Explaining the industrial transition: a non-Malthusian perspective

GEORGE GRANTHAM
McGill University, Department of Economics, Montreal, Canada, george.grantham@mcgill.ca

The large-scale structure of world economic history exhibits three steady states punctuated by two phase transitions. The first transition arrived with the domestication of plants and animals; the second with the invention of engines capable of converting thermal to mechanical energy for applications in mining, manufacturing, and transportation. Yet, although both transitions led to increases in the absolute size of the economy, they affected the standard of living differently. Whereas the Industrial Revolution resulted in sustained growth in real per capita income for more than two centuries, over nine millennia the Agricultural Revolution spent itself in population growth that left per capita income insignificantly higher, and possibly lower than the level prevailing under hunting and gathering. This pattern raises three fundamental questions in economic history: why did the first great technological transition produce secular stasis in living standards? Why has the second yielded both steady growth in population and rising living standards? What triggered the transition from the stationary agricultural state to the progressive industrial state?

A Farewell to Alms treats these questions in a novel synthesis of classical economics and historical sociology that explains the Industrial Transition as the indirect consequence of the earlier agricultural state. The principle of diminishing returns and Malthusian demographic dynamics supply...
the classical component. The sociological element derives from Weber’s hypothesis that the advent of modern capitalism was triggered by a religious revolution giving moral force to the ‘bourgeois virtues’ of foresight, patience, prudence, diligence, and rationality. The synthesis proposes that a critical mass of these virtues sufficient to trigger the second economic phase transition emerged by natural selection on individuals operating in an economic state characterized by repeated negative income shocks. On the premise that the bourgeois traits giving superior ‘economic fitness’ to individuals possessing them are heritable, a population of persons economically responsive and open to innovation could be expected to emerge wherever demographic selection remained undisturbed long enough to produce the predicted outcome. The Industrial Revolution originated in England because precocious political centralization and insulation from the Continent’s dynastic and religious conflicts provided the necessary temporal space. The East Asian economies failed to capitalize on their stability because they were not sufficiently Malthusian.

Subject to reservations discussed below this is an elegant argument that explains the apparently universal secular stasis in living standards before the nineteenth century and the odd timing of the upswing in English productivity, which preceded the onset of massive technological change in manufacturing by a half century. It also accounts for the paradoxical rise in agricultural productivity at a time when technological breakthroughs were few and far between. The book thus addresses a critical analytical issue in the economic history of the second phase transition. The evidence points to continuity in the demographic, institutional, and technological context from the high Middle Ages that persisted well into the early phase of the Industrial Revolution. In the face of that continuity, why did productivity begin to rise? Clark dismisses the view shared by Marxists and neo-institutionalists that the upswing reflects the ‘rise of the market’ and the redefinition of property rights in land, goods, and capital in ways that improved the efficiency of private decisions on the grounds that these institutional foundations of market coordination were in place by the high Middle Ages and were not unique to European civilization. Europeans were fully ‘incentivized’ at an early date. Given the constancy of agricultural technology and the slight weight of manufacturing in total output, this leaves by the law of the excluded middle Clark’s preferred explanation that the upswing was due to a cultural revolution that led to greater work effort, greater saving, and more innovation.

Before considering whether the facts support his argument, we need first to review a hypothesis that it asserts but nowhere demonstrates. Evidence that wealthy persons left more surviving heirs than poor people is the critical link that joins the Malthusian constraint to the revolution in values. That wealth was due to possessing the bourgeois virtues would be hard to test in the best of worlds, however, because we have no metric that quantifies those virtues, and consequently no way of knowing whether people who left large
estates had more of the right stuff, more of the right connections, or were just plain lucky. Testing this theory of relative economic success also risks circularity of bourgeois traits, since it is impossible to know in what measure they were possessed by poor people. One cannot define the traits by their association with economic success and then explain success by the traits. The traits themselves may be endogenous to economic circumstances, as evidenced by higher saving rates among small businessmen and lower labor market participation among workers discouraged by the slight probability of finding work at an acceptable wage. That patience, prudence, and diligence ought to yield economic benefit is plausible, and surely desirable; but it is hardly inevitable. Context matters, and some contexts reward virtue more consistently than others.

