Demand vs. Supply in the Industrial Revolution

The intelligent application of the formal principles [of economic theory] is . . . chiefly significant negatively rather than positively, for showing what is "wrong" rather than what is "right."

Frank Knight (1940)

In the process of what could best be called "making sense of the Industrial Revolution," few articles have been more influential than Gilboy's eloquent plea to view demand as an equal partner in bringing about the most profound economic change in human history.[1] The notion that *both* sides of the demand and supply equation come into play in the explanation of the crucial questions, such as "why England first" or "why the eighteenth century," has made sense to more than a generation of economic historians better trained in handling Marshall's scissors than Occam's razor. The appearance of Keynes's *General Theory* a few years later lent additional support to the notion that demand was somehow important.[2]

The present paper reexamines what may be termed the "Gilboy thesis." The concept of demand itself is too vague without additional clarification. What I shall therefore attempt is to reformulate the Gilboy thesis in consistent and testable form.

Journal of Economic History, Vol. XXXVII, No. 4 (December 1977).

I would like to express my thanks to Donald N. McCloskey, on whose very elastic supply of knowledge in economics, history, and econography I have liberally drawn. Others who have been of great help with comments and suggestions include Pekka Ahtiala, Reuven Brenner, Karl de Schweinitz, Stefano Fenoaltea, Jonathan Hughes, Eric L. Jones, Glenn C. Loury, Jacob Metzer, F. Michael Scherer, and Richard O. Zerbe.

[1] Elizabeth Waterman Gilboy, "Demand as a Factor in the Industrial Revolution," in A. H. Cole, ed., *Facts and Factors in Economic History* (1932); rpt. in R. M. Hartwell, ed., *The Causes of the Industrial Revolution in England* (London, 1967), pp. 121-38.

[2] Very explicit reiterations of Gilboy's thesis can be found, for example, in David Landes, *The Unbound Prometheus* (Cambridge, 1969), p. 46; D. E. C. Eversley, "The Home Market and Economic Growth in England, 1750-1780," in E. L. Jones and G. E. Mingay, eds., *Land, Labour, and Population in the Industrial Revolution* (London, 1967), pp. 206-59; Phyllis Deane, *The First Industrial Revolution* (Cambridge, 1969), p. 34; A. H. John, "Agricultural Productivity and Economic Growth in England," in E. L. Jones, ed., *Agriculture and Economic Growth in England, 1650-1815* (London, 1967), pp. 172-93. A recent restatement is in François Crouzet, "Western Europe and Great Britain: 'Catching Up' in the First Half of the Nineteenth Century," in A. J. Youngson, ed., *Economic Development in the Long Run* (New York, 1967), pp. 113-14. A variation on the same theme is in Nicholas Georgescu-Roegen, *The Entropy Law and the Economic Process* (Cambridge, Mass., 1971), pp. 246-47. Doubts concerning the demand hypothesis were sounded by W. W. Rostow, *How It All Began* (New York, 1975), pp. 14, 129, 172-73, and in a slightly different context by Donald N. McCloskey, "Did Victorian Britain Fail?" *Economic History Review*, 23 (Aug. 1970), 446-59.

I. DID THE DEMAND CURVE SHIFT?

It is possible to interpret the Gilboy thesis as a shift in aggregate demand, or as a shift in the market demand curve for industrial goods. Leaving the former for section III, we first consider whether the Industrial Revolution could be triggered by an outward shift of the industrial demand curve.

In her original article, Gilboy maintained that

> The factory could not become typical until demand had been extended . . . throughout the entire population to consume the products of large scale industry. . . . In order that a shift in the demand schedule may occur, individuals must be able to buy more units of a commodity at the same price, or the same amount of the commodity at a higher price . . . the entire schedule must shift upward, indicating a greater buying power.[3]

In this simple form the demand thesis is based on circular reasoning. A shift *of* the demand curve for manufactured goods can occur only if income rises, the price of nonmanufactured goods falls, or if a change in tastes occurs. Ruling out the latter for the moment, the shift in the demand curve must be caused by a rise in real income, and can therefore not serve at the same time as an explanation of it.

It is transparent that if a shift in the demand curve for industrial goods is to be used for the explanation of the rise of industrial output, the shift in the demand curve must be caused by factors other than the rise of output itself. Three alternative theories have been presented in the literature in this context, namely agricultural growth, expansion of foreign demand, and population growth.[4] It is worthwhile to examine these theories in more detail. The first and in many ways the most attractive hypothesis is agricultural progress. A decline in prices of agricultural goods will lead to an increase in the demand for nonagricultural goods if, ceteris paribus, the demand for agricultural goods is inelastic, as is usually assumed. While unimpeachable on a priori grounds, the theory runs into a timing dilemma: whereas agriculture prices fell in the first part of the eighteenth century, they started to rise after 1750, a trend that persisted until the end of the Napoleonic Wars both in Great Britain and on the Continent.[5] More-

[3] Gilboy, "Demand," pp. 122-26.

[4] W. A. Cole, "Eighteenth-Century Economic Growth Revisited," *Explorations in Economic History*, 10 (Summer 1973), 327-48.

[5] The evidence for France is summarized in C. E. Labrousse, *Esquisse du mouvement des prix et des revenus en France au XVIII^e siècle* (Paris, 1933), pts. II, III, IV, esp. pp. 137-66. For the nineteenth century, too, the timing does not work. The period usually associated with the most rapid economic growth in France (1840-70) was also one of rising wheat prices. See ibid., p. 141, and B. R. Mitchell, *European Historical Statistics* (New York, 1975), p. 742.

over, it is not always sufficiently emphasized that there should be a decline in the price of agricultural goods *relative* to the price of nonagricultural goods. The movements of relative prices are less clear-cut.

What was the contribution of falling agricultural prices to the rise in demand for nonagricultural goods? The order of magnitude of the impact of changing relative prices on demand can be estimated approximately following a procedure outlined in a recent study by R. A. Ippolito.[6] Since the price elasticities as well as the actual movement of prices are subject to large margins of error, the procedure followed has been to estimate a lower and an upper bound of the expansion of industrial demand due to the fall in relative prices.[7] The upper bound of the estimate is 26.1 percent for the entire period 1750-1850, or .26 percent annually, while the lower bound is −.25 percent annually.[8] The average rate of growth of individual output for the entire period can be estimated conservatively at 3 percent annually. The changes in relative prices thus account for at best 8 percent of the industrial expansion, while it is quite possible that their net effect was on the whole negative.

Some efforts have been made to rescue the argument by maintaining that rising agricultural prices due to an exogenous decline in

[6] Richard A. Ippolito, "The Effect of the 'Agricultural Depression' on Industrial Demand In England, 1730-1750," *Economica*, 42 (Aug. 1975), 298-312. Ippolito's overall conclusion is that the agricultural depression of 1730-50 did not constitute a major source of demand for industrial goods either. It should be emphasized that the estimates do not truly reflect the contribution of the agricultural revolution (i.e., shifts of the supply curve of agricultural goods) to the industrial revolution. This would only be the case if the supply curve of agricultural goods were perfectly elastic (as is assumed by Ippolito) or if demand were stationary. Neither of these assumptions seems plausible; hence, what is measured is the impact of price changes only, without further identification of their source.

[7] Let the demand function have the form $Q_M(D) = Y^a P_A{}^b P_M{}^c$. This means that we can write the relationship between the demand for industrial goods and the relative price $P = P_A/P_M$ as

$$\frac{\Delta Q_M(D)}{Q_M(D)} = \left[\frac{\Delta P}{P} + 1\right]^b - 1$$

An approximate value of b can be estimated by using the so-called Cournot relationship between the own and the cross elasticities of demand:

$$\gamma_A \epsilon_A + (1 - \gamma_A)b = -\gamma_A$$

ϵ_A is the own price elasticity of demand for agricultural goods and γ_A is the share of agriculture in total output. A reasonable estimate of γ_A would put it at around $\frac{1}{3}$. The value of ϵ_A is unknown, but is generally believed to be less than one in absolute value. Setting it equal to −.3 and −.7, we obtain estimates of b between −.35 and −.15.

[8] The relative prices underlying the calculations were obtained and computed from a number of series which overlap to some extent, or which had to be converted to relative prices using some assumptions about weights. The procedure followed was in each case to choose the extreme estimates, thus obtaining upper and lower bounds.

supply (for example, harvest failures) could, under certain conditions, stimulate industrial demand. The mechanism supposedly responsible for this is that a bad harvest raises agricultural prices. Since the elasticity of demand for farm products is usually thought to be less than unity, farm income would rise. If the farmers have a very high propensity to spend on manufactured goods, a harvest failure could allegedly cause a net rise in the demand for nonagricultural goods.[9] In a simple general equilibrium framework it can be proven that if both goods are normal, this cannot occur. Only if agricultural goods—taken as a whole, including both consumption and investment goods—are inferior goods (that is, the income elasticity for them is negative), demand for nonagricultural goods may (but does not have to) expand as a result of harvest failures. For all practical purposes, this version of the demand thesis can thus be ruled out. The proof of the theorem is relegated to Appendix A.

