CHAPTER

1

Understanding growth in Europe, 1700–1870: theory and evidence

Joel Mokyr and Hans-Joachim Voth

Contents

Incomes of ordinary citizens in developed countries today dwarf those enjoyed even by the wealthy elite during most of mankind's history. As John Maynard Keynes, with slight incredulity, observed in 1930 that the economic problem of mankind (in Europe and North America at least) has been solved (Keynes, 1930). People no longer go hungry. Clean clothes, shelter, and warmth have gone from luxuries to necessities. By 1870, developments that would eventually deliver this full complement of riches were already in full swing. This chapter summarizes recent research by growth economists on how mankind escaped from a life that was, in the words of Thomas Hobbes, "nasty, brutish, and short." It contrasts these interpretations with the existing historical evidence and recent findings of economic historians. Four areas are of particular concern – demography, institutions, human capital, and technology. We conclude with suggestions for future research.

Theoretical approaches

In the late 1980s and early 1990s, macroeconomists began to turn their attention from business cycles to the determinants of long-run economic growth. Papers in the endogenous growth literature sought to explain why some countries had grown more rapidly than others. The main period of interest to which these models were applied was the post-war era. They returned to Kuznets's classic argument that current growth rates, when extrapolated backward, implied absurdly low incomes in early modern times and before. Therefore there must have been a long period of stagnation before modern growth started. But what was the source of the phase transition from a world of very low or zero rates of growth to a modern world of rapid and sustained growth?

From the 1990s onwards, scholars started to search for an overarching theory that could encompass both slow growth and the transition to rapidly increasing per capita incomes – a "unified growth model." The field has flourished since. A number of themes stand out – demography, the influence of institutions, human capital and culture, and the role of technology. We first summarize some of the most prominent contributions in the theoretical literature. In the main part of the chapter, we compare the theorists' predictions with the main facts unearthed by economic historians. Our conclusion offers some suggestions on how progress can be made.

Early models in unified growth theory, such as Kremer's (1993) paper, modelled the transition from stagnation to growth as one long, gradual acceleration of growth rates. As in some other papers in the endogenous growth literature, Kremer's model assumes that more people spells faster technological change, since the probability of a person having a bright idea is more or less constant. Because ideas are non-rivalrous, growth accelerates. Kremer showed that some of the basic predictions derived from such a simple growth model hold both over time and in cross-sections. Since 1,000,000 B.C., growth rates of population can be predicted from the current size of the population. Also, geographically separated economic units with greater surface areas produced bigger populations and higher densities. As population size and technology increase jointly, there is no steady state in Kremer's model. To avoid all variables showing explosive behavior, a demographic transition is necessary, so that fertility responds negatively to higher incomes above some threshold level.

In contrast, in exogenous growth models, technology "just happens," and adoption decisions are not explicit. Size itself does not affect technology or productivity change. In one application of exogenous growth to the transition to self-sustaining growth, Hansen and Prescott (2002) model the transition "from Malthus to Solow" by assuming that technological change in both the land-using (diminishing returns) and the non-land-using modes of production is exogenously given and constant. Initially, only the Malthus technology is used. In every generation, each lasting thirty-five years, productivity in their model increases by 3.2 percent in the "Malthus sector" (i.e. agriculture, where labor is subject to declining marginal returns) and by 52 percent in the "Solow sector" (where all factors of production are reproducible). Eventually, as the productivity of the unused technology increases exponentially, the Solow technology becomes competitive and is adopted. In this setup, an Industrial Revolution is inevitable, and does not depend on anything other than the differential growth rates of productivity used in the calibration.

A second class of models in which size matters also takes technological change to be exogenous. Here, the focus is on the conditions under which new techniques will be adopted. Early models in the tradition of Murphy, Shleifer and Vishny (1989) relied on demand effects, and hence the size of economies, to explain when a "big push" might occur. By "big push," authors in the tradition of Rosenstein-Rodan mean the simultaneous adoption of advanced technologies in many sectors. In order to pay the fixed cost necessary for adopting modern production, demand needs to be sufficiently high. This will often be the case only if a whole range of industries industrializes. The chances of this occurring increase with total output. One implication of these models is that industrialization might have been feasible long before it got under way – if only everybody had decided to invest earlier in fixed-cost technology, profits would have been high enough to justify the expense. Advances in technological knowledge themselves need not translate into greater output. Coordination failure can thus undermine the transition to modern technology.

High fixed costs and indivisibility also play a crucial role in models that put risk diversification at the heart of adoption decisions. Acemoglu and Zilibotti (1997) present a model with a tension between production requirements and household investment. Productive projects using new technology require substantial set-up costs. At the same time, households want to diversify their investments to minimize risks. Because of this, investment in the new, productive technology is initially very low, and so is output. This changes as households become richer – their savings become sufficiently large, relative to the capital requirements of new technologies, to avoid "putting all their eggs in one basket." Industrialization, once under way, generates the means with which to sustain itself. A number of lucky draws can get it started. Two identical economies may end up on very different paths, depending on whether they get lucky in the first round or not. Acemoglu's and Zilibotti's model also has the feature that households do not take into account the effect of their investment decisions on aggregate productivity. Industrialization may not occur, while being feasible. The model incorporates a stochastic component – industrialization may partly be the result of chance. One implication is that not every aspect of actual industrial transformations is fraught with meaning – and the country that actually went first may simply have been lucky.[1]

Many unified growth models link human capital accumulation with technology and the ideas-producing properties of population growth. These papers have argued that the transition to modern growth is accompanied by a growing importance of human capital (Becker and Barro, 1988; Lucas, 2002; Becker, Murphy, and Tamura, 1990). Galor and Weil (2000) made the nexus between human capital and technological change a cornerstone of the transition to rapid growth. They argue that the escape from stagnation took place in two steps – a transition from the Malthusian to a post-Malthusian state, and then to a modern-growth regime. Galor and Weil's key assumption is that, as technological change accelerates, human capital becomes more valuable: it allows people to cope with a rapidly changing workplace. Technological change accelerates as more people produce more ideas during the long Malthusian period. Because of a delay in the response of population to income growth, per capita incomes grow, if very slowly. Eventually, parents invest more in the human capital of their offspring. This in turn accelerates the growth of knowledge. Higher incomes make it easier for parents to have more children. At the same time, a growing value of human capital produces incentives to increase the quality of one's offspring, reducing quantity. Initially, after the start of

[1] Following Crafts's (1977) original contribution, this idea has been the subject of substantial debate among economic historians.

modern growth, the income effect dominated, leading to more births; later, the substitution effect became more important, and fertility declined.

Cervellati and Sunde (2005) as well as de la Croix (2008) alter this setup by arguing that life expectancy rose quickly with productivity. This in turn encouraged investment in human capital, as payback horizons lengthened. Even if technological change is only slightly skill-biased, a self-reinforcing cycle of better technology, greater life expectancy, and higher investment in human capital can get started. Boucekkine, de la Croix and Peeters (2007) show how rising population density may encourage higher literacy, through the cheaper provision of schooling services. Jones (2001) combines the population-ideas mechanism with a property rights regime that reserves a share of output for innovators. Based on his calibrations, Jones concludes that the single most important factor leading to a take-off in growth after the nineteenth century was more effective enforcement of intellectual property rights, which created the necessary incentives for the sector that produced the ideas.

Some observations from economic history

The population–idea nexus is key in many unified growth models. How does this square with the historical record? As Crafts (1995) has pointed out, the implications for the cross-section of growth in Europe and around the world are simply not borne out by the facts – bigger countries did not grow faster.[2] Modern data reinforce this conclusion: country size is either negatively related to GDP per capita, or has no effect at all. The negative finding seems plausible, as one of the most reliable correlates of economic growth, the rule of law (Hansson and Olsson, 2006), declines with country size. Even if we substitute "population" with more relevant concepts like market size, which might have influenced the demand for new products, the contrasting growth records of Britain and France are hard to square with endogenous growth models emphasizing size.[3] Moreover, it is disconcerting for these models that in 1750, on the eve of the Industrial Revolution, Britain had just experienced half a century of virtual demographic stagnation. One could also point out that if population size

[2] It is indeed striking that prior to the coming to the fore of the British economy, Europe's most successful economies tended to be city states (Hicks, 1969, p. 42). These, with high density but relatively small populations, had an advantage in solving the problems of setting up effective institutions of commerce and finance. Market size was less of a problem, in part because the fixed costs of setting up these institutions were not all that high, and because they tended to be open economies. The main source of economies of scale was not economic but military. Military power depended on total income and population.

[3] Some later models in the spirit of Kremer, such as Jones (2001), attempt to provide a solution to this problem by assuming increasing returns in the production of goods, and by allowing the number of new ideas to be a function of the existing stock of ideas.

is critical, China's early modern record is a puzzle. Its population rose from 130 million in 1650 to 420 million in 1850, yet no Industrial Revolution occurred. An interesting argument is made by Lin (1995). Lin argues that the relationship between population size and technological change depends on the source of innovation. In a world in which new technology is based entirely on learning by doing, greater size would imply more innovation, assuming that the advances were disseminated effectively over the larger unit. Once progress begins to depend more on experimentation and theory, such advantages disappear. Lin maintains that the success of China in the Song period (960–1279), as opposed to its relative stagnation in the seventeenth century and beyond, reflects a change in the source of innovation.