Let us, however, give the argument from values the benefit of the doubt, since there remains much else that is doubtful. Does the secular stability of real per capita income imply the kind of demographic constraint needed to produce Malthusian selection of fitness traits? Are the economics of pre- and early industrial economies adequately represented by an aggregate production function? Is the timing of the technological inventions that marked the second economic transition better explained by the ‘black box’ of a new species of economic man or by the known history of the posing and resolving of technical problems in particular spheres of intellectual endeavor? These questions lead to the larger question of whether the analytical shortcut proposed by *A Farewell to Alms* reveals a hidden unity in what we already know about the economic development of Europe between 1200 and 1800 or simply suppresses that knowledge in the interest of a quick analytical fix.

### 1. Malthusian demographics

Classical wage theory rests on the propositions that given an inelastic supply of land and static agricultural technology, inputs of labor and capital are subject to diminishing return, and that population endogenously adjusts to changes in the real wage. The first proposition establishes a negative relation between living standards and population density; the second establishes a positive one. Interaction between the productivity and demographic schedules jointly determines the real wage and population density in a steady state to which the economy converges in response to exogenous shocks. A crucial, though usually unstated, assumption is that the economy is closed with respect to migration and trade in foodstuffs.

At first glance, which is all Malthus got, the English price and wage data seem to confirm the homeostatic hypothesis. Real wages rose following the Black Death and declined in the long demographic recovery to the early seventeenth century. Although the precision of the price and population data has vastly improved since Malthus’s time, nothing has altered his
initial impression of a negative correlation between population density and real wages before the nineteenth century. Malthus’s explanation of that pattern, however, was not the only possible one. J. Thorold Rogers, whose compilation of prices and wages was the foundation of the English wage and price series, believed that the ups and downs of ordinary Englishmen’s standard of living documented the downs and ups of the landed aristocracy’s political power, which had weakened in the years following the break-up of manorial serfdom and the fifteenth-century struggle for power, but recovered following the dissolution of the monasteries and the rise of the gentry. Momentarily disrupted during the Civil War, the aristocracy consolidated its domination of the land after the Restoration, sealing its authority over the rural population by means of the Law of Settlement. As with Clark’s ‘revolution of values’, this narrative argument cannot be empirically tested because of the difficulty of defining a metric for political power.

The uncontestable fact is that real per capita income must have been roughly constant over the nine to ten millennia of the agricultural epoch. No other hypothesis is compatible with the level prevailing at the end of the eighteenth century. While that constancy supplies a presumptive argument for Malthusian adjustment in the long run, it does not necessarily imply homeostatic mechanisms strong enough to induce the kind of demographic selection based on strong demographic adjustments to deviations from the long-term subsistence wage. Lee found that the English real wage and population data indicate instead indicate that the response of fertility and mortality to random technological and demographic shocks was slow. Instead, the shocks were initially absorbed by wide swings in the price level and real wage. Convergence to the Malthusian steady state occurred on the time scale of centuries, not generations, which suggests a weak Malthusian constraint at best. Nicolini’s vector autoregression analysis of the post-1540 data yields similar findings. The impulse responses indicate an extremely weak positive check that disappeared around 1640, a slightly stronger preventive that vanishes after 1740, and no real wage response. The best case one can make for the Malthusian hypothesis is ‘not proven’. These findings have broader implications for the analysis of the economy as a whole, since they throw doubt on the usefulness of the classical representation of the economy as a stylized agricultural production function.

Analysis of price and population data from other European countries casts further doubt on the Malthusian hypothesis. On the assumption that the supply of land was fixed, one would expect the elasticity of the real wage