Eversley has maintained that the post-1750 rise in agricultural prices was not caused by a decrease in supply but by a rise in demand.[10] The absence of detailed output data precludes a direct test of this argument, but the work of Eric Jones and others demonstrates that agricultural productivity followed an upward secular trend throughout the eighteenth century, including the period 1750-1780. But it does not help to explain the growth in industrial production. If the increased demand for agricultural goods came at the expense of industrial goods, say through a change in tastes or income distribution, it runs directly counter to the Gilboy thesis. If it did not come at the expense of the industrial goods, where did it come from? Surely in this case rising aggregate income must be the underlying mechanism. To the extent that a higher aggregate income resulted from population growth, it will be dealt with below. But if higher income per capita is the source of the expansion of demand we are back in the same circular argument as before.[11]

[9] David S. Landes, "The Statistical Study of French Crises," this JOURNAL, 10 (Nov. 1950), 200-01. Phyllis Deane and W. A. Cole, *British Economic Growth, 1688-1959* (2nd ed.; Cambridge, 1969), p. 93. John D. Post, "A Study in Meteorological and Trade Cycle History: The Economic Crisis Following the Napoleonic Wars," this JOURNAL, 34 (1974), 338-39.

[10] Eversley, "The Home Market," pp. 240-46.

[11] A general discussion of the impact of agricultural fluctuations on the English economy is in J. D. Gould, "Agricultural Fluctuations and the English Economy in the Eighteenth Century," this JOURNAL, 22 (Sept. 1962), 313-33. Gould's attempts to explain the price movement by the operation and later repeal of the Corn Laws in the eighteenth century cannot be viewed as definitive—the export bounties were repealed in 1773, and it is hard to see how that could help raising prices.

A more convincing rationalization of the apparent paradox is provided by Eric Jones.

> It is unlikely that higher food prices after 1750 ever seriously eroded the expansion of the home market for consumer goods by eliminating the margin for such spending among the working population. Agriculture constantly demanded more labour. The "golden age" of the labourer in the second quarter of the eighteenth century apparently engrained in him tastes for manufactured goods which he was willing to work harder thereafter to gratify.[12]

In other words, there seems to have been a shift of the labor supply curve due to a reduced leisure preference. Following this argument, it can indeed be maintained that demand factors mattered insofar as the supply of labor, the demand for leisure, and the demand for goods are simultaneously determined. If there was an increase in the "demand for income," economic growth would occur, but only at the expense of leisure.[13] Possibly this change in attitudes may have been of importance in the initial stages of the Industrial Revolution but it can hardly account for sustained economic growth. It should be added that others have attributed the lengthening of the labor day to different factors such as coercion (for example, Marx and the Hammonds) or improved diets (Freudenberger and Cummins).

The second exogenous source of demand which allegedly was a necessary condition for rapid industrial growth in Britain in the last third of the century—and by implication Belgium, Switzerland, and France in the half century following Waterloo—was exports.[14] It is important, however, to phrase the question sharply; it is not sufficient to state that a "leading" nation had an export sector which pulled the entire economy behind it. In the case of early industrial Europe, if export demand was a main determinant of the patterns of industrialization and growth, it is necessary to show that the successful industrial nations were somehow favored over other potential sellers for whom the expansion of world demand did not lead to rapid industrialization. Moreover, as Kindleberger has pointed out, the

[12] Eric Jones, "Agriculture and Economic Growth: Economic Change," in *Agriculture and the Industrial Revolution* (Oxford, 1974), pp. 116-17.

[13] According to Freudenberger the effective labor input per worker doubled during the second half of the eighteenth century. Cf. Herman Freudenberger, "Das Arbeitsjahr," in Ingomar Bog et al., eds., *Wirtschaftliche und Soziale Strukturen im saekularen Wandel* (Hannover, 1974), pp. 307-20.

[14] Walter E. Minchinton, ed., *The Growth of English Overseas Trade* (London, 1969), ed.'s intro., pp. 36-52. H. J. Habakkuk and Phyllis Deane, "The Take-off in Britain," in Walt Whitman Rostow, ed., *The Economics of Take-off into Sustained Growth* (London, 1962), pp. 77-78.

coincidence between exports and growth is not self-evident. It is necessary to specify in which way foreign expansion brings about growth over and above the simple gains of trade.[15]

The issue of export as a critical component of industrial demand is complicated by the continuous improvements in the productivity of international trade. These had the effect of reducing the transactions and transportation costs, which led to increased international specialization along the lines of comparative advantage. The proportion of national product exported thus rose, and at the same time national income itself grew as the gains from international trade were realized. It could thus be argued that the ultimate causes of growth were changes on the supply side, even though from the point of view of the economy in question the changes were perceived as shifts of the demand curve. More substantial than this taxonomy is the objection that exports may have increased as a result of supply shifts which caused prices to fall, resulting in rapidly growing exports. Unless the shifts in demand and supply curves are separated it is impossible to attribute rising exports to growing foreign demand.

How important was export in the demand for British industrial output? A well known remark in McPherson's *Annals of Commerce* dismisses the importance of exports altogether, estimating the ratio of exports to home consumption at 1:32.[16] This number is obviously too low, but it illustrates the fact that some contemporaries were sceptical of the preponderance of exports in the expansion of industrial demand. The available data support that impression. Schlote's index of the relation between total overseas trade and industrial product displays a striking stability in the period in which export is supposed to have led industrialization. The index which moves slightly above 50 (1913 = 100) in the first half of the eighteenth century, leaps to 65 in the 1760s, but averages only 54 for the rest of the century, and 55 for

[15] Charles P. Kindleberger, *Economic Growth in France and Britain, 1851-1950* (New York, 1964), pp. 264-66. A longer list of possible links between exports and growth is suggested by Richard E. Caves, "Export-led Growth and the New Economic History," in Jagdish N. Bhagwati, ed., *Trade, Balance of Payments, and Growth* (Amsterdam, 1971), pp. 433-37. Some of the mechanisms suggested by Caves work in the wrong direction, however, and other nexuses seem far from easy to test in the case of Europe's industrialization, 1750-1850. For instance, does an expansion in exports stimulate and create ex nihilo entrepreneurial talent and initiative, or does it simply divert those resources away from alternative uses? Do increased exports stimulate the formation of overhead capital (e.g., harbor facilities), or should this be viewed as an additional cost imposed on the export-oriented economy?

[16] Minchinton, *English Overseas Trade*, p. 38. Deane and Cole, *British Economic Growth*, p. 42. T. S. Ashton, *An Economic History of England: The Eighteenth Century* (London, 1972), p. 63.

the period between the Napoleonic Wars and 1850.[17] Eversley stresses the powerful counterexample provided by the events of the 1770s: while exports collapsed, industrial output kept growing.[18] The aggregate data for the eighteenth century may be deficient but data for individual goods display the same features.[19] Even in the case of cotton in Great Britain it appears that the relation between exports and total output is not very pronounced. One crude test of the export-led hypothesis would expect to find a positive correlation between the rate of growth of cotton industry and the proportion of total output exported. But the correlation coefficient between the quinquennial rates of growth of cotton goods production (approximated by raw cotton consumption) and the proportion of total cotton goods exported is not statistically significant for the period 1750-1829. Turning to the nineteenth century we have access to somewhat more aggregated data, which are inconsistent with an "export pull" or staple theory of growth. Total domestic exports as a percentage of national income fell from 18 percent in 1801 to 11 percent in 1841, and then rose to 14 percent in 1851. The corresponding ratios of exports to industrial output are 76 percent, 32 percent, and 42 percent.[20] Had export demand been the one factor that singled out Great Britain as the economy most suitable to industrialization, one should observe that the role of exports increased when industrial growth was fastest, that is, after the Napoleonic Wars. Quite clearly, this version of the Gilboy thesis fails the test.[21]

Furthermore, it is clear that the British Empire, where expansions in demand would benefit British manufacturing more than other countries, was a small proportion of total industrial demand. From the Peace of Paris on, the share of the colonies fluctuates above 35 percent of total exports, although the data for 1793-1814 are not easy to interpret. After Waterloo the British Empire accounted for less than 30 percent of exports. In 1846-1850, total exports to India were

[17] Werner Schlote, *British Overseas Trade from 1700 to the 1930's* (Oxford, 1952), p. 51.

[18] Eversley, "The Home Market," pp. 247-49.

[19] Exports of woolen goods remained virtually unchanged between 1730-39 and 1780-89, while output of woolen cloth milled in West Riding grew by 396 percent (broadcloth, 1735-85) and 97 percent (narrow cloth, 1739-85). Exports of tin grew by 53 percent (1730/39-80/89), while output increased by 86 percent. Between 1750 and 1790 exports of coal rose by 60 percent while coal output more than doubled.