Even if "size mattered" in the data, it would not be clear what the relevant channel was. A larger population (without a collapse in per capita incomes) may be accompanied by positive externalities of a different kind. Regardless of whether size mattered to the generation or adoption of new technology, as the endogenous growth models suggest, greater size could simply have enhanced the division of labor. This in itself could have contributed to an acceleration of output growth. Kelly (1997) presents a model of "Smithian growth," where trade integration is furthered by improvements in transport infrastructure, leading to an acceleration of growth. He applies this model to Song-dynasty China. Similarly, in Europe, higher population density may have generated the scope for positive externalities, partly through improvements in turnpikes and canals, partly through long-distance trade (Bogart 2005a, 2005b; Daudin 2007). In this sense, it becomes easier to rationalize the commercial successes of the medium-sized, but densely populated and internationally integrated Dutch Republic in the seventeenth and eighteenth centuries.

Models in the "big push" tradition run into problems similar to population-based endogenous growth; the European experience after 1700 does not suggest that the absolute size of economies is a good predictor of the timing of industrialization. The size of most industrialization projects was small – even the largest textile mills, had they been financed by a single person, hardly constituted a large concentration of risk. Before the late nineteenth century, fixed costs in manufacturing were limited. Much diversification, moreover, could take place *within* the existing business structure of Britain during the Industrial Revolution.[4] When it comes to production technology with high fixed costs, adoption decisions *after* 1870 could possibly be explained by the

[4] Pearson and Richardson (2001) show that the typical entrepreneur in the Industrial Revolution was heavily diversified. Rather than describing the entrepreneur as a single-minded owner-manager who spent his entire life on the one business, they show the extent to which early entrepreneurs were involved in non-core ventures. Cotton masters and other textile producers in Manchester, Leeds, and Liverpool, for example, could be found as directors of insurance companies, canal and turnpike companies, gas companies, banks, and companies in other sectors.

big-push framework. Yet by that point in time, international trade was already doing much to break down the link between the size of the domestic economy and the possibility of technology adoption. If there were large fixed costs before 1870 they were in infrastructure, not in manufacturing. In Britain, these infrastructure investments – canals, turnpikes, harbors – do not appear to have suffered a great deal from capital scarcity. This is despite the numerous shortcomings of the British financial system, which ranged from the Bubble Act to usury laws that squeezed private credit, and the relentless borrowing by the Crown for much of the eighteenth century (Temin and Voth, 2008). On the whole, infrastructure projects were apparently financed without too much difficulty, mainly through local notables (Michie, 2000).

Finally, unified growth models that emphasize differences in productivity growth between the agricultural ("traditional") and industrial ("modern") sectors, such as Hansen and Prescott (2002), also encounter substantial empirical difficulties. At the point in time when overall growth rates began to accelerate, both the land-using sector as well as the industrial sector became more productive – according to some measures, at relatively similar rates (Crafts, 1985a). By definition, the Hansen – Prescott model has little to say about which country industrialized first, and why – the entire world is its unit of observation.

These observations are not meant as final verdicts on the merits or otherwise of unified growth models. They explain why we believe that theorists, applied economists, and economic historians should dig deeper – especially into the interactions between fertility, human capital, institutions, and technology. This is what the following sections attempt to do.

Malthus vanishing

Populations grew in most parts of Europe during the early modern period. In some parts, they surpassed the levels seen before the Black Death. Demographic growth accelerated decisively in many European countries in the late eighteenth century. There was substantial variation in timing, with Britain and Ireland leading the way, and France avoiding a major jump altogether. During the period 1500–1870, the economic impact of demographic factors changed. It went from being a crucial determinant of per capita incomes in most parts of Europe to a factor of declining importance as technological change accelerated after 1800. Growth theorists often refer to the period before 1750 as the Malthusian epoch. We first describe the Malthusian model and key changes in demographic–economic interactions after 1800. We then review the evidence and summarize what we know about how population pressure eventually fell away as a key economic variable.

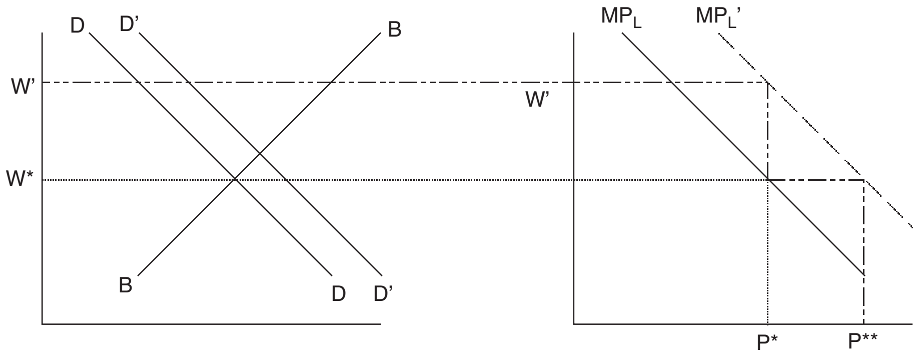

Figure 1.1 The Malthusian model

The Malthusian model relies on two main assumptions: first, that population growth responded positively to per capita incomes. As wages or per capita income fell, fertility declined (the "preventive check") and death rates increased (the "positive check"), as indicated by the upward sloping fertility schedule BB, and the downward sloping mortality schedule DD, in Figure 1.1. The second assumption is that income per capita was negatively related to population size due to diminishing returns to labor, illustrated by the downward sloping marginal product of labor curve, MP_L, whose position reflects *inter alia* the level of technology in the economy. A widely cited example illustrating the trade-off between incomes and population size is the Black Death. As European populations fell by approximately one-third to one-half of their pre-crisis levels, wages everywhere surged. Living standards in fourteenth-century England, conventionally measured, reached a high not seen again until the nineteenth century.

Together, the two assumptions underlying the Malthusian model imply that whatever advances in incomes occur will inevitably be frittered away through more babies. In Figure 1.1, birth and death schedules intersect at a wage W*. The technology schedule in the right-hand panel then translates this into a feasible population size P*. If a temporary technological shock moves the MP_L to the right, to MP_L', and thus drives the wage up to W', death rates fall and population starts to grow. Eventually, because of declining marginal returns, this will force wages down to their previous level (at population level P**). As H. G. Wells put it, mankind "spent the great gifts of science as rapidly as it got them in a mere insensate multiplication of the common life" (Wells, 2005).[5]

[5] Galor and Weil (2000) assume that the response of fertility to incomes is delayed. Hence a one-period acceleration in technological change can generate higher incomes in the subsequent period, and a sequence of positive shocks can lead to sustained growth. While this solves the problem in a technical sense, it is unlikely to explain why fertility responses did not erode real wage gains over hundreds of years.

Clark (2007a) even goes as far as to argue that the English population in 1800 were no better off than their ancestors on the African plains millennia before.

Higher death rates (depicted by a rightward shift of the mortality schedule from DD to D'D') imply higher per capita living standards. Unhygienic conditions and a deterioration in the microbial environment, for example, will boost incomes as they reduce the number of surviving children. Lower fertility rates can achieve the same effect. Welfare is not necessarily any higher, but the incomes of those who live will be. Europeans in the early modern period also reduced population pressure by ensuring that a high proportion of women never gave birth at all. The rest postponed marriage, further reducing fertility rates. This pattern is unique to Europe, and only occurred west of a line from St. Petersburg to Trieste (Hajnal, 1965, ch. 1.2). Other parts of the world, such as China, used infanticide for the same purpose, but with less effectiveness.

There are two variants of the Malthusian model. The model in its strongest form has its roots in the classic "iron law of wages." Without shifting mortality and fertility schedules, it predicts stagnant real wages. Without technological change or other supply shocks, population size will stagnate. The weaker version emphasizes equilibrating mechanisms, not outcomes. The positive and preventive checks identified by Malthus influence demographic growth. Only if these responses are sufficiently large, and only without further perturbations to the system, does the weak version lead *in the limit* to a return to a subsistence wage.

It is clear that the strong version – with stagnant wages at the subsistence level – can claim little empirical support. Stock variables like population size are invariably slow-moving. Shifts in mortality schedules (possibly as a result of urbanization) could produce new equilibria, but our chances of observing them will depend on the relative magnitudes of short-term and enduring shifts. For England, the real-wage data computed by Clark (2005, p. 1311) replace the traditional wage series computed by Phelps-Brown and Hopkins. They are based on a broad array of commodities and a comprehensive set of nominal wages. Both the Clark and the Phelps-Brown series show the same, surprising sharp decline in wages in Tudor England between 1495 and 1575. This decline is puzzling, since it was accompanied by a stable and then rising population, as well as by unusually long life expectancy. Recent calculations by Allen (2001) and others show that, in the long term, wages in Europe followed divergent trajectories. Northwestern Europe saw marked rises in wages, often at the same time as population increased. This would contradict the strong version if both north and south were subject to Malthusian forces.[6] Furthermore, some of the

[6] Real wages may not reflect changes in welfare, because some of the wage premia available in towns only compensated for higher mortality risk. We are also not sure how much payments in kind varied over time, and if higher payments in cash compensated for declining payments in wheat, etc.

debate regarding the outcome of Malthusian processes conflates real wages with real per capita GDP or income. This is problematic because participation rates and hours worked may have changed, leading to considerable changes in incomes per capita and per family even at more or less constant wages. Indeed, rising participation rates could, all other things being equal, lead to real wages and real income per capita moving in *opposite* directions. The rise of cottage industries in the countryside after 1650, the famed "proto-industrialization" phenomenon, would do exactly that. There is also reasonable evidence to believe that labor participation rates were rising in the century before the Industrial Revolution (De Vries, 1994, 2008; Voth, 1998, 2001a, 2001b).

