimates are far from reliable" and that they should be used with caution. To be sure, they were lent much reinforcement by the Lindert-Williamson wage data, but most scholars felt that more confirmation was needed to disperse remaining doubts.

Such confirmation has not been forthcoming. On closer inspection, the real wage data is found to suffer from a number of rather serious defects. One is that they cover only limited data points and that the choice of the end year (1851) by Lindert and Williamson is unfortunate, because that happened to be a year of unusually low prices.[119] The nominal wages fluctuated a lot but their secular movement was quite stationary in this period, so that the rise in real wages came almost exclusively from falling prices. Hence, the optimist conclusion is highly sensitive to the correct specification of the price deflator, and its deficiencies weaken the optimist finding even further.[120] When those two biases are corrected together, real wages rise so slowly that Huck (1992, chap. 2, p.

[119] The only price index covering the entire nineteenth century, the Rousseaux index, points to 1851 as the cheapest year before 1885, and the index is about 17 percent lower than the average for 1840-1850. Had Lindert and Williamson chosen 1847--an unusually expensive year--the rise in real wages would have been half of what they report.

[120] This point was made by Crafts (1985d), who points out that Lindert and Williamson use only cotton as their textile price and that cotton prices fell faster than wool. Correcting for these defects, he concludes that the index rose slower before 1820 and fell slower after 1820 than Lindert and Williamson estimate.

22) concludes that "1850, or some point in the 1840s, should be seen as the key turning point, as opposed to [the] 1820s."[121] Some of the new series produced are illustrated in Table 1.4.

In a series of recent papers, Charles Feinstein has recalculated the real wage series de novo. His coverage is considerably wider and deeper than the Lindert and Williamson calculations and his deflators have many of the defects that mar the Lindert and Williamson figures removed from them.[122] Feinstein's finding are nothing short of devastating to Lindert and Williamson's newly discovered optimism. As table 1.5 shows, the increase of real earnings shows a barely perceptible increase from the end of the Napoleonic wars till the mid 1840s, and shows an acceleration in the third quarter of the nineteenth century, although this movement is still full of leaps and bounds. Feinstein himself summarizes his findings that "it was not until the mid-1840s that real earnings broke away and by the middle of the nineteenth century they had moved to a new level" (1997, p. 45).

TABLE 1.4 Nominal Wages, Real Wages, and Prices, 1787-1851

1	2	3	4	5	6
	Nominal Wages (male adults)	Real Wages (Blue Collar)	Real Wages (all)	Real Wages (blue Collar, revised)	Cost of Living, Revised
1797	58.97	53.61	42.48	60.6[a]	146.3[b]
1805	75.87	51.73	40.64	--	177.5
1810	84.89	50.04	39.41	--	207.1
1815	85.30	58.15	46.71	--	164.3
1819	84.37	55.68	46.13	69.9	166.6
1827	83.11	69.25	58.99	79.5	131.9
1835	88.77	83.43	78.69	88.0	109.4
1851	100.00	100.00	100.00	100.00	100.0

a - 1781.
b - 1795.
SOURCES: Cols. 2-4: Williamson (1985, pp. 14, 17). Col. 5: Huck (1992, p. 48)
Col. 6: Lindert and Williamson (1985b, p. 148).

[121] Lindert and Williamson's nominal wage series shows virtual stability: In 1819 the wage of all "blue collar workers" was 101.84 (1851 = 100). The revised price index they themselves propose in response to Crafts's critique is 166.6 in 1819 and 141.4 in 1847 (1851 = 100). If we assume that nominal wages in 1847 and 1851 were the same, the implied rise in real wages between 1819 and 1847 is only .52 percent per year. To be sure, 1847 was a year of extreme dearth (although less so than 1839), but the rate of deflation proposed by Lindert and Williamson is sharper than that of Crafts.

[122] Among the many corrections introduced by Feinstein in his new cost of living index is the use of a chained indices rather than single based indices, the inclusion of a host of products omitted by Lindert and Williamson, the introduction of a new index for clothing and replace Lindert and Williamson's very weak component for rent.