[20] Deane and Cole, *British Economic Growth*, p. 166. Mitchell and Deane, *Abstract*, pp. 282-83.

[21] A similar conclusion has been reached recently by Paul Bairoch, who states flatly that the industrial revolution reached quite an advanced stage in England before seeking foreign outlets. Cf. Paul Bairoch, "Geographical Structure and Trade Balance of European Foreign Trade from 1800 to 1970," *Journal of European Economic History*, 3 (Winter 1974), 569.

only 87 percent of total exports to Germany and only 30 percent higher than exports to the Low Countries. The demand exerted by colonial markets is thus not a very persuasive explanation of why Britain became the "workshop of the world."[22]

The third source of demand cited as important to the industrial revolution is population growth.[23] England's rapid population growth is supposed to have created a large market for industrial commodities, while France's slow population growth is viewed as a cause of retardation. As a matter of economic logic it simply is false that population growth, all other things held equal, will invariably increase the demand for industrial goods. Demand, after all, depends on consumers' income, not merely their numbers. Population growth will increase the number of consumers but decrease income per capita due to diminishing returns. The net effect is indeterminate. It can be shown that in general the following proposition holds: all other things equal, a rise in population will lead to an increase in the demand for any good if and only if the income elasticity of the demand function of that good is less than the reciprocal of the elasticity of non-labor in the production function. The proof of this proposition is straightforward and is deferred to Appendix B. The implications are quite striking. Assume that the economy was sufficiently competitive to approximate the elasticity of non-labor by its share in national income. Deane and Cole estimate the share of labor in national income in 1801 at about .45, which implies that all goods with an income elasticity of 1.8 and higher would experience a decline in demand as a result of population growth.[24] It can also be seen from eq. (6) in Appendix B that if we assume that the income elasticity for agricultural goods was about .5, which implies that the income elasticity for nonagricultural goods was about 1.25, a 1 percent increase in population would lead only to a .31 percent increase in the demand for nonagricultural goods as a whole. Thus population growth *alone* increased demand for nonagricultural goods by 14 percent between 1751 and 1801 and by another 29

[22] The colonial market version of the demand hypothesis is even less convincing for other European economies. Spain, Portugal, and the Netherlands—all late industrializers—had access to substantial colonial markets. Belgium and Switzerland, on the other hand, were confined to the demand structure dictated by the competitive world market, and yet underwent an industrial revolution before 1850. Ireland could have benefitted from the enormous demand exerted by Great Britain and its colonial empire, yet Ireland failed to industrialize.

[23] Deane, *First Industrial Revolution*, p. 34. Brian Murphy, *A History of the British Economy, 1740-1970* (London, 1971), p. 333. Gilboy herself expressed doubts about the importance of population growth by itself.

[24] Deane and Cole, *British Economic Growth*, p. 255. King's estimates for 1688 indicate that the share of labor income was only 39 percent, which would make the critical income elasticity equal to 1.64.

percent between 1801 and 1851. The increment of population increased demand by *less than 10 percent* of total output growth for the first half of the nineteenth century, and probably by even less than that for the period 1751-1801. Even if it is assumed that population growth was fully exogenous, its significance in generating the demand for increased industrial production was marginal.

It thus appears that cost-reducing and factor-increasing changes occupy the center of the stage: supply rules supreme. Technological change, capital accumulation, improvements in organization and attitudes, all made it possible to produce food, clothing, pots, and toys cheaper and better. But the sceptical reader may still be unconvinced. Will these goods be sold? Will it be "possible to find people with income and demand schedules capable of absorbing this increased output?" Is it not true that "the growth of industry has to be explained in terms of its markets, the reasons why increasing quantities of its products could be sold?"[25] Contrary to the first intuition, when considered in a competitive, multiproduct economy these questions are meaningless. As John Stuart Mill put it: "The demand for commodities determines in what particular branch of production the labour and capital shall be employed; it determines the *direction* of the labour, but not the more or less of the labour itself, or of the maintenance or payment of that labour."[26] Nothing has since been added to the body of economic theory that could refute this view—at least as long as the economy is in full employment. After all, the decisions on how, what, and how much to produce are made by the firm. And the individual firm always faces a demand curve which is elastic, irrespective of the demand elasticity for the industry as a whole.[27] An inelastic demand curve facing an *industry* will not affect the impact of a cost-reducing innovation on the rate of growth of the economy. In this case the income effects of the falling prices will simply be siphoned off to other industries.

II. INDUCED TECHNOLOGICAL CHANGE AND ECONOMIES OF SCALE

One interpretation of the Gilboy thesis maintains that the supply curve shifts outward as a result of increases in demand. It is to this issue we now turn.

[25] Eversley, "The Home Market," p. 211. Ralph Davis, *The Rise of the Atlantic Economies* (London, 1973), p. 304.

[26] John Stuart Mill, *Principles of Political Economy*, ed. W. J. Ashley (London, 1929), pp. 79, 87. Emphasis in original.

[27] This is true even in the case of a monopoly in which the industry and the firm are the same. A monopolist will always operate on the elastic segment of the demand curve facing him.

The idea that technological change is demand-induced is far from new, though its full impact on economic theory has not been felt until relatively recently.[28] The precise meaning of demand-induced innovation is more difficult to formalize if we keep in mind that historians have to deal—often simultaneously—with both process and product innovation. In any event, the theory that the intensity of technological change depends on demand factors is different from the Kennedy concept of "induced innovation." The latter shows how market conditions affect the location of an economy on a *given* "innovation possibility frontier." But for the Gilboy thesis in its dynamic version to be true, it is necessary that demand conditions should be capable of causing a shift *of* the frontier outward. This implicitly assumes that there is a "market" for inventions.[29] The demand for technological progress becomes essentially a derived demand, dependent on the demand for the final good. An increase in the demand for the consumption good will generate more inventions precisely in the same way as it will generate increased employment of other inputs. But in the present context that hardly resolves the dilemma first encountered in section I: whence the initial shift in demand?

Moreover, formidable doubts have been expressed concerning the "market" for inventions.[30] If such a market existed at all, it was plagued by at least three sources of market failure, namely the preponderance of externalities, the uncertainty surrounding all stages of innovative activity, and the fact that often new knowledge rapidly becomes a public good.[31] Most damaging, however, to the view that

[28] The seminal article is Charles Kennedy, "Induced Bias in Innovation and the Theory of Distribution," *Economic Journal*, 74 (Sept. 1964), 149-55. For a review of the ensuing debate, see Paul A. David, "Labor Scarcity and the Problem of Technological Practice and Progress in Nineteenth Century America," in *Technical Choice, Innovation, and Economic Growth* (Cambridge, 1975).

[29] The classic work is Jacob Schmookler, *Invention and Economic Growth* (Cambridge, 1966), esp. pp. 88-103, 202-09. Samuel Lilley, "Technological Progress and the Industrial Revolution," in Carlo M. Cipolla, ed., *The Fontana Economic History of Europe* (London, 1973), vol. 3. See also, for example, Landes, *The Unbound Prometheus*, pp. 77, 137. H. J. Habakkuk, "The Historical Experience of the Basic Conditions of Economic Progress," in L. H. Dupriez, ed., *Economic Progress* (Louvain, 1955), pp. 150-51. The debate is admirably summed up in A. E. Musson, ed., *Science, Technology and Economic Growth* (London, 1972), ed.'s intro.

[30] M. Ishaq Nadiri, "Some Approaches to the Theory and Measurement of Total Factor Productivity," *Journal of Economic Literature*, 8 (Dec. 1970), 1148-49 and references cited there.

[31] The list of inventors who failed to capitalize on their major inventions is very long, with Cort, Crompton, Goodyear, Whitney, W. Kelly, Lenoir, and Ericsson being some of the most notorious examples. Schiff has shown that there is little evidence to support the hypothesis that a national patent system makes a significant difference in the rate of industrialization. Cf. Eric Schiff, *Industrialization without National Patents* (Princeton, 1971). For a sceptical evaluation of

the "quantity" of technological progress will respond to shifts on the demand side is that the price on the vertical axis of this supply and demand model is not precisely specified. Since each invention is, by definition, produced only once, the producer has no firm basis on which to estimate his returns. It is often argued, moreover, that financial gain is only a minor consideration for many inventors: scholarly achievement, desire to improve, and pure love of inventing are often cited as prime motives of inventors.[32] This does not imply in itself that inventors are unaware of or indifferent to the economic needs of the society they live in. But a distinct possibility of market failure exists all the same. The material needs of society as reflected by the structure of market prices, and the same needs as viewed by an inventor in search of fame or satisfaction can diverge enormously. Often the full potentialities of major inventions were not recognized initially by the progenitors.[33]

Furthermore, as Rosenberg has recently stressed, it is misleading to assume that the supply of new technology was very elastic so that demand conditions set the "output" of technological change.[34] It is tautological that cost-reducing innovations will be sought after and adopted irrespective of the size of the market; competition forces firms to minimize costs. But is there strong evidence that an increase in demand will result in an increase in inventive activity? Schmooklers results seem to bear this out, but his dependent variable is patenting, not invention nor adoption of an innovation.[35] In fact, there is some evidence indicating increased efforts on technological improvement in times of reduced demand.[36] While this line of rea-

---

the impact of the patent system in Britain, see T. S. Ashton, *The Industrial Revolution, 1760-1830* (New York, 1964), p. 11.