Confronting the model's predictions in its weaker form – with an emphasis on equilibrating mechanisms – is less demanding. We can observe flow variables such as births and deaths at high frequencies, and relate them to food prices and real wages. Over the short run, movements in population before 1750 seem to offer some limited support for a Malthusian response.[7] Mortality and nuptiality can adjust even over the short run. High-frequency events such as famines, wars, and epidemics had much smaller long-term effects than has often been assumed: a sharp decline in population was normally followed by higher wages. Within a few years, unusually high birth and low death rates would compensate for the initial decline in population (Watkins and Menken, 1985; Watkins and van de Walle, 1985). Lee's original work on the Wrigley–Schofield population data showed nuptiality responding (weakly, and with a lag that stretches credulity) to wages, but life expectancy to be largely independent of the wage.

In testing both the weak and the strong version of the Malthusian model, endogeneity is a major challenge. Wages influence population size and vice versa (Lee and Anderson, 2002). One potential way forward is to use an exogenous source of identification. Recent work by Kelly (2005) suggests that weather is a useful instrument for wages – the part of real wage variation that is driven by it is not the result of a feedback from population. Estimated in this way, there is strong evidence that Malthusian restrictions bound in England before 1800, with marriage rates reacting strongly (and positively) and death rates strongly (and negatively) to wage changes. Kelly's findings suggest that passing real-wage fluctuations had a larger effect on nuptiality than on mortality. This implies that, in the short run, the preventive check was stronger than the positive one, but both were significant.

Vector autoregressions offer an alternative method. Nicolini (2007) and Crafts and Mills (forthcoming) use them to model the dynamic feedback

[7] See e.g. Galor (2005, pp. 183–84) for some graphs that indicate that in pre-industrial Britain population and real wages moved roughly in opposite directions, and that crude birth rates and crude death rates were negatively correlated.

between fertility, mortality, and the real wage in England. In this way, they examine the strength of the preventive and the positive checks. Both papers find much stronger evidence in favor of Malthusian checks and balances for the period up to the middle of the seventeenth century than for later decades. The fertility channel is particularly potent, while the mortality channel appears weaker. After 1650, the fertility channel declines in strength. Nicolini (2007) concludes that "perhaps the world before Malthus was not so Malthusian." As is the case with all negative results, it is not always clear if it is lack of power in the statistical procedures used, a shortage of identifying variation in the data, or the true absence of a causal link that it is responsible. Overall, the IV-procedure used by Kelly appears more promising as a way to pin down causality and the strength of interactions.

Some progress has thus been made in terms of analyzing short-term responses. However, the precise contribution of demographic factors to divergent per capita incomes in early modern Europe remains largely unclear. Golden-age Holland had exceptionally high wages compared with the rest of Europe, and a stagnant population. It is not clear what particular feature of fertility behavior or of death schedules (if any) accounts for this beyond the high levels of urbanization. The Dutch example suggests that, while Malthusian adjustment mechanisms may have operated in the short run, many interesting shifts were caused by other factors. Since the late Middle Ages, there were throughout Europe regions and towns in which incomes exceeded subsistence levels, traditionally defined, without a concomitant rise in population size. Some unified growth models (Galor, 2005; Jones, 2001) predict (modestly) rising per capita incomes before the Industrial Revolution. This is on the whole confirmed: living standards drifted up, albeit slowly, in some parts of Europe in the centuries before 1800. The reason proposed – a delayed population response to technological advances – is not altogether persuasive: total fertility rates for females in many pre-modern populations (and especially European ones) were substantially below their biological maximum. Birth rates rebounded vigorously after each famine. This suggests that they could respond to rising living standards. One important question, then, is why Europeans curtailed their fertility, and why they did so in a peculiar way that involved delayed marriages for some women, and a life of celibacy for others. What social institutions underpinned the "European marriage pattern"? One interesting hypothesis links the emergence of fertility restrictions to the high price of labor after the Black Death (van Zanden and de Moor, 2007), which made female workers more valuable. This would have made it beneficial to keep them in the workforce as long as possible, and to delay motherhood. But why did this mechanism work in the Netherlands and not, say, in Italy, China, or India? This question is particularly relevant, since all these areas suffered plague outbreaks.

One way of linking the persistence of high wages with specific European features involves interactions between cities and death schedules. European cities were veritable death traps, with far higher mortality rates than the countryside. In contrast, in China and Japan urban and rural mortality rates were broadly similar (Woods, 2003). Different cultural practices, such as the regular removal of excrement from Far Eastern cities for use as fertilizer in the countryside, may have played an important role. Not only were European cities far more unhealthy places to live in under normal conditions (due to congestion and poor sanitation), but they were especially sensitive to contagious epidemics and military disasters such as sieges and plunder. Hence the DD curve in the graph, which is a composite of rural and urban demographic behavior, could slope upwards *over some part of the w-D space* because of a composition effect. There could then be multiple equilibria: societies could move from one state, where population was large, wages were low, cities small, and aggregate death rates low, to another, where wages were higher, cities larger, death rates higher, and the population smaller. A major shock, such as the Black Death, could push the economy from one equilibrium to another.[8]

Cities mattered for reasons other than excess mortality. They were the loci of much international trade, and of private-order institutions that supported the operation of markets in goods, capital, and labor. They were also centers of inventive activity. Urban activities produced a higher likelihood of inventing new techniques with a large economic impact: technology itself could have improved as a result of urbanization (Clark and Hamilton, 2006; Voigtländer and Voth, 2006). City growth may therefore have gone hand in hand with a slow, gradual outward shift of the technology schedule, making higher wages compatible with bigger populations. What this means is that, at any level of population, income would be higher with a larger urban sector, which would go some distance to explain the Dutch "anomaly." This means that far from being simply an *indicator* of productivity, urbanization itself could become a driving force increasing output per head. In this case, Malthusian forces could still dominate short-run changes, but the key *explanandum* would no longer follow from its basic tenets.

At some point, in the majority of European countries, population growth accelerated in an important way. Often, a rise in fertility and/or a decline in mortality signalled the end of the previous regime. Eventually, fertility rates followed the downward trend of mortality – completing the "demographic transition."[9] The latest revisions of the Wrigley – Schofield English population estimates (Wrigley et al., 1997) show that fertility increases dominated as a cause of more rapid growth; mortality played a role, but it was responsible for

8 Such a model is developed in Voigtländer and Voth (2008).

9 A good summary is Chesnais (1992). The concept goes back to the work of Warren Thompson in the 1920s.

only about one-third of the acceleration.[10] It seems that by 1750 the old demographic regime was breaking down. The work of Patrick Galloway (1988) shows that in the middle of the eighteenth century the short-term behavior of British vital rates was no longer responsive to changes in prices.

While some of the population explosion in Europe after 1800 derived for a while from higher fertility, declining mortality eventually became more important. Fertility eventually followed the downward trend, in many cases with a delay measured in decades (Lee, 2003; Coale and Watkins, 1986). Most of this fertility decline was concentrated in a few decades, starting in 1870 and accelerating after 1890. In some countries, such as the United Kingdom, Germany, Sweden, the Netherlands, Finland, and Belgium, there were sustained and sometimes marked increases in fertility before decline set in. For example, the average number of children per woman rose from 4.5 to 5.5 in the Netherlands between 1850 and 1880. By 1900, it had returned to its earlier level. In most European countries the first significant reductions in fertility occurred after the 1880s, long after industrial change had started to take hold on the continent. Some countries saw large reductions in infant mortality before fertility started to decline (Sweden, Belgium, Denmark); in others, both series show a concurrent downward movement (France, Germany, Netherlands) (Chesnais, 1992).

Finding an economic reason for fertility decline has not been easy, and there is currently no consensus on the principal contributing factors (Alter, 1992). Variations both across Europe and over time present challenges of interpretation. The biggest comparative project on the fertility transition, the Princeton European Fertility Project (EFP), concluded that there was no clear link between socioeconomic factors and fertility change. Instead, ethnic, religious, linguistic, and cultural factors appeared to be dominant (Coale and Watkins, 1986). Woods (2000) reached a similar conclusion for Britain, attributing the Victorian decline in fertility to changing ideology, primarily "the desire or willingness to limit family size from the 1860s on" (p. 150), and suggests, more provocatively, that "the very question 'how many children should we have' was new to most Victorians" (p. 169). The leading explanation for fertility change is the "diffusion model," where knowledge about prophylactic techniques spread along linguistic lines. The principal reason why scholars have accepted the findings of the EFP is the remarkable similarity in the timing of the transition, and its spread along linguistic lines.[11]

[10] Wrigley (1983) showed that without mortality decline, eighteenth-century growth would have accelerated by 1.25 percent; without fertility change, growth would have improved by 0.5 percent. This implies that over 70 percent of the acceleration was driven by changes in fertility. Wrigley et al. (1997) qualify these conclusions to some extent, finding a faster decline in mortality, but the relative rankings are unlikely to change significantly.

[11] As Cleland and Wilson (1987) argue, "the simultaneity and speed of the European transition makes it highly doubtful that any economic force could be found which was powerful enough to offer a reasonable explanation."

Studies that go beyond the broad aggregates and look at regional data sometimes reach different conclusions. For example, in Bavaria the opportunity cost of women's time, religion, and political affiliations appear to have played a big role (Brown and Guinnane, 2002). Furthermore, the statistical basis for some of the EFP's conclusions may be less robust than had previously been assumed.[12] The simultaneity of the drop in reproduction rates across Europe in the decades before 1914 makes it unlikely that economic factors can account for the fertility decline all by themselves. Exogenous, non-economic factors probably dominated in the great decline of European fertility. This need not present a challenge to all growth models. Yet for the more ambitious class of structural models in the unified growth tradition, the apparent incapability of economic factors to have a clear bearing on fertility outcomes represents a challenge.