A major defect in Lindert and Williamson's calculations and corrected by Feinstein as much as possible is that the wage data cover only selected workers. By definition it covers only those employed in the "formal" sector, that is, receiving a wage. Under labor market equilibrium conditions, this objection is unimportant because the wage rate in the formal labor market and the implicit wages earned by the self-employed would move together. But much of the argument for the "modernization" of industry suggests that while factory wages were rising, the real income of most domestic workers and independent artisans were falling (Allen, 1992b, pp. 255-256; 296-297; Lyons, 1989). This discrepancy constituted the market "signal" that the death bell was sounding for much of the traditional sector; for our present purpose it means that using formal wages as a proxy for "labor income" may be quite misleading. Furthermore, not all formal market wages are equally useful. The estimates of agricultural wages are especially fragile, and because agricultural workers still constituted over 20 percent of the labor force in 1841, their fate is quite important. The income of farm laborers was determined in part by other factors, such as access to commons and a growing seasonal unemployment, especially of women (Allen, 1992a; Huck, 1992; Snell, 1985). Thus rising *wage rates* might well have been accompanied by falling *incomes* and *living standards* as growing redundancies in agriculture were not met by a rising demand for labor from nonagriculture, leading, in Allen's words, to "structural unemployment rather than increased manufactured output" (1992, p. 32). This complication was exacerbated by the decline in the custom of paying workers partly in kind, so that the rise in observed real wages could in part be spurious. Changes in nominal wages in agriculture differed from a 13 percent fall in the east to a 10 percent rise in the southwest between 1824 and 1851.

Most wage data used by Lindert and Williamson pertain to adult male wages. The justification for this is explicitly stated by them (Lindert and Williamson 1985, p. 194) to be that wage rates of women and children advanced as fast as those of adult male farm laborers (which was considerable slower than that of "all workers"). This conclusion, they feel, will not be overthrown by correcting for changes in employment. Recent research, however, has been divided on this issue. Horrell and Humphries (1992) confirm Lindert and Williamson's findings about the rise of adult male real wages, though not without some misgivings.[123] Yet their work clearly shows than male and female earnings did not move all the time in the same direction. Robert Allen (1992b, pp. 255-256, 296), who has studied the fate of rural laborers, has emphasized the sharp decline in employment opportunities suggesting that family income fell relative to male earnings. As Allen's males hardly experienced much real income growth, he concludes that before 1850 real family income in rural Britain declined.

Furthermore, rising real wages may have different interpretations. Even a firm believer in the efficiency of labor markets will concede that a rise in real wages may not be an indication of rising living standards if these rising real wages were a compensation for deteriorating labor conditions. If factory work and life in industrial towns and villages became more onerous, dangerous, or unpleasant, rising real wages would have the interpretation of a compensating differential. This effect has been measured in an ingenious paper by Brown (1990), who, like Lindert and Williamson, finds a significant rise in real wages yet concludes (pp. 612-613) that "there was virtually no improvement in living standards until at least the 1840s and perhaps the entire first half of the nineteenth century."

One way to try to circumvent these and similar problems is to look at microeconomic series for the consumption of a popular and income-elastic consumer good. Any such series would have the advantage that it would reflect living standards of both employed and self-employed workers and take into account both the level of income per capita and the inequality of its distribution. Food consumption series are shrouded in rather serious statistical uncertainty. Recent work on the problem, based on fragmentary and indirect data, seems to cast

[123]Horrell and Humphries add that secular income growth was interrupted by setbacks that tend to be underestimated by trend analysis based on a limited number of observations. They also note, as we have before, that the optimist findings depend crucially on price movements (they deflate their nominal series by Lindert and Williamson's "best guess" cost-of-living index), and insist that questions still hang over the speed by which price falls filtered down to the working class.