[32] Joseph Rossman, *The Psychology of the Inventor* (Washington, 1931), p. 152.

[33] Edison, for example, "failed to understand the future of his invention [the phonograph] as part of the world of art and entertainment. He thought of it mainly [as a] modern dictaphone, [or] a speaking family album, to preserve the speeches of great statesmen, to teach languages . . . ." See J. R. T. Hughes, *The Vital Few* (London, 1973), p. 175. The divergence between the social and the private rates of return on innovations does not seem to have vanished over time. For a recent study see Edwin Mansfield et al., "Social and Private Rates of Return from Industrial Innovation," *Quarterly Journal of Economics*, 91 (May 1977), 221-40.

[34] Nathan Rosenberg, "Science, Invention and Economic Growth," *Economic Journal*, 84 (March 1974), 90-108.

[35] For critiques of Schmookler's findings see, e.g., Musson, *Science and Industry*, pp. 25-29, and John Jewkes, David Sawers, and Richard Stillerman, *The Sources of Invention* (2nd ed.; New York, 1969), pp. 210-11.

[36] See esp. Nathan Rosenberg, "The Direction of Technological Change: Inducement Mechanisms and Focusing Devices," *Economic Development and Cultural Change*, 18 (Oct. 1969), 23, and sources quoted there. P. Schöller, "La Transformation économique de la Belgique de 1832 à 1844," *Bulletin de l'Institut des Sciences Economiques*, 14 (Dec. 1948), 585. William

soning has its problems as well, and the entire debate is highly speculative, it underlines the weakness of a theory that associates technological change with increasing (or simply "large") demand. Even if the supply of new knowledge were highly responsive to changes in demand, the basic problem of the *primum mobile* remains. If technological change occurred when "demand conditions were ripe" (as Lilley has put it), what changed in this respect in the late eighteenth century? As was shown in section I, most of the traditional sources of demand expansion were of negligible size.

If necessity was indeed the mother of invention, surely the conception was no case of parthenogenesis. A frequently employed notion in this context is that of "bottlenecks" or "challenge and response." To the extent that what is meant here is technology's "responding" to demand, it has been discussed above. But there is a more sophisticated interpretation of this idea in terms of "technological bottlenecks."[37] Some find it useful to distinguish between external bottlenecks which pass through the market mechanism, and internal imbalances which take place at the level of the firm, but this is immaterial for the present purpose. The essence of the argument can be sketched as follows. Suppose a firm or an industry produces a good X by means of two perfectly complementary processes, a and b. The factors employed are L_a and K_a in process a and L_b and K_b in process b. Total initial output is thus given as $X_0 = \min [a\ (L_a, K_a), b\ (L_b, K_b)]$. Now suppose that a technological breakthrough occurs which increases the productivity of both factors in process a by a factor of $1 + \alpha$. Due to the strict complementarity of the two processes, however, output will not rise initially at all; a bottleneck in process b has emerged. Obviously any innovation that will raise the productivity of the factors in b by a factor of $1 + \alpha$ or less will be translated directly in an equiproportional growth in output. One could surmise that in this case the incentive to innovate would be especially high, whether the bottleneck is external or internal. But one should not forget that an alternative solution exists to the "imbalance" created, namely the reallocation of labor and capital from process a to process b. In the present example, simple calculations show the reallocated amount to be $\dfrac{\alpha L_a L_b}{L_a + (1 + \alpha)L_b}$ for labor and a similar quantity for capital.

Brown, "Innovation in the Machine Tool Industry," *Quarterly Journal of Economics*, 71 (Aug. 1957), 406-25.

[37] David, "Labor Scarcity," pp. 82-83. J. R. T. Hughes, "Foreign Trade and Balanced Growth: The Historical Framework," *American Economic Review*, 49 (May 1959), 335-36. Nathan Rosenberg, "The Direction," pp. 4-11.

When will a "bottleneck" result in a technological breakthrough and when will it result in a reallocation of factors? Unfortunately we have no good theory that will predict this. It seems reasonable that the ultimate result will depend on the supply side of technology, that is, whether and at what cost the existing stock of scientific and technical knowledge is capable of solving the bottleneck.[38] In addition, it is likely to depend equally on the cost and time involved in reshuffling the factors of production. As reallocation is, by comparison, a relatively continuous process, whereas inventions are more or less discrete events, it is reasonable to suggest that the more time passes without a technical resolution, the more actual reshuffling of resources will take place, gradually lessening the extraordinary payoff of a technical breakthrough. Many cases can be cited in which apparent "bottlenecks" were solved by reallocation of resources before technological changes altered the required input ratios again. Those "bottlenecks" that were resolved by a spectacular masterstroke such as Eli Whitney's or William Perkin's are likely to receive better coverage in the literature than resource reallocations, which tend to be more gradual and evolutionary. An example of the latter is the production of energy in the second half of the nineteenth century. Output of coal increased at an annual rate of 2.5 percent in Great Britain between 1854 and 1911, and even faster in the U.S. and Germany. The increased demand for energy generated by rapid growth created a bottleneck in coal mining, which was highly labor intensive. Yet there were few major technological breakthroughs in coal mining, especially in England.[39] The winning of coal continued to depend on arduous manual labor under exhausting and perilous conditions. The "bottleneck" was resolved by a reallocation of resources.[40] The proportion of male workers in mining and quarrying in Great Britain rose consistently from 4.2 percent in 1841 to 9.2 per-

[38] A good illustration is the case of steel. It has been argued that Henry Bessemer's attention was first directed to steel when a cast iron cannon was unable to fire a new projectile he devised, and that the Bessemer process was therefore a result of a "technical imbalance." It is noteworthy that a much more costly imbalance existed in constructing railroad tracks from wrought iron, which had to be replaced frequently. Experiments conducted in England indicated that steel rails could outlast wrought iron rails by a factor of seven. Fogel has estimated that the distribution of the wear of wrought iron rails was such that 55 percent of all rails laid wore out in the first ten years. Since the life of equipment was much longer and the life of the area-clearing infinite, the inability to produce cheap steel constituted a severe imbalance—which for the first forty years of the railroad was solved by factor reallocation, not by the invention of cheap steel.

[39] See, e.g., T. K. Derry and T. I. Williams, *A Short History of Technology* (Oxford, 1960), pp. 473-74. A. J. Taylor, "Labor Productivity and Technological Innovation in the British Coal Industry, 1850-1914," *Economic History Review*, 14 (Aug. 1960), 58.

[40] Taylor, "Labor Productivity," pp. 62-64.

cent in 1911. In the U.S. the proportion of miners in the labor force rose from 1.2 percent to 2.8 percent between 1850 and 1910; in Germany the proportion rose from .9 percent to 2.8 percent in the same period. About two thirds of all mine workers were employed in coal or lignite mines. Another example is cotton, often cited as the prime case of an industry in which technological change occurred in "compulsive sequences." It is rather arbitrary to cite the technological solutions of the "bottlenecks" in dyeing, carding, weaving, and spinning, while ignoring that similar problems existed in the planting and picking of cotton at one end of the process, and the tailoring of clothes at the other end. Bottlenecks here were clearly resolved by reallocation, not invention.

Repeating the words "challenge and response" endlessly constitutes no more of a theory of technological change than Oscar Wilde's parrot's "supply and demand" constituted a theory of prices. As long as challenges occur without a forthcoming response, while other major developments occur without discernible stimuli, the "bottleneck theory" ought to be treated with utmost caution. In any event, the "challenge and response" or "compulsive sequence" mechanism is a weak defense of the Gilboy thesis. After all, even if technology did respond positively to the emergence of "bottlenecks," such a pattern of technological progress is essentially a description of the precise operation of the supply side rather than a "link" between supply and demand.