In many models of long-term growth, the fertility transition plays a crucial role, and the timing of fertility decline is central to many theories explaining the transition to self-sustaining growth. The decline is normally modelled as a response to changing economic incentives. Leading interpretations by Becker and Barro (1988) and Lucas (2002) emphasize the quantity–quality trade-off facing parents in a context of faster technological change and higher returns to human capital. The standard arguments are that (i) skill premia surged, often because of technological change; and (ii) parents limited fertility as a response to this change in the trade-off between child quantity and quality. This is not unproblematic. Returns to human capital, conventionally measured, probably did not increase significantly before 1870. Models that link population dynamics to technological progress itself, such as Galor and Weil (2000), run into timing problems in the case of Britain, because demographic growth accelerated there in the mid-eighteenth century, *before* any serious impact of technological change on output per capita can be discerned. Furthermore, since the economic benefits of formal education were probably minor for working-class employment, any model of parental fertility choice based on quality–quantity trade-offs faces problems, explaining at best the demographic behavior of a minority group.

Definitive evidence for a quantity–quality trade-off is lacking. What is more plausible is to argue that the net *costs* of child quantity increased in the second half of the nineteenth century. An alternative interpretation thus emphasizes the importance of government intervention through compulsory schooling laws and child labor regulations. Doepke (2004) argues that the latter were

[12] A much larger research project on German fertility decline is now under way (Sheilagh Ogilvie, Timothy Guinnane, and Richard Smith, with Markus Küpker and Janine Maegraith, Economy, Gender, and Social Capital in the German Demographic Transition, available at www.hpss.geog.cam.ac.uk/research/projects/germandemography/2005), using that country's extraordinarily rich data sources.

crucial, and further argues that other government policies (such as education subsidies) could not have had a similar influence. If the importance of government intervention is confirmed, examining the economic and other factors behind the adoption of child labor laws or educational reforms becomes crucial (Doepke and Zilibotti 2005). Galor and Moav (2006) emphasize the Balfour Act, which introduced compulsory schooling. In their view, support for the reform by capital owners, who needed more skilled labor, was critical.[13]

Yet we do not know with certainty that government intervention was crucial in moving children out of the factories and into the classrooms. For the United States, state schooling laws may only have had a small influence on child labor (Moehling, 1999). At the same time, data problems bias any estimate of such an effect towards zero. In the United Kingdom, Nardinelli (1980) and Kirby (1999) argue that child labor laws came in at the same time as technological changes made the employment of children less useful. There is therefore considerable tension between the views of theorists, who emphasize either rapid, skill-using technological change or effective government intervention, and the assessment of economic historians, who largely reject the former and find limited evidence of the latter.

Some data constraints will be hard to overcome. We have little information on what determined completed fertility rates, educational investment, age at marriage, and the like in the industrializing cities throughout Europe. There are no cohort-specific studies of fertility behavior at the micro level that would unambiguously identify the impact of discontinuous changes in schooling laws and the like. Wrigley and Schofield's famous *Population History of England* is based on family reconstitutions that focus on rural parishes, and their data end in 1837. Everywhere in Europe, family reconstitutions are harder to construct for the nineteenth century than for earlier periods because mobility increased. Future research should aim to improve our understanding of fertility behavior, and of the relevant costs of child-rearing. More detailed demographic analysis of the fertility choices of the working class – combined with information on rates of school attendance, the economics of apprenticeships and the like prior to and after the introduction of compulsory schooling laws – could do much to further our understanding of the demographic transition.

Institutions, good and bad

A good part of the modern debate about growth centers on the relative importance of institutions versus human capital (Acemoglu and Johnson, 2005; Rodrik,

[13] The failure of skill premia to rise (which we shall describe in more detail below) could then be explained by this supply shock.

Subramanian, and Trebbi, 2004; Glaeser et al., 2004). In cross-sections of countries from the late twentieth century, constraints on the executive tend to be positively correlated with higher per capita output. Because of the potential for reverse causation – with higher income per capita improving institutional quality – work on modern data has principally focused on finding an exogenous factor that affect institutions, but not economic outcomes, and so can be used to instrument for institutions. One such factor that has been used with great success is historical settler mortality. In a series of pathbreaking papers, Acemoglu, Johnson, and Robinson show that countries in which European settlers survived easily also ended up with more desirable institutional arrangements (e.g. Acemoglu, Johnson, and Robinson, 2001). They are markedly richer today, making it more likely that the link between institutions and efficiency is causal.

How much can institutional interpretations help us understand growth in Europe before 1870? What is the role of institutional change in the transition to self-sustaining growth? Any analysis of institutions in a historical setting needs to look beyond the role of the state and constraints on the executive. We do not know nearly enough about how institutions worked in Europe between 1500 and 1870 to pass judgement on their overall contribution to economic growth. In particular, we need to learn much more about legal processes, the role of "state building" and the importance of informal institutional arrangements.

Possibly the single best-known statement in the institutional tradition was formulated by North and Weingast (1989). They conclude that the Glorious Revolution and the Bill of Rights in England in 1688–89 did more than just put government finances on a firmer footing. Because of the boost to the role of Parliament and the greater influence of the common-law courts, the English monarch's power had been curtailed. Crucially, it was widely viewed as such because of credible commitment. High-handed breaking of contracts and seizure of property came to an end.[14] North and Weingast argue that, once property rights and constraints on the executive had been firmly established, risk premia fell. Capital accumulation accelerated, and investing in new ideas became much more profitable. Eventually, Britain's growth rate took off.[15]

Most institutional interpretations of the early modern period similarly focus on capricious despotic rule falling away. DeLong and Shleifer (1993) return to Montesquieu's famous argument that growth is more vigorous in republican states because they suffer fewer arbitrary interventions by the authorities.[16]

[14] These had previously been possible both through the legal system – namely the Star Chamber – and brute force (such as in the raid on the Tower of London, when the precious metals of goldsmiths were seized).

[15] North and Weingast are cautious not to link the events of 1688 directly to the Industrial Revolution, which began about seventy years later. Others have not been so prudent (e.g. Dam, 2005, p. 84).

[16] "An opinion of greater certainty as to the possession of property in these [republican] states makes [merchants] undertake everything … [T]hinking themselves sure of what they have already acquired, they boldly expose it in order to acquire more."

They argue that absolutist rule was harmful for three reasons – states run by ambitious, powerful princes fought more wars, taxed more comprehensively, and respected property rights less. Autocratic states also happened, on average, to be further away from the new trade routes to the Americas and Asia. It should be noted that only one of these channels is directly associated with the institutional interpretation in its narrow form, and DeLong and Shleifer cannot show that it is particularly potent.

A more recent paper by Acemoglu, Cantoni, Johnson, and Robinson (2005) argues that two of the channels identified by Delong and Shleifer interacted in a particular fashion to strengthen institutions. Countries that had opportunities for Atlantic trade experienced a gradual strengthening of bourgeois forces in society. Hence, "constraints on the executive" in Britain and the Dutch Republic grew, according to their estimates. Acemoglu et al. demonstrate that this improvement in the quality of institutions mattered for growth – urbanization rates surged wherever geographically determined "exposure" to Atlantic trade was high.

Institutional interpretations of the Industrial Revolution and its aftermath have emphasized the role of political economy. For the case of Britain after 1800, Acemoglu, Johnson, and Robinson (2005) and Acemoglu and Robinson (2006) have argued that political power mattered in large part because it made the redistribution of income possible. They distinguish between *de jure* power, that is, the power to pass formal laws and statutes, and de facto power, which includes the physical ability to overthrow the regime if those who have it do not find the policies to their taste. While by 1720 Parliament had concentrated a great deal of *de iure* power and thus elevated itself to the status of a meta-institution, it still needed to be concerned with the de facto power of the large masses of middle-class people who accumulated increasing economic wealth, and yet were to a great deal disenfranchised until the reforms of 1832 and 1867. Acemoglu et al. (2006) argue that the French Revolution acted as an exogenous shock to the political system of neighboring states. Defeat at the hands of Napoleon's armies prodded rulers in Prussia and Austria into major reforms. Elsewhere, conquest by the French swept away old political and social institutions in their entirety. Combined with the rising tide of technological change after 1850, the authors argue, the improved institutions of the set of countries ringing France accelerated growth in the early nineteenth century in an important way.

Institutions have thus gained a great deal of credence in the modern growth literature (Rodrik et al., 2004; North, 2005). Yet interpretations of Europe's growth record that rely primarily on institutions face numerous challenges. To start with, few scholars agree what institutions are, and how the concept should be applied to the past. North defined them as "a set of rules, compliance procedures, and moral and ethical behavioral norms designed to *constrain* the behavior of individuals in the interests of maximizing the wealth or utility of principals." Greif (2005, emphasis added) includes other modes of behavior that create historical regularities. In Greif's model, beliefs and ideology act as "deep" parameters that determine the efficacy with which societies set up rules making exchange and investment possible. Yet there are few good theories that explain in detail how institutions change and why some economies end up with "better" ones than others. Standard measures in the literature such as the (perceived) risk of expropriation, government effectiveness, and constraints on the executive, can all easily reflect choices by governments, and may change quickly. For any model that implies that better institutions work wonders through capital accumulation or technological progress, this would be problematic. Glaeser et al. (2004) show that all three standard measures of institutions often change after a single election. Presumably, property rights that are simply protected because of a ruler's whim are not worth a great deal. The volatility of these measures over time makes it less likely that they identify some structural parameter of the political system. Other, more obvious variables, such as judicial independence, proportional representation, and constitutional review, vary much less and are more likely to proxy for the structural constraints on governments that North had in mind. Yet in modern-day data, the effect of these variables is small and insignificant. What would be needed to settle the matter is a "deep" parameter of a country's political constitution that does not change quickly, and that is not simply a reflection of current economic and political conditions.