growing doubt that food consumption per capita was rising sharply during the Industrial Revolution.[124] More accurate are the series for domestic consumption of imported consumer goods, such as tobacco and sugar. After correcting for changes in prices and other effects, we can employ these data to infer what kind of income data (given estimated income and price elasticities) would have generated these consumption figures (Mokyr, 1988). The results lend no support to the view that living standards increased before the late 1840s. These findings have been corroborated recently by Horrell (1995) who has computed the change in consumption levels of the British working class from a sample of budget studies. She found (p. 580) that for working class families, real expenditures per household between 1787-96 and 1840-54 increased by about one half of a percentage point per annum, and that in fact they declined from the 1830s on, the period for which Lindert and Williamson argue that the highest rate of growth in real wages occurred.The dilemma is thus clear: If real incomes of the bulk of British workers increased, and yet they did not eat appreciably more, lived in crowded and unhealthy houses, drank no more sweetened tea, smoked no more pipes -- where did this money go? The consumption of a few small items like hard soap and iron goods may well have increased, but many of the commodities on which we have data, such as bricks, coal, and glass, were as much investment as consumption goods and cannot be used readily for the standard-of-living debate. The only commodity that clearly figures prominently as an item in the improving budgets of workers is cotton textiles.[125]

An alternative approach to the standard-of-living problem is to look at biological indicators of the standard of living. It has long been recognized that indicators such as life expectancy and physical health are strongly correlated with economic living standards. Indeed, some economists (notably Sen, 1987) maintain that such physical measures *are* the standard of living. Thus in the absence of unambiguous economic measures of living standards, economic historians have increasingly turned to biological measures to try to test the hypothesis of rising economic welfare before 1850. On the whole, these measures have failed to support the optimist case. The broadest measure is the crude mortality rate, which declined more or less in the same period identified by the new optimists as the period of rising living standards: At about 1760, the crude death rate for England was still about 27.5 per 1000, declining steadily (with a few reversals) to about 22.5 per 1000 by 1850. Gross mortality rates, however, are flawed indicators for many reasons, primarily because of their dependence on the age structure of the population. A better measure is the life expectancy at birth. This variable, too, shows some improvement over the entire period, but its rise stops in 1820, and it remains essentially static at about 40 years until 1860 (Wrigley and Schofield, 1981, p. 529). The sharp rise in consumption and real wages claimed by the new optimists should have produced, through improved nutrition and better living conditions, a rise in life expectancy, perhaps lagged by a few years. Nothing of the sort happened. Data on infant mortality, though not available on a national basis, tell very much the same story. In a sample of seven parishes, Huck (1992) finds rising infant mortality rates in the period between 1813 and 1836, with no appreciable decline until 1845, precisely the years identified by Lindert and Williamson as the period of rapid improvement. More recent data reported by the Cambridge Group, based on family reconstitution and thus more representative of England as whole contain no consolation for the optimists: infant mortality rates (acknowledged by Williamson (1982) himself as a good indicator of the standard of living) declined sharply in the last two decades of the eighteenth century, but then rose slightly to 1837 (Wrigley et al. 1997).

---

[124] A detailed attempt to patch together existing data is carried out by Helling (1977), whose estimates of per capita grain and meat consumption show no improvement until the mid 1840s. Lindert (1992) argues that workers spent their incomes on rapidly expanding nonfood items. Clark, in his essay later in this book, concludes that given what happened to British agricultural output, sharply rising food consumption is unlikely.

[125] Of some interest here is the very peculiar paper published by Clark, Huberman and Lindert (1995) in which they face the problem that under the premise that real wages increased as rapidly as Lindert seems to believe, British workers were not spending their higher incomes on food either. The paper goes through a number of rather contrived arguments to settle this "puzzle." There is no puzzle, of course, once we realize that the premise is false and that the food consumption data, much like those of tea, sugar, and tobacco, neatly track the real wage data delineated by Feinstein.