More powerful in the present context is the argument that views the new technology as a shift in cost curves in such a way as to reduce average cost only at a given level of output; that is, the new technique is subject to economies of scale. The Industrial Revolution implied a manifold increase in plant size in manufacturing and transportation. Central power sources, a more sophisticated division of labor, and efficient supervision and discipline imposed on factory workers, brought about an unprecedented increase in firm scale. Hence demand conditions *could* have been important in determining where and when the shift to the new technique occurred, and whether its full potentialities were exploited. What is necessary for this line of reasoning is that the optimal plant size is large with respect to the extent of the market. More precisely, the crucial variable is the ratio of the horizontal distance of the demand curve from the vertical axis at the minimum cost price to the output at which this minimum cost is attained. If this ratio is less than one, the scale economies are not fully realized on account of insufficient demand. Even if the ratio is larger

than unity but still small, insufficient competition may slow down growth by reducing overall efficiency. The question whether demand factors could have operated in this fashion is particularly complicated because Europe was, comparatively speaking, a well integrated economy, so that economies of scale could lead to specialization of large economies in goods in which increasing returns were important. In other words, if country A produced goods under conditions of economies to scale, country B would experience growth too if it could import the goods from A at a cheaper price.

It is not an easy task to substantiate the case for increasing returns in manufacturing anywhere before, say, 1870.[41] Rigorous empirical tests of microeconomic data for early European industry are unavailable. But most econometric studies for American data seem to indicate that the hypothesis of no increasing returns cannot be rejected.[42] There are simpler ways to measure increasing returns than estimating production functions. An early study by G. T. Jones assigned a very modest role to scale economies in Britain for the period 1850-1920, and there is no reason to believe that before 1850 the situation was any different.[43] The alleged existence of economies to scale in the Lancashire cotton industry is criticized in a recent article by Gatrell who concludes that "size in itself guaranteed neither efficiency in good times nor viability in bad . . . one may be most impressed by the ability of small, not to say middle-sized, units to exploit their opportunities in an increasingly competitive industry."[44] A suggestive but not entirely unambiguous procedure is to argue that if there had been industries in which potential scale economies were present but in part

[41] Even today, there is considerable difficulty in interpreting the evidence whether national market size conveys a clear-cut advantage or not, although it is likely that in some industries small nations fail to achieve scale economies or sufficient competition. Cf. F. M. Scherer, *Industrial Market Structure and Economic Performance* (Chicago, 1970), pp. 93-95.

[42] Paul A. David, "Learning by Doing and Tariff Protection: A Reconsideration of the Case of the Ante-bellum United States Cotton Textile Industry," in *Technical Choice*, pp. 142-43. See also Thomas J. Weiss, "Economies of Scale in Nineteenth Century Economic Growth," (Summary of Research Workshop), this JOURNAL, 36 (March 1976), 39-41.

[43] G. T. Jones, *Increasing Returns* (Cambridge, 1933). Jones's results are upper bound estimates of the importance of scale economies since he is unable to separate scale effects from technological change. Still, he finds an "observed elasticity" (i.e., the proportional growth of output divided by the proportional fall in price) of about 5 for the Lancashire cotton industry between 1850 and 1870. This implies a degree of homogeneity of 1.25, assuming a Cobb-Douglass production function. Sandberg has disputed the accuracy of Jones's index and insists that costs in the British cotton industry fell substantially between 1885 and 1914. He attributes this decline in costs to technological change, however, not to scale economies. Cf. Lars G. Sandberg, *Lancashire in Decline* (Columbus, 1974), pp. 93-119, 131-33. See also J. D. Gould, *Economic Growth in History* (London, 1972), pp. 229-35.

[44] V. A. C. Gatrell, "Labour, Power, and the Size of Firms in Lancashire Cotton in the Second Quarter of the Nineteenth Century," *Economic History Review*, 30 (Feb. 1977), 125.

unrealized due to inadequate demand, a process of concentration should have taken place in these industries, possibly resulting eventually in the emergence of "natural" monopolies. There is very little evidence for such phenomena anywhere in Europe before 1850—and certainly none before 1760. One—admittedly rough—indication is simply to look at the number and size of firms active in manufacturing. In 1834, there were 1134 cotton mills in Britain, which rose to 1932 in 1850, stabilizing around 2500 in the 1860s.[45] In Ghent there were 29 cotton spinning firms in 1817 (a crisis year), which rose to 48 firms in 1826 and 78 in 1839.[46] In France, Lévy-Leboyer has observed the proliferation of small firms in the textile industry.[47] In the iron industry, similarly, there is little evidence of concentration; as late as the mid–1850s British iron works were small and dispersed.[48]

It seems unwarranted to rely on scale economies to rescue the Gilboy hypothesis. Scale economies were prominent on the level of the firm, and yet could be relatively insignificant for the economy as a whole if most firms operated on the horizontal segments of their cost curves. The most significant exception to this rule is inland transportation, especially canals and railroads. Here a combination of large fixed costs and non-tradeability would be consistent with the argument that in the absence of a high level of demand there would have been much less reduction of transportation costs. Some gains might have been made on account of increased interfirm specialization, but evidence for such vertical disintegration is not strong. Stigler's rash assertion that England's early start was because "as the largest economy in the world it could carry specialization further than any other country" not only puts the cart before the horses, but has no basis in

[45] A summary of these figures, collected from various Parliamentary Papers, is provided in Mark Blaug, "The Productivity of Capital in the Lancashire Cotton Industry during the Nineteenth Century," *Economic History Review*, 13 (April 1961), 379. The average size of the 1105 cotton manufacturers surveyed in the Horner Report of 1841 was 175 workers, with the median size almost exactly 100 workers. Cf. Gatrell, "Labour," p. 98.

[46] H. Coppejans-Desmedt, "De Statistieken van E. C. Van Der Meersch over de Katoenindustrie in Oost Vlaanderen," *Bulletin de la Commission Royale d'Histoire*, 128 (1962), pp. 121-81. Xavier Heuschling, *Essai sur la statistique générale de la Belgique* (2nd ed.; Brussels, 1841), p. 96n.

[47] Maurice Lévy-Leboyer, *Les banques européenes et l'industrialisation internationale dans la première moitié du XIX^e siècle* (Paris, 1964), pp. 170-71.

[48] See, e.g., A. Birch, *The Economic History of the British Iron and Steel Industry* (London, 1967), p. 205. Duncan Burn, *The Economic History of Steelmaking, 1867-1939* (Cambridge, 1940), pp. 191, 194-95. The number of firms engaged in iron production mentioned in the Coal Commission report was 342, of which 187 engaged in pig iron production. Cf. Shinichiro Kurimoto, "A Statistical Arrangement of the Royal Commission on Coal in 1871," *Quarterly Review, Nara Prefectural College*, 22 (Aug. 1974), 65-93. Kurimoto concludes that "small scale firms had not only overwhelming weights in number but also excellent activity in operating equipments [*sic*]."

the evidence.[49] An alternative approach proposed by Burnet maintains that the supply curve is continuously downward sloping, so that the entire growth process is to be viewed as an explosive disequilibrium process. In this way, he suggests, one could lend theoretical legitimacy to Rostow's take-off.[50] But Burnet's theory does not distinguish sufficiently between historical and Marshallian supply curves. The "historical" supply curve contains technological progress, resource discoveries, and capital accumulation which bring about shifts *of* the Marshallian supply curve. Burnet's bold statement that "the entrepreneur lucky enough to discover a virgin field of consumer demand can look forward to a golden age of self generating growth" is thus largely based on a misconception, although local gains are of course to be expected.

It appears that the preponderance of the idea that economies of scale were somehow crucial in spite of the absence of evidence is caused by a misunderstanding. In many cases an invention was followed by additional, gradual progress which resolved relatively minor but vital bugs in the application of the new technique.[51] These subsequent improvements are often mistaken as increasing returns—shifts *of* the supply curve are likely to be identified erroneously as points lying *on* a given supply curve when they follow a major invention as aftershocks.[52] It is possible, however, that the opportunities for "learning by doing" type of technological progress were far larger in those industries for which demand was comparatively elastic, and that the learning process was chiefly determined by the quantities produced. "Learning by doing," it appears, is the most convincing prima facie nexus between the structure of demand and the rate of growth of the economy. Whether the correlation between demand elasticity and "learning by doing" potentialities actually existed is an empirical issue.[53]

[49] George Stigler, "The Division of Labor Is Limited by the Extent of the Market," rpt. in W. Breit and H. M. Hochman, eds., *Readings in Microeconomics* (New York, 1968), p. 158.

[50] I. D. Burnet, "An Interpretation of Take-off," *Economic Record*, 48 (Sept. 1972), 424-28. Rostow himself is lukewarm about this attempt to formalize his take-off, and notes correctly that Burnet's falling supply curve has to level off somewhere. Cf. Rostow, *How It All Began*, p. 141.

[51] Nathan Rosenberg, "Factors Affecting the Diffusion of Technology," *Explorations in Economic History*, 10 (Fall 1972), esp. pp. 10-14.

[52] An interesting case in point is the debate about the adoption of the reaper in ante-bellum midwest grain farming. See Alan L. Olmstead, "The Mechanization of Reaping and Mowing in American Agriculture, 1833-1870," this JOURNAL, 35 (June 1975), 327-52.