For the period 1500–1800, the "constraints on the executive" variable, as compiled by Acemoglu et al., successfully predicts urbanization rates. The same is true of Delong and Shleifer's absolutism indicator. Yet both concepts are troubling for the early modern period. Data problems abound, and coding variables based on the complex institutional arrangements in place in many European states before 1800 is not a challenge for the faint-hearted. The Habsburgs ruling Spain, coded as perfectly absolutist by Acemoglu et al. for the sixteenth century, often failed to get tax or other concessions from the Cortes of Castile. In their other kingdoms, such as Aragon, many medieval "freedoms" and the assemblies that protected them curtailed the monarch's powers. For instance, even for that epitome of absolutist rule, the Sun King of France, Louis XIV (and coded as a perfectly unconstrained "1" by Acemoglu et al.), there are important question marks. Historians have largely rejected the idea that his rule can meaningfully be described as implementing a successful, far-reaching absolutist agenda. For a generation, a new consensus inspired by the works of, *inter alia*, Georges Pagès and of Roland Mousnier (1970) has emphasized how much French kings at the height of absolutism still governed through social compromise and consensus, maintaining the stability of a traditional society and the influence of old elites for much of the time. It seems doubtful that the currently available classification schemes capture enough of what is directly relevant to the argument that institutions and restrictions on the executive caused economic growth before 1800.[17]

States with extensive checks and balances, such as Venice, the Holy Roman Empire, and Poland, constrained their rulers' freedom of action. Yet they did not become hothouses of economic dynamism. This may be because fettering the prince was neither unambiguously good nor bad in economic terms before the rise of modern, centralized states with clearly defined and stable borders. The late S. R. Epstein (2000) has emphasized the advantages conveyed by a powerful state that could curb local rent-seeking and resolve coordination problems. Most "constraints on the executive" took the form of rent-seeking groups ensuring that their share of the pie remained constant. Unsurprisingly, large parts of Europe's early modern history read like one long tale of gridlock, with interest groups from local lords and merchant lobbies to the Church and the guilds squabbling over the distribution of output. One particularly striking example of the inefficiencies produced in French agriculture is provided by Rosenthal (1992). None of the groups that offered resistance to the centralizing agendas of rulers in France, Spain, Russia, Sweden, and elsewhere were interested in growth. Where they won, they did not push through sensible, long-term policies. They often replaced arbitrary taxation by the ruler with arbitrary exactions by local monopolies.[18]

In the early modern period, states with good institutions were often weak. The cut-throat nature of international politics undermined the viability of good governance. The League of Cambrai (1516) laid Venetian power low, and contributed to the eventual decline of Venice's prosperity. Today, constraints on the executive go hand in hand with lower probabilities of military conflict, as democracies are unlikely to go to war with each other (and tend to win in wars against non-democratic powers). In the early modern period, the correlation probably had the opposite sign. Political entities with highly effective constraints on the executive quickly became victims of outside powers whose rulers operated without being hamstrung by domestic opponents. Thus in early modern Europe less-developed but large and militarily strong political units, such as the young nation-states of Philip II, Gustavus Adolphus, and Louis XIV,

[17] For a recent critique of the revisionist argument, cf. Beik (2005).

[18] The case of Venice is instructive. In terms of its institutional setup, it is hard to think of a political entity that would more closely approximate the modern ideal. Property rights were well protected. The high bourgeoisie controlled politics and the courts. Doges were elected officials, theoretically for life but in reality subject to good performance. A patent system was in place as early as the fifteenth century. Yet despite its early riches and success as a sea power, Venice declined both militarily and economically.

threatened the richer but smaller city states in Italy, Germany, and the Low Countries. Economically successful but compact units were frequently destroyed by superior military forces or by the costs of having to maintain an army disproportionate to their tax base.[19] The only two areas that escaped this fate enjoyed unusual geographical advantages for repelling foreign invasions – Britain and the northern Netherlands. Even these economies were burdened by high taxation, the cost of surviving in a mercantilist world based on the notion that the economic game between nations was zero-sum, and that foreign trade was a servant of political and dynastic interests.

A fundamental trade-off was thus created: a powerful central government was more effective in protecting an economy from foreign marauders, but at the same time the least amenable to internal checks and balances. Carried to the Polish extreme, strict constraints on the executive were not conducive to economic development – not least because they could contribute to the disappearance of the state itself at the end of a sequence of gruelling military defeats. Rather than a weak or strong government, the most important institutional feature of a society might have been institutional agility, the capacity for institutions to adapt to changing circumstances with a minimum of pain and friction.

Two observations to summarize the importance of institutions in the post-1750 transformation of Europe are in order. One is that throughout western Europe we observe after 1750 a rising tide against the rent-seeking institutions that were associated with the mercantilist *ancien régime* (Mokyr, 2006). The roots of this development involve some combination of the changing political influence of economic elites and the influence of a more liberal ideology. Second, change often occurred by force (e.g. in the United States and France). The main exception is Britain, where the existence of a meta-institution (i.e. Parliament) permitted adaptation to changing circumstances and beliefs, and reform of the system peacefully and without major upheavals (Mokyr and Nye, 2007). Yet even here it could be argued that the settlement following the Glorious Revolution would not have been possible without the bloodshed of the Civil War.

The literature on institutions in Europe, 1500–1870, has primarily focused on the state and formal institutions. Non-governmental institutions, both formal and informal, have received much less attention (but see Mokyr, 2008). This is striking since work on the medieval period has given prominence to these arrangements (Greif, 2005). If, as the institutions literature argues at a fundamental level, respect for property rights and recourse to due legal process are key for economic development, then we need to construct variables that

[19] In modern data, there is a robust, negative correlation between military conflict and political instability on the one hand, and growth on the other (Alesina et al., 1996).

more closely capture this dimension.[20] A more comprehensive and historically meaningful set of indicators should measure effective legal or customs-based constraints on the actions of the executive or of local power groups – anything that makes it harder for might to be right without due recourse to the law. In addition, opportunistic behavior leading to Pareto-dominated equilibria could be overcome by a host of mechanisms (besides the standard of third-party enforcement) in which members of select groups were able to establish their trustworthiness through a variety of costly signals (Greif, 2005) and play cooperatively. The modern literature on institutions has shown that such arrangements may still have a fair amount of explanatory power today (Ellickson, 1991; Posner, 2000). They need to be investigated for periods not covered by Greif, and their significance relative to that of formal institutions such as Parliament explored.

Eighteenth-century Britain is a case in point. In a recent paper Mokyr (2008) has argued that informal arrangements and cultural change had effects similar to public institutions, facilitating the operation of markets. Within a larger group of people, a stable equilibrium emerged that allowed signalling of trustworthiness. Middle-class people adopted certain virtues associated with gentlemanly behavior; since gentlemen were supposedly not greedy, they could be expected to cooperate in one-shot Prisoner's Dilemma games (Clark, 2000). Moreover, Britain in the eighteenth century experienced an enormous growth in formal and informal social networks, through the growth of friendly societies, masonic lodges, and eating clubs, with an estimated membership of 600,000 in 1800. The effect of this growth was to make reputational mechanisms more effective, since reports of non-cooperative behavior would soon be disseminated. This may have mitigated free-riding behavior in the private provision of public goods in eighteenth- century Britain, which included local administration, overhead projects, education, and health. It can be argued that such informal institutions not only supported markets, but also helped Britain take the technological lead, because the success of these informal institutions made its apprenticeship system particularly effective (Humphries, 2003). The apprenticeship contract was particularly vulnerable to opportunistic behavior, and in Britain the guild system that enforced it elsewhere was weak – yet it functioned well in Britain. As a consequence Britain could count on a large number of highly skilled craftsmen and mechanics, whose role in the Industrial Revolution may well have been critical.

Power-sharing arrangements between nobility and the rich bourgeoisie after 1830 underpinned some of the smooth functioning of institutions in Britain, as

[20] Acemoglu and Johnson (2005) argue that institutions protecting property are crucial, and that "contracting institutions" only influence the type of financial intermediation one ends up with.

Acemoglu and Robinson have emphasized. Workers may not have had *de iure* power, but their implicit ability to rebel gave them de facto power. Analysis based on "realpolitik" overlooks the growing influence of Enlightenment ideology on political institutions; analyzes of what de facto power consisted of are still incomplete. The British military suppressed popular unrest in the 1790s as well as the Luddite riots very effectively, and the Chartist movement remained a mostly non-violent movement, its few more threatening outbursts readily suppressed. It may thus be that *de iure* and de facto power coincided to a great extent. Perhaps this was the key to the success of Britain's political model. It is striking, nonetheless, that while those who had political power did use it at times to redistribute income to themselves (most blatantly by the reformulation of the Corn Laws in 1815), the tendency to do so lessened as the nineteenth century wore on, and by 1860 rent-seeking in Britain was at a historical nadir.

Human capital and culture

In many models of long-run growth, the transition to self-sustaining growth is almost synonymous with rising returns to education, and a rapid acceleration in skill formation. Becker, Murphy, and Tamura (1990) model an economy without a fixed factor of production. Improvements in human capital directly feed into higher output. Human capital is produced, it is assumed, by investments of parental time. Parents maximize their own utility, derived by their own consumption, the number of children they have, and their quality. When parents start to invest massively in the education of their offspring, growth rates rise. Once incomes are high enough, fertility falls, leading to yet more investment in child quality. In this model, human capital and growth are almost identical. Lucas (2002) extends the approach of Becker et al. by adding a land-using sector with diminishing returns, and a modern sector where human capital enters linearly. Many unified growth models have followed in the same direction, adding interactions with the rate of technological change.