A biological indicator that has enjoyed considerable interest in the last few years is human height. It has become widely accepted that height is a function of net nutritional status, that is, the amount of food taken in by children and adolescents net of demands made on their bodies by labor and diseases. All other things equal, a child born in a family that enjoyed a higher standard-of-living would grow up to be taller. The idea that observed height data could therefore be used to approximate the elusive standard of living was proposed by Fogel (1983) and his associates and has since then stimulated a large number of research projects. The research that is most pertinent to the standard of living debate in Britain is Floud, Wachter, and Gregory (1990) and Komlos (1997). Their finding is that net nutritional status, as measured by stature, increased between about 1760 and 1820 and then went into a secular decline for half a century. Indeed, the cohorts born in 1850-1854 are shorter than any cohort born in the nineteenth century, and the levels attained in the first decades of the century are not attained again before the last decade (Floud, Wachter and Gregory, chap. 4, passim). Based on this evidence, they maintain, the debate on living standards during the second and third quarters of the nineteenth century is still very much open, and (p. 305) "if there were significant gains in real incomes for the working class between the 1820s and the 1850s they were bought at a very high price." Komlos's figures are even more pessimistic, and prompt to regard this as a major "puzzle" in economic history. In other words, if there were economic gains, they did not lead to *physical* improvements in the lives of English men and women.

The incongruity of the biological indicators, which tend more to support the pessimist case, and the aggregate economic indicators, which on the whole present a mixed case, can be reconciled in three different ways. One is that the biological indicators pertain to the population as a whole, including the domestic sector, paupers, and the "informal" economy of the urban poor, whereas the real wage data pertain largely to the modern and formal sector and thus are not as representative. To put it differently: The Industrial Revolution brought forth losers and gainers. Real wage data alone tend to reflect more the situation of male employed workers, who were predominantly gainers, than upon domestic workers, many of whom were female and self-employed and who, by and large, ended up on the losing side. The failure of microlevel consumption data to reflect the rise in real wages is consistent with this view. A complementary explanation may suggest that while real wages improved, other aspects of living standards deteriorated. These would reflect not only urban living conditions and the harsh conditions of factories but also some less obvious factors, such as the loss of flexible choice between leisure and income brought about by the factory system. Thus rising real wages simply compensated the workers for other losses and there is no obvious case for "improvement." Moreover, as Komlos notes, rising real incomes could be consistent with changes in relative prices that made healthy (protein-rich) foods more expensive. Finally, it can be argued, of course, that biological indicators such as height, while easy to measure and estimate, are difficult to interpret and that economists should treat them with even more caution than wage or income data (Mokyr and Ó Gráda, 1996).[126] In view of the fragility of much of the statistical material on aggregate income and consumption, however, this view seems difficult to sustain. At this stage, therefore, it has to be inferred that the evidence of a rise in living standards before 1850 is simply too weak to be convincing.

The pessimist case itself, however, should be tempered by acknowledging the well-known pitfalls involved in measuring changes in living standards in an age of rapidly changing technologies. All quantitative studies of living standards measure in the final analysis quantities of goods that incomes can buy. They fail to account for changes in quality. A typical textile product in 1830 was not only cheaper than in 1750 but was also better in terms of the evenness of its fabric, its durability, its ability to absorb and maintain color, its ease of laundering, and so on. The same is true for a wide range of products, from iron pots to glass to steel pens to printed illustrations in books. Moreover, a number of inventions made during this period created completely new

---

[126] A particular difficulty with interpreting height data is that changes in measured adult height reflect changes in living standards in the past, but any attempt to time this accurately is difficult, since physical growth occurs over more than two decades (and possibly longer, since malnourished women tended to have babies who grew up shorter) and any prolonged "dietary insult" could lead to stunting. Moreover, height was also determined by morbidity, and such exogenous changes as changes in disease regime could contaminate the relationship between observed health and anything resembling an economic standard of living (see e.g. Voth and Leunig, 1996)