---

1 Thorold Rogers 1891. For a larger review, see Kadish 1989, pp. 22–35.
2 A growth rate of 0.02 percent projected backward 2,000 years would reduce that income by one-third, which, given consumption levels in 1800, would seem to be the minimum level compatible with survival. On food consumption levels, see Fogel 2004.
3 Lee 1993, pp. 1–30.
with respect to population density to be lower in countries where the land-labor ratio is relatively low. The data from early modern Europe indicate just the opposite. In presumably land-scarce Italy and England, a 1 percent rise in population was associated with a 1.0 and 1.2 percent declines in the real wage. For more land-abundant Germany and Austria, the elasticities are \(-2.4\) and \(-2.5\), and for France and Spain, \(-2.2\) and \(-2.4\). Only Poland’s insignificant coefficient of \(-0.9\) supports the Malthusian view.\(^5\) Given the problematic nature of the population data and difficulty of interpreting as national prices affected by international trade, one should not put too much faith in these numbers. They nevertheless fail to support a major leg of the Malthusian argument. The same is true of recent econometric analysis of the French population series, which do not find the predicted strong fertility and mortality responses to price shocks.\(^6\) With few exceptions, the great early modern famines were attributable to the inelasticity of the distribution system in the face of moderate to severe, but hardly catastrophic harvest shortfalls.\(^7\) Probably the greatest post-plague mortality was manmade. Between 1618 and 1648 the warring armies equal to medium-size cities that foraged the countryside and conducted scorched-earth campaigns against each other may have carried off 30 to 40 percent of the population of Central Europe. The war zones of Mecklenburg, Pomerania, and Württemberg lost over half their mainly rural inhabitants.\(^8\) One does not need to appeal to Malthus to account for the long-term demographic stasis.

The data are too noisy to confirm or reject Clark’s hypothesis. English population is plausibly projected to 1540, beyond which estimates rest on conjectures derived from early sixteenth-century militia rolls and problematic extrapolations from the 1380 poll tax returns. Clark has attempted to correct that deficiency by using a production function to infer population density from the real wage, but any such estimate is for testing the hypothesis on which they rest.\(^9\) Nominal wage data are notorious for their medium-term inflexibility, and although Clark fleshes out the conventional series with threshers’ piece and day rates, the real wage index remains dependent on movements in the general price level unrelated to the Malthusian dynamics of a closed economy. The response of nominal wages to changes in prices was extremely slow, because apart from the great debasement of the 1540s and 1550s and the first decade of the nineteenth century, changes in price to which rational workers might respond were unnoticeable. From 1200 to 1840, the rate of inflation was below one percent 87 percent of the time, and below half a percent 67 percent of the time.\(^10\)

\(^5\) Lee 1987, p. 450.
\(^6\) Chevet 1993; Weir 1984.
\(^7\) Grantham 1997.
\(^8\) Parker 1997, pp. 188–9.
\(^9\) Clark 2007.
\(^10\) Calculated from data published in Clark 2007.
In the face of that glacial change it is hardly surprising that nominal wages stayed constant for long periods of time and only moved in response to triggering shocks. The single exception occurred in the 1350s, when both wages and prices accelerated sharply in response to drastic increase in the stocks of money and capital relative to population.\textsuperscript{11} The primary evidence for the Malthusian mechanism thus comes down to the single episode of the Black Death and its successive visitations in the second half of the fourteenth century. Real wages rose with the decline in population, and fell when population recovered. Short of triggering economic implosion of the kind that destroyed the Native American economy in the sixteenth century, it is hard to envision any other outcome of demographic catastrophe.

\section{The production function}

Clark’s analytical tool of choice is a neoclassical aggregate production function characterized by an inelastic land input, which he deploys to interpret the indices of real wages and rent as the marginal productivity of labor and land for the economy as a whole. While this time-worn technique falls well within the disciplinary boundary of unified growth theory, one may question its appropriateness for pre-modern economies in which the required degree of economic integration is problematic. A high degree of economic integration is a necessary condition for treating empirical proxies for that function as continuously binding constraints. That is something that has to be demonstrated, not merely assumed on the grounds that everyone does it.

Traditional agriculture supplies a test of this approach because we are well informed about its technology and have enough information to track the methods of cultivation and productivity to the early thirteenth century. The conventions of historical exposition represent medieval and early modern farming in terms of ideal types that epitomize its organization and technology in different eras. Clark correctly observes that the technology of English farming in the thirteenth century was essentially the same as it was in at the end of the eighteenth century, when productivity was roughly twice as high.\textsuperscript{12} He incorrectly infers from this observation that in the absence of a cultural revolution that encouraged more effort to innovate, that technological stasis imposed strict limits on agricultural productivity and economic growth, in large part because of a nitrogen constraint on the level of crop yields. There is reason to doubt the effectiveness of that constraint in limiting potential agricultural output.