[53] The one case that has been investigated in this respect is the New England cotton industry. Robert Zevin has found that demand was indeed elastic (the elasticity being between -2 and -3 in the 1820s, falling to about -1.5 after 1833). Paul David has found evidence for "learning by doing," although he views the learning more as a function of time than of

We return to economies of scale. As Allyn Young noted, it would be wasteful to make a hammer to drive a single nail. But one is hard pressed to come up with many examples in which the number of nails was so small to obviate the purchase of at least one hammer. And even when this was the case—could nations not borrow each other's hammers? Perhaps Argentina was too small to found a special school for railroad engineers. For that very reason the Argentine railroad system was built by British engineers and the trains hauled by British locomotives. The Belgian textile industry in the 1840s imported its largest machines, while producing most of the smaller machines itself. The average capacity of a domestically produced steam engine in the East Flanders industry was about 50 percent smaller than that of imported engines. In Verviers the capacity of imported engines was more than three times larger than that of domestic ones.[54] The one Belgian machine manufacturer who produced largely for export markets, Cockerill's works in Seraing, made machines with a capacity of 145 horsepower. His competitors, who worked largely for the domestic market, made much smaller machines—an average of 20 horsepower.[55] Today, Belgium is perhaps too small to build its own civilian airplanes—so they buy them in the United States. Luxembourg, which is even smaller, may not find it profitable to operate its own airlines—so the *Luxembourgeois* fly Sabena.

III. A MACROECONOMIC ANALYSIS

An alternative interpretation of the Gilboy thesis maintains that an expansion of demand for industrial goods does not necessarily have to come at the expense of other goods. This would be the case if the preindustrial economy had large amounts of underutilized resources. Indeed, without unemployed reserves of factors, economic expansion initiated by demand would have run into difficulties.[56] If there were large reservoirs of involuntarily unemployed labor, increases in demand would set into motion a multiplier mechanism, which, enforced by induced investment, could have led to the Industrial Revolution.[57]

accumulated output. Cf. Robert B. Zevin, "The Growth of Cotton Textile Production after 1815," in R. W. Fogel and S. L. Engerman, eds., *The Reinterpretation of American Economic History* (New York. 1971), p. 135, and David, "Learning by Doing," p. 167.

[54] Royaume de Belgique, Ministère des Travaux Publics, *Statistique de la Belgique: Mines, usines minéralurgiques, machines à vapeur, 1839-44* (Brussels, 1846), p. 48.

[55] Lévy-Leboyer, *Les banques*, p. 361.

[56] The same point is made by McCloskey, "Did Victorian Britain Fail?" p. 455.

[57] See, for example, Murphy, *A History*, p. 387; Habakkuk, "The Historical Experience," p. 153; Jan de Vries, *The Economy of Europe in an Age of Crisis, 1600-1750* (Cambridge, 1976), pp. 177, 241. The assumption that the early industrial economy was subject to serious Keyne-

This view is reflected in the famous but infelicitous note on the final page of Hicks's *Value and Capital* that "the whole Industrial Revolution of the last two hundred years has been nothing but a vast secular boom, largely induced by the unparalleled rise in population." It is well known that unemployment and pauperism were widespread in preindustrial societies, and Western Europe was no exception.[58] But unemployment is by no means the same as underutilized resources. The latter would only be the case if aggregate demand was insufficient so that people willing and able to work could not find employment. In a preindustrial or early industrial economy a large amount of "natural" or frictional unemployment is to be expected.[59]

Decisive evidence of whether eighteenth- and early nineteenth-century Europe was in fact in something like Hansen's secular stagnation or Joan Robinson's "limping" golden age is not easily obtained. It is a difficult problem to distinguish in a meaningful way between voluntary and involuntary unemployment of employable workers. But there is one body of evidence which requires scrupulous attention in this context, namely the writings of the so-called mercantilist or "pre-Adamite" school. Between William Potter's *The Key to Wealth* (1650) and James Steuart's *An Inquiry into the Principles of Political Economy* (1767), a long list of political economists expressed opinions which all contain an element of Keynes's theory of aggregate demand. It is unmistakable that many of these writers thought that unemployment could be remedied by an exogenous expansion of effective demand. Keynes cited with enthusiasm Bernard de Mandeville's *Fable of the Bees*, and one could provide much more sophisticated and elaborate illustrations from such eminent writers as William Petty, Nicholas Barbon, and George Berkeley. There can be no doubt that from many points of view these writers should be viewed as precursors of Keynes.[60]

---

sian unemployment is made also by J. L. Anderson, "A Measure of the Effect of British Public Finance, 1793-1815," *Economic History Review*, 27 (Nov. 1974), 610-19.

[58] See esp. D. C. Coleman, "Labour in the English Economy of the Seventeenth Century," *Economic History Review*, 8 (Dec. 1955), 280-95.

[59] For a view that attributes "unemployment" to dietary inadequacies, see Herman Freudenberger and Gaylord Cummins, "Health, Work, and Leisure Before the Industrial Revolution," *Explorations in Economic History*, 13 (Jan. 1976), 1-12.

[60] See, for example, William D. Grampp, "The Liberal Elements in English Mercantilism," *Quarterly Journal of Economics*, 66 (Nov. 1952), 465-501. Douglas Vickers, *Studies in the Theory of Money, 1690-1776* (Philadelphia, 1959). T. W. Hutchison, "Berkeley's *Querist* and its Place in the Economic Thought of the Eighteenth Century," *British Journal for the Philosophy of Science*, 4 (May 1953), 52-77. N. G. Pauling, "The Employment Problem in Pre-Classical English Economic Thought," *Economic Record*, 26 (June 1951), 52-65. S. R. Sen, *The Economics of Sir James Steuart* (Cambridge, Mass., 1957).

But even if we accept that these writers thought that the unemployment they were witnessing was of the Keynesian type the issue is not settled. How can we be certain that they actually saw what they thought they saw? Heckscher, Blaug, and Coleman among others have pointed out that there are powerful alternative explanations to the observed unemployment in the pre-industrial economy.[61] One is a high preference for leisure resulting in a backward bending supply curve, which could account for much of the unemployment observed by contemporaries.[62] High seasonal variance in the demand for labor combined with high adjustment and transportation costs led to widespread seasonal unemployment.[63] The importance of this factor appears to have been diminishing with the increasing availability of nonagricultural employment in agrarian communities.[64]

[61] Eli Heckscher, *Mercantilism* (2nd rev. ed.; London, 1955), vol. II, pp. 340-58. Mark Blaug, *Economic Theory in Retrospect* (Homewood, Ill., 1968), p. 15. Coleman, "Labour," p. 289.

[62] I. D. S. Ward, "George Berkeley: Precursor of Keynes or Moral Economist on Underdevelopment," *Journal of Political Economy*, 68 (Feb. 1959), 31-40; Landes, *Unbound Prometheus*, p. 59.

[63] Seasonal unemployment struck agriculture and industry alike, since weather conditions affected water and wind mills, road conditions, bleach fields, and so on. John Law, in his *Money and Trade* (1705), seemed to think that seasonal unemployment could get as high as 50%. A 1752 pamphlet cited by Mantoux places the proportion of time in which journeymen tailors were unemployed at about 40%. (Cf. Paul Mantoux, *The Industrial Revolution in the Eighteenth Century* [rev. ed.; New York and Evanston, Ill., 1961], p. 71n.) In the mid-eighteenth century construction workers were idle at least "four or five months in the year." R. Campbell, *The London Tradesman* (1747), pp. 103-04; rpt. in M. Dorothy George, ed., *English Social Life in the Eighteenth Century* (London, 1923), pp. 32-33. Seasonal unemployment was particularly severe where rural industry was absent and the crops not diversified, particularly in Ireland. As late as 1836 G. C. Lewis noted that two thirds of the Irish work force was not employed all year round, and that this irregularity was the true cause of poverty in Ireland. Cf. G. C. Lewis, *On Local Disturbances in Ireland* (London, 1836), p. 312. Other examples are cited by T. S. Ashton, *An Economic History*, pp. 202-03; idem., *Economic Fluctuations in England, 1700-1800* (Oxford, 1959), p. 6. Wages fluctuated seasonally, so that it is not easy to distinguish between voluntary unemployment (i.e., consumption of leisure) and involuntary seasonal unemployment.