products, making welfare comparisons very difficult: Traditional measures of real wages and national income do not adequately capture the economic value (or additional consumer surplus) of the decline of smallpox, the introduction of gaslighting, or the use of anesthesia during surgery. Against this we have to weigh the increased adulteration of the food and drink bought by the working class (Burnett, 1966, pp. 99-120), the negative effects of the ‘disamenities’ of urbanization, and the loss of outdoor relief with the reform of the poor laws in 1834. Feinstein (1997, p. 47) reckons that these three effects reduce the gain in standard of living of the average family in the United Kingdom between the 1770s and the 1850s from about 28 per cent to 8-13 per cent. At the same time, it should be added that certain highly aggregative measures of economic well-being such as the Human Development Index (Crafts, 1997) paint a more optimistic picture. Crafts's index rises steadily and in rather stable fashion from 1760 to 1850. Even corrected for changes in equality and the disparity in the achievements of genders, the various indices display a steady rise. A closer look at the Living Standards Indicators produced by Crafts does, however, weaken his optimist inferences regarding the critical years between 1815 and 1850: Height of army recruits show a decline, while the critical demographic variables show no improvement. Much of the effect clearly derives from the sharp increase in GDP per head, a variable as we have seen is still in serious dispute. Moreover the HDI is computed additively and thus trades off demographic variables and income per capita linearly against literacy. Given the somewhat questionable role of literacy in this age (being able to read may not mean much if people rarely actually read), this procedure is thus not wholly satisfactory and certainly does not rehabilitate Lindert and Williamson's now-discredited optimism.

Part of the standard-of-living controversy is the debate over what happened to the inequality in income distribution. The famous Kuznets curve hypothesis (Kuznets, 1955) suggests that during the first stages of industrialization, income distribution became more unequal, eventually reaching a peak and then improving afterward. A worsening income distribution is one obvious way to reconcile rising per capita income and stagnant living standards for the majority of the population. The argument is discussed by O'Brien and Engerman (1981, p. 174), who maintain that given the rate of growth of income per capita between 1800 and 1850, an unchanging income level of the bottom 80 percent of earners would have meant that their share in income decreased from 75 percent in 1800 to 41 percent in 1850. Such a sharp worsening being unthinkable, they dismiss the argument. The revision of per capita growth rates, however, makes this argument less compelling. At a per capita income growth of perhaps 0.7 percent per year between 1800 and 1850 (instead of the 1.2 percent estimated by Deane and Cole and used by O'Brien and Engerman), a relatively slight sharpening in income distribution might have reduced the growth of income of the bottom of the income distribution to little more than a trickle.[127] The most dedicated proponent of the applicability of the Kuznets curve to Britain during the Industrial Revolution is Jeffrey Williamson (1985), although this belief in part undermines his view that living standards improved rapidly after 1819. There is, however, some doubt about what precisely happened to income equality during the critical years between 1800 and 1867, and until this doubt is cleared up, it is hard to draw any firm conclusions about how changing inequality affected living standards.[128] In a critique of Williamson's work, Feinstein (1988b) denies the applicability of the Kuznets curve to the British experience during the Industrial Revolution and argues that inequality remained more or less unchanged. Some complicating factors, however, still have to be fully accounted for. For instance, there is a difference between the inequality of the distribution of income among households and the distribution among individuals. If poorer families tended to increase family size over time relative to

[127] Using the rise in inequality estimated by Lindert and Williamson (1983a) and assuming the share of the poor was little changed between 1850 and 1867 yields a growth of slightly over 0.4 percent in the incomes of the bottom 90 percent of the income distribution between 1800 and 1850. However, the decline of the share of the bottom 90 percent from about 54 percent to about 47 percent is, by Lindert and Williamson's computations, entirely accounted for by the sharp decline in the earnings of the people in the bracket between the bottom 65 percent and the bottom 90 percent, that is, the upper bracket of the bottom 90 percent. Removing these "lower middle class" people and concentrating on the bottom 65 percent reverses the picture, and incomes in this group increased by 0.90 percent per year.