The doubts arise from the wide range of agricultural productivity observed in pre-modern farming. By the late thirteenth century, grain yields in

\begin{footnotesize}
\begin{enumerate}
\item Findlay and Lundahl 2002.
\item The conventional estimates are summarized in Wrigley 2006.
\end{enumerate}
\end{footnotesize}
Flanders and the adjacent districts of northern France had attained the maximum levels obtainable using natural fertilizer in conjunction with traditional varieties of cereals on soils by no means superior.\textsuperscript{13} Yields on the large farms that provisioned Paris from the plains of the Île de France were as high as yields on comparable operations in eighteenth-century England, and in England itself, as Campbell and Overton have shown for Norfolk, they were higher in the early fourteenth century than they would be for next three centuries.\textsuperscript{14} Although these performances were achieved at the cost of high inputs of labor and capital, they were not achieved at the expense of diminishing returns to labor, as Karakacili’s painstaking reconstruction of labor inputs on a half dozen Ramsay manors has demonstrated. Her estimates indicate a level of labor productivity equal to the better managed English farms around 1800.\textsuperscript{15} While these findings, of which the examples can be multiplied, are consistent with the hypothesis of fundamental stasis in the technology of traditional husbandry before the early nineteenth century, they are not consistent with the hypothesis of a tight subsistence constraint. The latent productivity of traditional husbandry was significantly higher than average productivity, which was the weighted average of the productivity of farms exposed to different economic contexts.

Clark’s hypothesis implies that these differences are due to cultural factors selected by Malthusian demographic forces. But there is a simpler testable explanation. The high-yield districts were exposed to stronger market opportunities, which induced the effort and investments he correctly supposes underpinned productivity growth prior to the massive technological breakthroughs of the 1830s and 1840s. Most of these districts were exposed to strong urban demand, which apart from the episodic appearance of armies constituted the only concentration of buying power sufficient to overcome the limitations on the extent of the market imposed by high transportation and storage costs.\textsuperscript{16} The effect of demand on productivity, however, goes beyond the relatively bulky cereals and fodder crops. In the century following the Black Death the average weight of the eminently tradable English fleece declined by 34 percent.\textsuperscript{17} It is implausible that in an age when the price of wool was increasing relative to grain cereals that English farmers forgot how to breed productive sheep. A more plausible explanation is the collapse of foreign demand caused by the late fourteenth-century Flemish

\textsuperscript{13} Derville 1987.
\textsuperscript{14} Campbell and Overton 1993, pp. 38–105.
\textsuperscript{15} Karakacili 2004.
\textsuperscript{16} The role of military demand in the development of markets for subsistence foodstuffs in the seventeenth and eighteenth century is a much overlooked topic. By the 1630s armies operating in central Europe regularly exceeded 40,000, and by the 1670s and 1680s were exceeding 100,000 men plus camp followers, by which time, unlike the earlier period, they were regularly supplied by quartermasters.
\textsuperscript{17} Stephenson 1988.
urban implosion. Wool-growing was relatively more profitable that cereal farming, but it was not so profitable relative to the cost of production to induce the investments that had sustained earlier high productivity. A similar phenomenon can be seen in the reduction in the size of oxen at the end of the Roman era that was not reversed before the early industrial age. These and other examples were consequences of specialization, which affected the productivity not only of biological matériel, but physical and human capital as well. What Clark assumes to be the agricultural production function is in fact conditional on the state of the market.

Much of the book is given over to dismissing the ‘institutional’ hypothesis. As with stasis in technology, European systems of property rights were to use the author’s neologism, adequately ‘incentivized’. There is no reason to believe that in regions and periods undisturbed by international or domestic violence property rights in most parts of Europe were so defective as to significantly impede private response to opportunity. Opportunities, however, were not uniform. Clark dismisses a simple version of the multiple equilibrium explanation of economic phase shifts drawn from recent work in unified growth theory that focuses on demographic behavior. The more relevant economic source of multiple equilibria, however, lies in the phenomena of increasing returns in commerce, finance, transportation, and handicraft manufacture first described and analyzed by Adam Smith.

The effects of high-level equilibria were spatially constricted by high transportation costs, which is why with the exception of sixteenth-and seventeenth-century Netherlands, they do not show up in noisy estimates of aggregate productivity at the national level. Clark is right in arguing that much of the institutional framework evolved endogenously; he is wrong in supposing that only a putatively exogenous event like a change in morals could have shifted the system’s economic trajectory.