[64] A possible explanation of the widely observed phenomenon of "unemployment" in preindustrial Europe could be built on the idea of structural unemployment. If the marginal productivity of labor was lower than some accepted minimum of subsistence (possibly zero), it follows that people willing to work cannot find employment. For a precise formulation of this idea see R. Eckaus, "The Factor-proportions Problem in Underdeveloped Areas," in A. N. Agarwala and S. P. Singh, *The Economics of Underdevelopment* (London, 1958). Under these circumstances a change in the composition of demand could reduce unemployment if demand shifted toward comparatively more labor-intensive goods, thus increasing the total demand for labor. William Petty's recommendation to employ idle workers to "build a useless pyramid upon Salisbury Plain, bring the stones at Stonehenge to Towerhill or the like" sounds reasonable on this background. It is quite clear, however, that such compositional effects were becoming rapidly less important as the Industrial Revolution proceeded and industrial production was becoming more capital-intensive. The argument could, however, go a long way in explaining the rapid expansion of domestic industry before 1750, which was much more labor-intensive than factory production.

The above is not meant to imply that shocks and fluctuations did not have a profound impact on the level of economic activity. Harvest failures, fluctuations in exports, political upheaval, tariffs, and wars caused widespread distress and unemployment.[65] But as a *long-run* description of a normal state of affairs, involuntary unemployment in preindustrial Europe seems a dubious proposition. One reason is that in a barter economy Keynesian unemployment cannot occur. While Europe and North America were gradually becoming more monetized in the seventeenth and eighteenth centuries, they still contained large pockets of barter exchange. Moreover, in the seventeenth and eighteenth centuries, both in England and in Europe, long-term investment in capital goods was typically financed out of own funds (family loans, retained profits). This implies that investment cannot exceed *ex ante* savings, but as Heckscher pointed out, the "pure" savers who did not plow their savings back into their business, had relatively few alternatives to hoarding. As capital markets improved, the gap between savings and investment narrowed. Heckscher "tests" this hypothesis by looking at secular price movements in the eighteenth and nineteenth centuries.[66] The tests are suggestive but not definitive, since the supply of money is not held constant.

One possible answer to the attempts to search for long-run Keynesian unemployment in pre-modern Europe is provided in the *General Theory*. Keynes's rather cavalier description of the pre-modern economy has it that:

> It is impossible to study the notions to which the mercantilists were led by their actual experiences, without perceiving that there has been a chronic tendency throughout human history for the propensity to save to be stronger than the inducement to invest. The weakness of the inducement to invest has been at all times the key to the economic problem. Today the explanation of the weakness of this inducement may chiefly lie in the extent of existing accumulations; whereas, formerly, risks and hazards of all kinds may have played a larger part.[67]

What Keynes overlooks is that in order to have continuous involuntary unemployment, hoarding has to exceed dishoarding over prolonged periods of time. In other words, a secular upward trend in the demand for money should be discerned. Why would such a trend

[65] See, for example, Joel Mokyr and N. Eugene Savin, "Stagflation in Historical Perspective: The Napoleonic Wars Revisited," in Paul Uselding, ed., *Research in Economic History*, vol. 1 (1976), pp. 198-259.

[66] Heckscher, *Mercantilism*, vol. II, pp. 348-54.

[67] John Maynard Keynes, *The General Theory of Employment, Interest, and Money* (New York, 1936), pp. 347-48.

occur? Keynes's explanation of "risks and hazards" explains the absolute level of the demand, but it is hard to argue that alternative assets were becoming *gradually riskier* than money. But deflationary pressures could result also if the economy was growing or monetizing, however slowly, and the consequent increment in the demand for money exceeded the growth in the money supply. Such deflationary pressures could have produced unemployment if the price level did not adjust sufficiently.

The absence of unemployment or aggregate output data exclude any direct testing of this hypothesis, but some suggestive facts cast doubt on it. First, if it is true that the demand for money for any reason rose faster than its supply, it is reasonable to suppose that interest rates would have been subject to upward pressure. Available data do not support this hypothesis: on the whole, interest rates declined during the seventeenth century while in the eighteenth the downward trend tended to be obscured by wartime borrowing after 1740.[68] Secondly, the evidence indicates that the supply of high-powered money started to grow at an accelerated rate at some point after 1680.[69] Equally important was the emergence of nonmetallic money as a means of exchange. Here one should count as money not only the notes issued by the bank of Stockholm, the Bank of England, and the American colonies, but also bills of exchange which were increasingly made negotiable. The latter, especially, provided a true source of "inside money," making the money supply more responsive

[68] De Vries, *The Economy of Europe*, p. 211; Sidney Homer, *A History of Interest Rates* (New Brunswick, N.J., 1963), pp. 133-43, 155-80.

[69] The best indicator is still total output of the silver and gold mines in Mexico and South America collected by Von Humboldt and published in 1809 in his *Essai politique sur le Royaume de Nouvelle-Espagne*. These figures were refined and completed by Adolf Soetbeer, *Edelmetallproduktion und Wertverhältnis zwischen Gold und Silber seit der Entdeckung Amerikas bis zur Gegenwart* (Gotha, 1879). Soetbeer's figures show a marked acceleration of bullion output after 1680. Moreover, if his figures are corrected in the ways suggested by W. Lexis, the acceleration is even more marked. Cf. Wilhelm Lexis, "Beitrage zur Statistik der Edelmetalle," *Jahrbücher für Nationalökonomie und Statistik*, 32 (1879), pp. 361-417. The results of applying Lexis's critique to Soetbeer's figures are provided below:

Total World Annual Output of Bullion
(in millions of piastres)

Year	*Soetbeer estimate*	*revised Soetbeer estimate*	*Year*	*Soetbeer estimate*	*revised Soetbeer estimate*
1601-20	99.9	81.9	1701-20	99.8	89.5
1621-40	94.0	76.9	1721-40	130.8	120.5
1641-60	90.4	74.1	1741-60	164.6	155.3
1661-80	86.5	71.2	1761-80	175.3	162.4
1681-1700	91.6	76.4	1781-1800	207.8	192.4

to the needs of the economy.[70] Thirdly, if considerable involuntary unemployment had existed in preindustrial and early industrial economies, it should be expected that sudden expansions of aggregate demand should have affected output and employment in a significant way. It seems, however, that in the one test case for which evidence exists, the Napoleonic Wars, this was not the case.[71] It may thus be inferred tentatively that preindustrial and early industrial economies were as a rule *on* their transformation curves. This is not to say that in no sense were there any underutilized resources in preindustrial Europe which could be brought into productive activity during the industrialization process. Obvious examples are a better allocation of resources due to increased efficiency of the market mechanism or the formation of additional factors of production complementary to labor. But such movements should be viewed properly as supply and not as demand related phenomena.

IV. CONCLUSIONS

The intention of this paper has been to examine the Gilboy hypothesis in every possible interpretation, and to decide whether we can assign an important role to demand factors in the explanation of the industrial revolution. Few of the various alternative interpretations withstand the scrutiny of a priori reasoning or empirical tests. The old schoolboy view of the industrial revolution as a "wave of gadgets" may not be so far off the mark after all, provided we allow for "more" as well as for "better" gadgets, and we include abstract improvements such as organizational change, changes in workers' attitudes, and so forth, as "gadgets" in a wider sense.

If demand was not a "factor," what exactly was its place in the industrial revolution? To start with, we observe that any supply shift will affect the economy in direct proportion to the proportional size of the industry affected relative to the economy as a whole. For a once and for all fall in costs the demand structure only matters for determining the composition of the increment in national income but not its size. If the supply shift is a continuous process, however, the shape of the demand curve does matter, because it determines the future

[70] Fernand Braudel and Frank Spooner, "Prices in Europe from 1450 to 1750," in *Cambridge Economic History of Europe*, vol. IV (Cambridge, 1967), pp. 386-87. It is worth noting that Braudel and Spooner claim, albeit without precise evidence, that not only the supply of money rose considerably in the eighteenth century, but the demand for money in fact declined. Cf. ibid., p. 450.

[71] For details, see Mokyr and Savin, "Stagflation in Historical Perspective," pp. 210-23.

pattern of the relative size of the industry in question in the economy. To be more precise: if the own price elasticity is more than unity in absolute value it can be shown that (1) the overall impact of a constant rate of cost reduction will increase over time, and (2) the spillover effects for all other goods taken together are negative. This casts a peculiar light on the role of so-called leading sectors which are supposed to grow due to a very elastic demand curve. Moreover, as output expands, the economy will move down into the inelastic segment of the demand curve. At that point the continuous fall in costs will result in increased demand for all goods (that is, for the good in question and for all other goods taken together), but this impact will slowly peter out over time. Something similar to this process happened in the market for textiles between 1760 and 1860. The elastic demand caused output to grow very rapidly as a response to initial price reductions, so that subsequent price reductions were applied to a much wider base.

Secondly, there is the stability of demand. While the absolute level of demand for industrial goods cannot be used to explain the timing and speed of the industrialization process, heavy fluctuations in demand had an adverse effect on growth, due to the substantial costs of resource reallocation, the acquisition of new information, and so forth. Wars, revolutions, blockades, tariffs, harvest failures, and other unanticipated catastrophes inhibited growth not so much through the first moments of the demand function parameters as through the second moments.