[128] Allen (1992b, p. 285) argues that in agriculture landlords were the only gainers from the agricultural revolution before 1850.

richer families, a constant distribution over household would in fact imply a growing inequality among individuals. A further complication is the decline in poor-relief support prompted by the Poor Law Reform Act of 1834, when spending on poor relief fell from over 2 percent of national income to about 1 percent (Lindert, 1992). Obviously, the reform sharpened the after-tax distribution of income, but it is as yet unclear to what extent changes in the poor law affect the standard-of-living overall.[129] Horrell's work (1995), although based on a very different set of sources, lends some indirect support to the Williamson view of sharpening inequality before 1850, as she finds steeply increasing consumption by middle class families but practically none for the working class. It might be added that there is something ironic about the historiography of inequality here. The debate between Williamson and Feinstein has been about real wages and inequality. Whereas Williamson has argued for steeply rising blue collar wages, he has also maintained that inequality increased; Feinstein has taken the opposite position on both issues. Yet rising blue collar wages might have been, at first glance, more consistent with declining or at least constant inequality. This paradox may be resolved once Feinstein completes his project of estimating National Income. If his results lower estimated growth rates significantly, stagnant real wages may well be consistent with no dramatic changes in inequality.

7: An Assessment

The New Economic History has traditionally been iconoclastic, and the Industrial Revolution has not been immune from attacks on the usefulness of the concept. Such attacks are to be welcomed simply because they force a reconsideration and reevaluation of the conventional wisdom. The Industrial Revolution may not, in fact, have been nearly as abrupt and as sudden as some of its historiography suggests. Furthermore, there has been a tendency among some economic historians to identify the economic history of Britain in the century after 1750 *as* the Industrial Revolution. Such an identification is misleading and a-historical. Much, perhaps most, of what happened in the British economy at that time had little or nothing to do with the Industrial Revolution. Before 1830, most of Britain's land and the majority of its population were only affected by it in a roundabout way, many perhaps were not affected by it at all.

Yet its importance as an event in economic history stands undiminished. Before the Industrial Revolution technological change and economic growth did occur sporadically in the experience of Europe and Asia but were invariably checked by stronger forces. Much of the growth that other scholars observe in Europe before the Industrial Revolution was due to the expansion of commerce, itself largely a function of institutional change and propitious political circumstances. Such cases were usually slowed down or even reversed by institutional breakdowns or military events. Technology, by its very nature, is much less reversible and less likely to run into diminishing returns that commercial expansion. What the Industrial Revolution meant, therefore, was that after 1750 the fetters on sustainable economic change were shaken off. There were lags and obstacles to overcome before technological creativity and entrepreneurship could be translated into sustained economic growth and higher living standards, but the secular trend pointed clearly upward. What ultimately matters is the irreversibility of the events. Even if Britain's relative position in the developed world has declined in recent decades, it has remained an urban, sophisticated society, wealthy beyond the wildest dreams of the Briton of 1750 or the bulk of the inhabitants of Africa or Southern Asia in our own time. Britain taught Europe and Europe taught the world how the miracles of technological progress, free enterprise, and efficient management can break the shackles of poverty and want. Once the world has learned that lesson, it is unlikely to be forgotten.

Regarded with the critical eye of statistical analysis, the events of the Industrial Revolution themselves may seem to us small and even insignificant because they affected only limited areas and products. But historians' judgment is inevitably colored by hindsight and rightly so. Examining British economic history in the period 1760-1830 is a bit like studying the history of Jewish dissenters between 50 B.C. and A.D. 50. At first provincial, localized, even bizarre, it was destined to change the life of every woman and man in the West beyond

[129] The estimates of the share of the bottom 40 percent in income distribution range between 10 and 14 percent of income (Williamson, 1985, p. 71). A decline in poor-relief transfers from 2 percent to 1 percent would have, by itself, reduced the incomes of the very poor by something between 7 and 10 percent.

recognition and to affect deeply the lives of others, even though the phenomenon remained confined primarily to Europe and its offshoots. Although the center of the stage has long been taken over by others, Britain's place of honor in the history books is assured: It will remain the Holy Land of Industrialism.