3. Technological innovation

Clark’s treatment of technological innovation in the Industrial Revolution is thin, and focuses primarily on its limited early impact on aggregate productivity as an element in the argument that it does not explain the timing of the initial upswing in output and productivity. While it is indisputable that the impact of the breakthroughs in ferrous metallurgy, mechanical preparation of cotton yarn, and the harnessing of heat energy for mechanical work on aggregate total factor productivity were initially limited, the sustained and accelerated surge in factor productivity to the middle years of the nineteenth century cannot be explained in their absence.

---

18 Chaix and Desse 1994; Audoin-Rouzeau 1995; Grant 1988.
19 Becker, Murphy and Tamura 1990.
20 For an intuitive discussion, see Grantham 1999.
Moreover, it is impossible to conflate the series of inventions that opened up these and other technologies to a datable change in culture that Clark believes explains the early upswing in output and productivity. The critical breakthroughs were in the sphere of mechanics, of which the steam engine with its multiple applications in industry and transportation, and semi-automated machine tools were the most important in terms of the long-run consequences for labor productivity. The history of the development of these technologies owes little to a ‘revolution in values’ and the inventions were far from random. This is particularly true of steam power, where the critical breakthroughs involved early seventeenth-century scientific analyses of the atmosphere demonstrating the phenomenon of air pressure and the distinction between steam and air, which opened the door to exploiting the latent energy of a vacuum created by condensing steam in a closed vessel.\textsuperscript{21} Exploitation of this discovery by means of a ‘steam engine’ was stimulated by concentrated demand for power exceeding the amount deliverable by waterwheels, windmills, and animals working against capstans. The original loci of that demand were waterworks for cities inadequately served by aqueducts and waterlogged mines, both of which engaged huge expenditures of energy needed to pump the required volumes of water.

It is difficult to see how a revolution in values or even in literacy and numeracy could have affected the timing of the successive posing and solving of the technical problems associated with the invention of this technology, which depended on the scientific discoveries in the first half of the seventeenth century and the expansion of cities and coal mining in the second half. The same is true of machine tools, though apart from the mathematics needed to calculate forces and angles, the scientific input was initially smaller. Much of the early critical work here was conducted in France in an effort to device machines to cut the high-precision groove for fusée chains in watches and clocks, which could also be applied to cut accurate screws in soft metal. The transfer of that technology to all-metal lathes capable of machining iron and steel pieces to high tolerances was a response to the opportunities opened up by the breakthroughs in smelting and refining in the second half of the eighteenth century. Work in this direction was equally intense and competent in France and England. In the case of the all-important turbine, the critical work in the eighteenth and early nineteenth centuries was carried out on the Continent. Here, the mathematics needed to analyze the complex aspects of fluid mechanics involved in designing efficient reaction machines became available only in the early eighteenth century.

The particular histories of critical technologies that powered the second economic transition owe little to contemporary evolution in values or attitudes that were peculiar or at least prominent in seventeenth- and eighteenth-century England. It is interesting and significant that the critical

\textsuperscript{21} Usher 1954.
breakthroughs in mechanical and power-generating technologies occurred in Europe rather than in East or South Asia, and it is clearly an important task of economic and technological history to track down the reasons why. It is also important – and here the chances of success are presently higher – to uncover the reasons why so many technological strands affecting productivity seemed to come together between 1650 and 1800. The level of specialization and scale of production that raised general mechanical proficiency in the metal-working trades and produced bottlenecks at points accessible to solution are products of the commercial and scientific developments between 1400 and 1750 common to western Europe. A counterfactual history in which the Thirty Years War was aborted probably would have resulted in the Industrial Revolution occurring in the Low Countries and the Rhineland, where the core of the European industrial economy emerged in the course of the nineteenth century. The cultural values that Clark argues were the source of English economic success were too widespread to explain the location of the early breakthroughs in organization and technology that marked the first phase of modern economic growth. Cultural explanations are quick fixes, and cannot substitute for the hard detailed rendering of the strands of causation that led to the second great economic transition.

References


Greg Clark is a master of the art of using one-liners in telling stories and *Farewell to Alms: A Brief Economic History of the World* is no exception. It offers the Malthusian hypothesis of population growth leading to misery as an all-purpose vehicle for all human history, except for the last 200 years. However,