Developments during the Industrial Revolution in Britain appear largely at variance with these predictions. Most evidence is still based on the ability to sign one's name, arguably a low standard of literacy (Schofield, 1973). British literacy rates during the Industrial Revolution were relatively low and largely stagnant. This is especially true if we take into account the fact that Britain was relatively rich before the Industrial Revolution, and that demand for literacy rises with income (Mitch, 1999). Britain's ability or willingness to educate its young did not appreciably improve during the years of the Industrial Revolution. School enrolment rates did not increase much before the 1870s (Flora, Kraus, and Pfenning, 1983).

Models in the Lucas tradition often predict an increase in the demand for human capital during the transition to self-sustaining growth. Also, technological change should be heavily skill-biased. This is historically problematic. Our knowledge of the behavior of the skill premium over time is incomplete, because estimates are based on a few skilled occupations, which may not be representative. Moreover, the skill premium is a reduced form measure, and changes in it could reflect any combination of changing supply and/or demand factors. There is little firm evidence of an increase in the returns to education in the eighteenth or nineteenth century. Williamson (1985) claimed to show that the skill premium surged between 1750 and 1850 in Britain, and declined thereafter. The consensus view amongst economic historians does not accept the Williamson interpretation. As Feinstein (1988) convincingly demonstrated, there is no clear evidence that skill premia changed at all over time.[21]

It is doubtful that the main developments in manufacturing during the Industrial Revolution, or even developments in its aftermath, depended on an increase in human capital. Possibly, some administrative tasks became more important. The rise in pay rates for highly literate workers observed by Boot (1999) suggests that there were some (small) parts of the economy where formal education may have paid off. Yet technological change itself was probably not skill-biased. As the machine-breaking Luddite riots highlighted, it may well be that de-skilling accompanied the first Industrial Revolution.[22] This would present an important challenge to the dominant economic models of long-run development. In the textile industry, the cotton mules, spinning jennies, and Arkwright frames replaced skilled labor with a mixture of capital and unskilled labor. For more traditional sectors, the evidence is much more mixed. Clark (2007a) examined the ratio of craftsmen to laborer wages in England, 1700–1850. His evidence shows a decline from a premium of 65 percent to 50 percent. Of course, the building trade experienced limited technological change. If the sector is nonetheless indicative of broader trends in the economy, then one would have to conclude that mild de-skilling occurred in the one and a half centuries before 1850. Perhaps focusing on the *average* level of human capital in industrializing societies is less useful than changes in its distribution. The technological changes of the nineteenth century created a demand for highly skilled mechanics and engineers in the upper tail of the distribution, while possibly reducing the need for skills among manual laborers. The failure of traditionally measured skill premia to show a rise may well mask an increasing polarization within the workforce, with

[21] See also van Zanden (2009).

[22] Adam Ferguson, Adam Smith's contemporary, noted in 1767 that "Many mechanical arts require no capacity … ignorance is the mother of industry as well as superstition … Manufactures, accordingly, prosper most where the mind is least consulted." For a recent model emphasizing the role of deskilling, see O'Rourke, Rahman, and Taylor (2008).

industrialization raising the returns to supervisory and advanced mechanical skills, and reducing those for standard ones (such as blacksmithing, carpentry, and weaving).

Also, for the time being the jury appears to be out on whether increased human capital formation from the middle of the nineteenth century onwards was an endogenous response to changes in factor prices and other economic incentives, whether it was a result of higher real incomes (education for one's children being a normal consumption good), or whether it was the result of "exogenous" shifts in the supply of education, such as the long-delayed effect of the Enlightenment, of nineteenth-century nationalism and nation-building, or attempts to strengthen social control over the lower classes.[23]

There is little evidence to support unified growth models if we identify human capital with formal education only, and the break with the pre-industrial period as occurring in Britain after 1750. The main conclusion appears to be that, while human-capital-based approaches hold some attractions for the period after 1850, few growth models have much to say about the first escape from low growth. Models that endogenize the transition from skill-replacing to skill-using technological change are just beginning to appear (O'Rourke, Rahman and Taylor, 2008). Our verdict changes somewhat as we widen our focus. As we turn from the particular case of England to trends in Europe as a whole, analyze the longer period 1500–1870, and include broader definitions of human capital to include factors such as numeracy and discipline, as well as informal education mechanisms such as apprenticeship, the fit between theory and history increases.

There is no doubt that some forms of human capital (such as literacy and numeracy) were on the rise in Europe long before the Industrial Revolution. In part this was due to the Reformation, in part due to slowly rising incomes, and possibly to a rising demand for literacy in the service sector during an age in which commerce and finance were growing rapidly. Measuring literacy rates in a consistent and comparable fashion is no minor matter, especially with the kind of pre-1800 sources available. A recent literature survey, focusing on the ability to sign one's name in and around 1800, rates this proportion at about 60 percent for British males and 40 percent for females, more or less at a par with Belgium, slightly better than France, but worse than the Netherlands and Germany (Reis, 2005, p. 202). Baten and van Zanden (2008) examine book production in early modern Europe. They find a veritable explosion of output per capita after the invention of moveable type, with production increasing between 10- and

[23] The latter effect would be in the spirit of Acemoglu and Robinson's (2000) paper, which sees the extension of the franchise as a reaction to revolutionary threats. A similar argument could possibly be made about the introduction of compulsory schooling.

100-fold. The Netherlands and the United Kingdom are far ahead of other countries – the richest areas consumed the largest number of books.[24]

One additional measure of human capital is numeracy. The ability to make sense of numbers, to remember them correctly, and to perform minor transformations is a crucial skill in many commercial transactions. Measurement in a historical setting can be achieved via a shortcut. As suggested by Mokyr (1985), we can use age-heaping as an indicator of numeracy. Many historical sources show a tendency for ages to be reported in multiples of five, while the true distribution should be smooth. A'Hearn, Baten, and Crayen (2009) compile a comprehensive database for the last two millennia. They find some evidence that numeracy was trending upwards in Europe from the sixteenth century onwards. Britain's experience after 1800 suggests stagnant literacy, despite sweeping changes in the economy. The more evidence we can collect on areas outside Europe, and the closer they can be linked to data on income differences as a result of higher numeracy, the more potent tests of models in the Becker and Lucas tradition become.

If we are to make progress, historians and economists need to broaden their conception of skills. They were not yet, as a rule, acquired at schools or similar formal institutions. Rather, they were mainly transmitted through personal contracts. Apprenticeship was the main form of training. Trainees would be "indentured" to a master. The contract involved a commitment by the trainee to work during his learning period, an obligation to teach for the master, and at times cash payments by parents (Humphries, 2003). New technology was put in place, made operational, and maintained by a small army of highly skilled men. They were clock- and instrument makers, woodworkers, toymakers, glasscutters, and similar specialists, who could accurately produce the parts, using the correct dimensions and materials, who could read blueprints and compute velocities, understood tolerance, resistance, friction, and the interdependence of mechanical parts. These anonymous but capable workers were an essential complement to inventors and engineers, and comprised perhaps 5 to 10 percent of the labor force. They turned models and designs into working machinery, operated and repaired it, and produced a rising tide of small, incremental, but cumulatively indispensable microinventions. Without them, Britain could not have become the "workshop of the world."

Most of the skills that the workers' elite of skilled craftsmen brought to the factories were the culmination of a century-long accumulation of expertise in traditional crafts. If the rise of new technology, and the high complementarity of their skills with the adoption of more productive machinery, made their human capital more valuable, we should find changes in the wage premium for this

[24] The authors argue that their indicator of human capital accumulation is a good predictor of subsequent growth. Baten and van Zanden (2008).

group. Collapsing wages of handloom weavers might have been compensated for by the growing demand for the highly skilled craftsmen who erected the new spinning and weaving machinery in Lancashire. While evidence is fragmentary, there is some indication that employers scrambled to find glass-cutters, millwrights or fine mechanics in eighteenth-century England (Musson and Robinson 1960). One conceptually appealing test of human-capital based models of the Industrial Revolution would focus on movements in the pay rate of this labor aristocracy compared with the rest, and on the supply response that these differences in pay engendered.

We may have to widen our definition of the relevant human capital yet further. The rise of the factory system required general skills that were not necessarily transmitted through formal schooling – discipline, punctuality, and respect, in addition to literacy and numeracy. Recent work in labor economics has highlighted the importance of non-cognitive skills (Heckman and Rubinstein, 2001). The equipment and materials used by workers belonged to the capitalist and were costly. Factory owners needed to instill in workers a culture of loyalty and sobriety, and a willingness to take instructions from and cooperate with other workers. While similar to the discipline exerted by masters over apprentices, it was reinforced by the expensive equipment in factories.[25] Beyond that, the more complex technology and finer division of labor created interdependencies between workers that required coordination between them that would have been hard to enforce unless workers were willing and cooperative. Wage premia for *disciplined* work in the factories were high vis-à-vis other, more self-determined forms of employment, and the factory system's profitability relied crucially on work intensity (Pollard, 1965; Clark 1994). In addition, steep experience-based earnings profiles in the textile industry offered high returns to those who could stand the habituation to factory work. During their early years, when unskilled workers such as brickmakers were better paid, skilled workers were effectively investing in their own human capital; by age 35, they could look forward to earning 2.3 times the wages of a brickmaker, and more than a coal miner (Boot 1995).[26]

For similar reasons, monitoring workers was an important task. If "discipline capital" mattered more for the first Industrial Revolution than education as conventionally measured, economic historians should compile more comprehensive wage measures that capture the rewards for workers who successfully

[25] This insight is hardly indebted to modern theory: Marx, in a famous passage, cites an industrialist telling the economist Nassau Senior that "if a labourer lays down his spade, he renders useless, for that period, a capital worth 18 pence. When one of our people leaves the mill, he renders useless a capital that has cost £100,000" (Marx, 1967, I, pp. 405–06).