A third way in which demand related factors could have been important in determining the speed and timing of the industrialization process is through their determination of the intersectoral terms of trade. As a consequence of a change in the terms of trade income distribution may change. For example, income may be redistributed from industrial workers and capitalists to landowners. It is, in fact, possible that industrialization could lead to "immiseration" of the modern sector in a way that is analogous to the well-known possibility of "immiseration" of a country increasing its exports. Note, however, that "immiserizing growth" can occur at the level of the economy but not at the level of the entire world. Similarly in a closed economy, it can occur at the level of a sector, but the economy as a whole is better off. Still, if the modern sector has a higher savings propensity or lower risk averseness than the other sector, such a worsening of the terms of trade may affect the rate of growth.[72]

[72] Joel Mokyr, *Industrialization in the Low Countries, 1795-1850* (New Haven, 1976), ch. 7.

To summarize, the traditional notion that supply and demand were somehow symmetric in the industrialization process is unfounded. The determination of "when," "where," and "how fast" are to be sought first and foremost in supply, not demand related processes. Statisticians warn that error can take the form of excessive credulity ("type II") as well as excessive scepticism ("type I"). The wide and uncritical acceptance of the Gilboy thesis is an example of a type II error which has crept into our textbooks and journal articles. It is hoped that the present essay will administer to the demand theory a dose of well-deserved compensating scepticism.

JOEL MOKYR, *Northwestern University*

Appendix A

Assume an economy with the following characteristics: (1) It produces two goods, agricultural goods (A) and nonagricultural goods (M). (2) The two goods are consumed by farmers and nonfarmers. (3) The demand for agricultural goods has a price elasticity of less than one in absolute value for both the farmer's and the nonfarmer's demand functions. (4) Both goods are normal; that is, all income elasticities are positive. (5) The supply of manufactured goods is a function of relative prices, but (6) the supply of agricultural goods depends in the short run only on weather conditions. Proposition: An exogenous decline in agricultural goods due to harvest failures will always result in a fall in the demand for nonagricultural goods.

Proof

There are four different demand functions in this system. Let the subscript 1 denote the demand of farmers for agricultural goods, 2 the demand of nonfarmers for agricultural goods, 3 the demand of farmers for nonagricultural goods, and 4 the demand of nonfarmers for nonagricultural goods. The demand system can thus be written as

$$Q_i(D) = f_i(Y'_j, P_A, P_M) \qquad \begin{matrix} i = 1,2,3,4 \\ j = A \text{ for } i = 1,3 \\ \quad M \text{ for } i = 2,4 \end{matrix} \qquad (1)$$

where Y'_j ($j = A,M$) is nominal income, P_A and P_M the money prices of agricultural and nonagricultural goods, and $Q(D)$ means the quantity demanded. Assuming that all demand functions are homogeneous of degree zero, we divide everything by P_M, thus using nonagricultural goods as our numéraire. Thus

$$Q_i(D) = f_i(Y_j, P, 1) \qquad \begin{matrix} i = 1,2,3,4 \\ j = A \text{ for } i = 1,3 \\ \quad M \text{ for } i = 2,4 \end{matrix} \qquad (2)$$

where Y_j is income in terms of M-goods and $P = P_A/P_M$.

The price elasticities of the functions described by eq. (2) will be denoted by b_i, $i = 1 \ldots 4$. Note that by assumption $-1 < b_1, b_2 < 0$. The *market* demand functions have demand elasticities which are weighted averages of the sectoral elasticities. Let θ_i ($i = A,M$) denote the proportion of good i consumed by the producing sector. Then

$$\delta = \theta_A b_1 + (1 - \theta_A) b_2 \tag{3}$$

is the overall price elasticity for agricultural goods with respect to P, and

$$\beta = (1 - \theta_M) b_3 + \theta_M b_4 \tag{4}$$

is the overall price elasticity for nonagricultural goods with respect to P, and should be interpreted as a cross elasticity. Finally we note that

$$Y_A = PA \tag{5}$$

denotes agricultural income, where A is exogenous, and

$$Y_M = M(P) \tag{6}$$

denotes industrial income. M is a function of P, that is, the supply of M-goods can have a positive price elasticity. This is expressed in γ, the elasticity of M with respect to P which is negative (since P contains P_M in its denominator), since some agricultural goods are used as inputs into manufacturing.

Let us now define $D = Q_3(D) + Q_4(D)$. We have to show that the elasticity of D with respect to A is positive under the assumptions stated. Using the chain rule, we obtain

$$\frac{dD}{dA} = \frac{\partial f_3}{\partial Y_A}\left[\frac{dP}{dA}A + P\right] + \frac{\partial f_3}{\partial P}\frac{dP}{dA} + \frac{\partial f_4}{\partial M}\frac{dM}{dP}\frac{dP}{dA} + \frac{\partial f_4}{\partial P}\frac{dP}{dA} \tag{7}$$

which can be written in terms of elasticities

$$\frac{dD}{dA}\frac{A}{D} = a_3(1 - \theta_M)\left[\frac{1}{\delta} + 1\right] + b_3\frac{1}{\delta}(1 - \theta_M) + a_4\gamma\frac{1}{\delta}\theta_M + b_4\frac{1}{\delta}\theta_M \tag{8}$$

where δ, b_3, and b_4 are as defined in eqs. (3) and (4) and the a's are the corresponding income elasticities.[73]

Assume that the theorem is false, that is, eq. (8) is negative. By multiplying the right-hand side of (8) by δ, it can be seen that (8) will be negative if

$$a_3(1 - \theta_M)[1 + \delta] + b_3(1 - \theta_M) + a_4\theta_M\gamma + b_4\theta_M > 0 \text{ (since } \delta < 0) \tag{9}$$

[73] The parameter δ is not precisely a demand elasticity. Rather, it is the elasticity of an envelope curve of shifting demand curves. An exogenous decline of A causes an initial rise in P as the economy moves up along the demand curve. In a general equilibrium context, however, δ incorporates the secondary effect produced by the decline in M (from eq. 6), causing a leftward shift in the demand for agricultural goods, partially offsetting the rise in P.

From eq. (4) we substitute into (9)

$$a_3(1 - \theta_M)[1 + \delta] + \beta + a_4\theta_M\gamma > 0. \tag{10}$$

The relationship between δ and β can be expressed by using the Cournot aggregation relation.

$$\beta = -(\delta + 1)\frac{PA}{M} \tag{11}$$

Substituting (11) into (10) we obtain

$$(1 + \delta)\left[a_3(1 - \theta_M) - \frac{PA}{M}\right] + a_4\theta_M\gamma > 0 \tag{12}$$

Since γ is negative and $\delta > -1$, this implies

$$a_3 > \frac{PA}{M}\frac{1}{1 - \theta_M} = \frac{PA}{Q_3} \equiv \frac{Y_A}{Q_3} \tag{13}$$

using the definition $I - \theta_M \equiv \frac{Q_3}{M}$.

Eq. (13) can be written as

$$\frac{\partial Q_3}{\partial Y_A}\frac{Y_A}{Q_3} > \frac{Y_A}{Q_3} \tag{14}$$

which implies

$$\frac{\partial Q_3}{\partial Y_A} > 1 \tag{15}$$

Eq. (15) says that the farmers' marginal propensity to consume manufacturing goods exceeds unity. That means that if their incomes increase by one penny, they will increase their spending on manufacturing goods by more than one penny; in other words, their consumption of agricultural goods will be reduced as their income rises. But we assumed that all goods were normal, and hence there is a contradiction. This completes the proof.

Appendix B

Proposition

All other things equal, population growth will increase the demand of those goods the income elasticity of which is *lower* than the reciprocal of the elasticity of non-labor in the production function.

Proof

Let total national income be denoted as

$$Y = f(L) \quad f' > o, f'' < o \tag{1}$$

Let y be income per capita, Y/L, and let the demand for any good by one individual be

$$q(D) = g(y) \tag{2}$$

Total demand is thus

$$Q(D) = Lg(y) \tag{3}$$

Differentiating (3) with respect to L yields

$$\frac{dQ(D)}{dL} = g(y) + L\left(g'(y)\frac{dy}{dL}\right) \tag{4}$$

By using the chain rule and simplifying we obtain

$$\frac{dQ(D)}{dL} = g(y) + g'(y)(f' - y) \tag{5}$$

which can be written in elasticity form

$$\frac{dQ(D)}{dL}\frac{L}{Q(D)} = 1 - g'(y - f')\frac{L}{Q} = 1 - \frac{(y - f')}{y}\frac{g'(y)y}{g} \tag{6}$$

which will be positive if

$$\frac{y}{y - f'} > \frac{g'(y)y}{g} \tag{7}$$

The left-hand side of (7) is the reciprocal of the share of non-labor in national income while the right-hand side is the income elasticity of the good under consideration. This completes the proof.