[26] Coal miners are arguably a better standard of comparison, since the wage of textile operatives will also reflect differences in the harshness of working conditions – and since those in coal mines were probably worse than in textile factories, skill accumulation is a good explanation.

internalized the demands of the machine age. Also, if the returns to disciplining workers were large, we should find high and rising pay premia for outstanding foremen and other members of the evolving hierarchies that ensured the smooth running of nineteenth-century factories. The most obvious testable implication of this idea is that early factory owners should have had a preference for the employment of comparatively more pliable workers, even if they were of low skill – that is, women and children. This was very much the case in the early stages of the textile mills. Similarly, one valid test of the human-capital approach would focus on highly skilled workers such as the textile operatives examined by Boot (1995) and Leunig (2001), and ask whether they received greater rewards for investing in their skills (by accepting years of poorly paid on-the-job training) than, say, apprentices in traditional sectors.

These observations suggest that non-cognitive skills and informal education may have mattered more in explaining the transition to self-sustaining growth in Europe than formal schooling and traditional reading and writing skills. In this sense, distinctions between education and human capital on the one hand, and culture on the other, become increasingly artificial. Since Max Weber's work on the spirit of capitalism, culture is one of the "usual suspects" that may determine wealth and productivity, and modern scholars (e.g. Jones, 2006, pp. 126–32; Temin 1997) have concurred. There is now increasing evidence that its impact can be demonstrated in modern-day data. This suggests that economic historians may want to revisit the issue. Guiso, Sapienza, and Zingales (2006) show an exogenous effect from culture to income. Culture also has a lot of persistence (see also Tabellini, 2006). Here, culture is defined above all in terms of the values and beliefs of individuals. While trying to control as much as possible for endogeneity, these studies show that when people trust one another, believe that if they work hard they will get ahead in the world, and that on the whole the formal institutions of power in the country are not threatening them, economic progress will ensue. We have no record, of course, of such poll-based data for historical times, and so there is no easy way that such findings can be reproduced. But the importance of private-order institutions in many early modern European societies is striking. These normally involved cultural beliefs that made people keep their promises and behave in an honorable way because opportunistic behavior was not a dominant strategy in a setting that ensured repeated interactions. Reputations became an asset to be managed carefully. In this interpretation, the middle classes of commercial societies adopted a more cooperative mode of behavior, leading to Pareto improvements. The focal point in such equilibria may well have been what middle-class people perceived as "gentlemanly behavioral codes." These signalled that one was uninterested in money, and hence would care more about honor than personal gain. This reduced the risk of cheating

(Mokyr, 2008). It may well be that such social norms were far more important than third-party enforcement of laws and contracts in the support of European markets, especially credit and labor markets.

But whence such middle class values? In an innovative paper, Doepke and Zilibotti (2008) emphasize the differences between aristocratic and middle-class behavior. Doepke and Zilibotti offer a model of class formation through endogenous, inheritable preferences. They argue that the rise of a bourgeois elite in industrializing Britain should be seen as a surprise. Before the transformation got under way, aristocrats had all the odds stacked in their favor – available funds, political connections, access to education. Yet few members of the old political elite actually got rich in manufacturing after 1750. Doepke and Zilibotti argue that this is because other groups in society – the middle classes – had accumulated a larger stock of "patience capital." A host of cultural practices and norms made them delay gratification. Artisans needed to acquire this skill, since it takes a long time to complete the training needed to become a craftsman. On the other hand, the old aristocracy taught their children how to enjoy leisure. It therefore provided them with a culture that worked against both hard work and investment. Through centuries of careful saving and educating their young, the middle class built up both financial capital and valuable cultural traits. As the new technologies of the Industrial Revolution suddenly offered greater returns to patience, the groups best placed to exploit them were not the elite but those with plenty of patience. Doepke and Zilibotti argue that it is this kind of culture that played a central role in the subsequent development of capitalist industrialism. In their account, the *absence* of well-functioning credit markets was key for the rise of the middle classes – only when financial markets are segmented do returns to patience differ across groups.

The concept of patience capital holds considerable promise. It may be no accident that the "nation of shopkeepers," as Adam Smith called Britain, became the first to industrialize. It offered an environment in which bourgeois values and practices flourished and gained in relative importance. If Europe saw a rise in bourgeois values prior to the Industrial Revolution, this was complemented by a rise in work intensity and the length of the working day for the lower classes, and a growing orientation towards the market at the expense of self-sufficiency. De Vries (1994) termed this change the "industrious revolution." By the eighteenth century, even Catholic rulers were abolishing holy days to boost labor input in their economies (de Vries, 2008). Clark (1987) found evidence that work intensity in the most economically advanced parts of Europe was much higher than elsewhere. Voth (1998, 2001a) argues that the work-year in Britain was already long in 1750, and that it lengthened further because of a decline in festivals, holy days, and the practice of taking Mondays off ("St. Monday"). Such changes are consistent with the model proposed by

Doepke and Zilibotti ((2008), in which those with relatively limited "leisure skills" eventually became the dominant classes.

In recent years there has been growing interest in Darwinian selection models as an explanation of cultural change. Galor and Moav (2002) offer a model in which the crucial state variable that changed during the pre-industrial period is not just population size, but "human quality" (genetic or behavioral). Households endowed with more desirable human characteristics (education, the right genes, economically beneficial attitudes) produce more surviving offspring and gradually but ineluctably change the composition of the population. Therefore the quality of the human population drifted up prior to the Industrial Revolution. The Galor–Moav approach has recently received some qualified empirical support. Clark and Hamilton (2006) find that the rich and literate in early modern England fathered more surviving children. Clark's and Hamilton's result that wealthier Englishmen had more surviving children may suggest that, instead of leading to an upward drift in some unmeasured, unnamed indicator of human quality, this simply enlarged the proportion of those who had learned to save (and invest), and those who passed such values on to their offspring. Given that European living standards far exceeded subsistence levels during the early modern period, many more children could have been born. Constraints on fertility behavior were mostly social and cultural (working through nuptiality rates). Such a change in population composition could also have contributed to the decline in English interest rates since the Middle Ages (Clark 1988), from 10–11 percent in the thirteenth century to 4 percent by the eighteenth. A gradual increase in savings, caused by compositional effects attained through the increase in the relative number of those who were more patient, would be an alternative to theories that attribute the rise in savings to the "Calvinist ethic." Compositional change may also help us to understand evolving demographic behavior. Fertility rates and age at first marriage often differ across subgroups, as both historians and economists have found.[27]

That natural selection improved in some definable dimension the quality of the population in the countries about to break out of the Malthusian model before the 1700s is still far from an established fact. Disentangling "inherent quality" changes from responses to new incentives seems a formidable challenge. Given that humans normally only start to reproduce in their late teens or early twenties, any process that relies on natural selection requires a very long time span – or strongly divergent fertility rates.[28] We know far too little about

[27] In addition to Clark and Hamilton (2006), one should mention the work by Herlihy (1997, pp. 56–57) and Galor and Moav (2002).

[28] Given that the earliest data are from the sixteenth century, there were only approximately five to six generations over which we can be reasonably certain that this selection effect might have worked – not a great length, given the modest reproductive advantage. All the same, recent genetic research has suggested that "evolutionary changes in the genome could explain cultural traits that last over many generations as societies adapted to different pressures" (*New York Times*, March 7 and 12, 2006).

the relative differences in reproductive behavior (as manifested, for example, in different marriage ages) and economic success in early modern Europe. Compositional change may have played a role, but at the current stage it is hard to tell. Grander claims about the prevalence of "survival of the richest" in Europe, and its absence in the Far East (Clark 2007a) rest on shakier foundations. What is needed is more evidence along the lines of the material gathered by Clark and Hamilton documenting differential fertility and survival over the long run, and in different parts of the globe.

Scholars in the field of cultural evolution such as Boyd and Richerson (1985, 2005) point to the fact that culture changes only in small part due to the compositional effects of natural section. Instead, they point to learning and imitation as sources of cultural change. People receive part of their culture from their parents, whether through genes or education, but they are exposed throughout their life to other influences that may make them different from their parents. Such "biases" may have different forms, but here perhaps the most interesting is what they call "model-based bias" in which individuals observe certain other persons who have an attribute they regard as desirable (e.g. social status or wealth), and thus choose this other person as their role model. In a highly stratified yet mobile society such as Britain, the incentives to imitate the behavior of others viewed as higher-up in the hierarchy were strong. This would make the most successful workers and artisans try to imitate the bourgeoisie, by mimicking their behavior. The growth of a middle class would thus be far faster than differential reproduction would predict.

Technology

Technological change has remained the backbone of modern economic growth, simply because all other potential sources of productivity growth tend to run into diminishing returns. Capital accumulation, improved allocation of resources, gains from trade, better institutions, and economies of scale will increase output, but they are eventually subject to diminishing returns. No historically accurate picture of modern growth can be formed without understanding the connection between science and technology in the Industrial Revolution and beyond. Historical scholarship has bifurcated into a minority view, which continues to view science and scientific culture as crucial to the Industrial Revolution (Musson and Robinson, 1969; Rostow, 1975; Jacob, 1997; Lipsey, Carlaw, and Bekar, 2005), and a majority, which has dismissed the role of science as epiphenomenal and marginal (Landes, 1969; Mathias, 1979; Hall, 1974; Gillispie, 1980). Examples of the importance of science and mathematics to some of the inventions of the Industrial Revolution can certainly be amassed. It is equally true, however, that many of the most prominent breakthroughs in manufacturing, especially in the mechanical processing of textiles, were not based on much more science than Archimedes knew. In other areas of progress, such as steam power, pottery, and animal breeding, advances occurred primarily on the basis of trial and error, not a deep understanding of the underlying physical and biological processes.

The debate between those who feel that science played a pivotal role in the Industrial Revolution and their opponents is more than just a dispute about whether a glass is half full or half empty. The glass started from almost empty and slowly filled in the century and a half after 1750. Scientists and science (not quite the same thing) had a few spectacular successes in developing new production techniques, above all the chlorine bleaching technique, the soda-making process, and the inventions made by such natural philosophers as Franklin, Priestley, Davy, and Rumford. While the Industrial Revolution in its classical form might well have occurred, with a few exceptions, without much progress in science, it is hard to argue that it would have transformed into a continent-wide process of continuing growth without a growing body of useful knowledge on which inventors and technicians could draw. It is not possible to "date" the time at which this kind of collaboration began. In some areas it can already be discerned in the mid-eighteenth century. In crucial "new" areas of technology in the post-1820 years, scientific knowledge became increasingly important to the development of new technology. Two of the most remarkable developments of the era, the electric telegraph and the growing understanding of the chemistry of fatty acids used in soap and candle manufacturing, took place in the final decades of the classical Industrial Revolution. Trial and error, serendipity, and sheer intuition never quite disappeared from the scene, but improved knowledge about how and why a technique works made it far easier to refine and debug a new technique quickly, adapt it to other uses, and design variations and recombinations that would not have occurred otherwise. In chemicals, steel, electricity, food processing, power engineering, agriculture, and shipbuilding technology, the ties between formally educated people who tried to understand the natural phenomena and regularities they observed, and the people whose livelihood depended on putting such insights to good use, became tighter and closer after 1750, and continued to do so (Mokyr, 2002).

The underlying institutions that made this growing collaboration possible have been investigated at length. Although intellectual property rights were of some importance, they cannot possibly explain the entire process. Jones (2001) is the only growth paper to date that models time-varying institutional parameters directly. They turn out to play a pivotal role in his model regarding whether the Industrial Revolution was "inevitable."[29] Jones's parameter π_t, which is the proportion of total consumption allocated to people employed in the ideas-generating sector, is computed to match the data. Broadly speaking, the rise of resource-intensive research and development is captured correctedly. At higher frequencies, the series displays a bizarre history (Jones, 2001, p. 24), falling from 0.44 percent to zero between the sixteenth and the eighteenth century, rising sharply in the eighteenth century, then falling to half that value in the nineteenth century, before leaping by a factor of 12 to 5 percent in the twentieth century. Nothing in Jones's model allows for the complex motivation that propelled the ideas sector in earlier history, when many natural philosophers and inventors were as much interested in signalling as in financial gains, much like a modern open-source technology (Lerner and Tirole, 2004). Models that purport to explain the growth of technology in this age must recognize the different ways of assigning property rights in the two separate segments of the "ideas sector." Whereas prescriptive knowledge, that is, techniques, could be patented and thus be allocated at least some form of property rights, this was never done with propositional knowledge in which priority credit assigned to the owner did not include exclusionary rights. Yet it is hard to understand the growth of technology during the Industrial Revolution and after without explicitly recognizing the feedback between these two forms of knowledge (Mokyr, 2002; see also Dasgupta and David, 1994). Scientists in the pre-1850 period were rarely interested in reaping the material gains that their findings could generate, insisting on credit rather than profit. As "gentlemen-philosophers" they refused to make a living from their discoveries and were suspicious of anyone who did (Bowler and Morus, 2005, pp. 320–21).

It is also worth noting that a recent attempt to estimate the value of inventions accruing to the inventors for modern America has found that only about 2.2 percent of the value of an innovation is captured by the inventor him- or herself (Nordhaus, 2004). Whether the number was higher during the Industrial Revolution seems unlikely. The patent system is central to this story, but its effect on the process of technological progress during the Industrial Revolution is still very much in dispute. The operation of the patent system awarded monopolies to inventors, yet infringements and other failures of the system implied that first-mover advantages and old-fashioned government and private-sector prizes were as important as the rents earned by inventors.[30] The British patent system was far from user-friendly: it was costly to file a patent and often hard to defend patents against infringers (Khan and

[29] The parameter π in Jones's model defines the proportion of total income that accrues to those who are employed in the "ideas sector," and in equilibrium equals the fraction of labor in the economy allocated to producing new ideas.

[30] The literature on the operation of the patent system in Britain is quite large; for an introduction see Dutton (1984), MacLeod (1988) and MacLeod and Nuvolari (2007).

Sokoloff, 1998; Dutton, 1984). The patent laws were widely condemned as ineffective in protecting the vast majority of inventors, and did so at a high cost (MacLeod and Nuvolari, 2007). Most inventions, even the most successful ones, were not patented (Moser, 2005). The fact that Britain's system was thus less likely to encourage potential inventors than the corresponding US system does not seem to have affected British technological leadership before 1850. Charles Babbage, never one to mince words, denounced the patent law as "a fraudulent lottery which gives its blanks to genius and its prizes to knaves" (1830, pp. 333, 321). Inducement prizes for inventions, such as those offered by Parliament, may have been at least as effective in generating new ideas as protection by the patent system (Brunt, Lerner and Nicholas 2008). The best we can say about the patent system is that it provided an ex ante belief that successful invention *could* pay off to a few lucky people, and thus provided a positive incentive.

Instead, a deeper and more encompassing social phenomenon was at play here, namely growing flows of information and improving interaction between people who made things (entrepreneurs and engineers) and people who knew things (natural philosophers). Not only did this interaction mean that what useful knowledge had to offer was accessible to those who could make best use of it; it also meant that the agenda of science was increasingly biased toward the practical needs of the economy. The bridges between *savants* and *fabricants* took many forms, from written technical manuals and treatises to academies and scientific societies, where they rubbed shoulders and exchanged ideas. By the closing decades of the eighteenth century it was normal for scientists to consult to manufacturers and farmers looking for improved bleaches, more efficient engines, or improved fertilizers.

By 1815, the need for this kind of collaboration had become a consensus, and European economies competed with one another in encouraging it. In Britain, the Society of Arts, established in 1764, the Royal Institution, founded in 1799, and the Mechanics Institutes (first established by George Birkbeck in 1804) were examples of how private initiatives could carry out this task in the land where people believed above all in private initiatives. Less formal institutions abounded, the most famous of all being the Birmingham Lunar Society, which brought together the top scientists with some of its most prominent entrepreneurs and engineers. Less well known but equally significant were the Spitalfields Mathematical Society, founded in 1717, and the London Chapter Coffee House, the favorite of the fellows of the Royal Society in the 1780s, where learned men discussed at great length the mundane issues of steam and chemistry (Levere and Turner, 2002). In France, Germany, and the Low Countries, government took a more active role in bringing this about (see, e.g., Lenoir, 1998). Not all of those efforts were unqualified successes: the engineers of the Paris École Polytechnique were often too abstract and formal in their research to yield immediate results. In Germany, the university system was on the whole rather conservative and resisted the practical applications that governments expected of them. New and more effective institutions were established, however, and the old ones eventually reformed.[31] The decades after 1815, then, were the ultimate triumph of the Baconian vision, which had formed the basis for the founding of the Royal Society in 1660. To achieve this triumph Europe had to undergo changes in its institutional set-up for the accumulation and dissemination of useful knowledge, yet these institutions were based on the scaffolds (to use North's term) of an Enlightenment ideology that firmly believed in material progress and advocated concrete programs as to how to bring this about.

Modelling the production of "new ideas" is a key challenge for growth models, and endogenous growth models have had to simplify away much of the historical richness. Thus the literature has not dealt effectively with the high riskiness of the inventive process, in which investing in the "ideas-producing" sector is more akin to purchasing a lottery ticket than to choosing an occupation.[32] While some models refer explicitly to "the number of ideas produced," such a concept is of course highly problematic, not only because ideas fail to meet the rules of arithmetic, but also because so many ideas generated were simply dead-ends, mistakes, or even pure fantasy. On the other hand, much of the new technology was the result of minor but cumulative improvements resulting from the experience and learning-by-doing of skilled craftsmen, rather than some kind of cognitive flash. While such artisanal advances were not all there was to technological change during the Industrial Revolution, their importance has been rightly stressed by historians (Berg, 2007).

Conclusion: progress out of misunderstandings

Theorists and economic historians interested in the transition to self-sustaining growth often appear as distinct tribes, separated by a common object of study. This has hampered progress in understanding how the switch from "Malthus to Solow" occurred. We highlight a few particular sources of misunderstanding, and offer suggestions for future research that should do much to reconcile the tribes and augment intellectual gains from trade.

[31] In Germany, universities had increasingly to compete with the technical colleges or *technische Hochschule*, the first of which was set up in Karlsruhe in 1825. In France new *grandes écoles* were set up to provide more practical education, such as the Arts et Métiers.

[32] Indeed, this aspect of technological progress may well be better analyzed by behavioral economists and decision theorists, who deal with models in which people systematically overestimate their own chances at succeeding. This was already understood by Adam Smith, who noted, "Their absurd presumption of their own good fortune is … still more universal [than people's overestimating their own abilities] … the chance of gain is by every man more or less over-valued, and the chance of loss … undervalued" (Smith, 1996 [1776], p. 120).

Economic theorists writing on long-run growth often apply their models to industrializing Britain, the classic, first case of an Industrial Revolution. This forces them to ignore or to play down inconvenient facts about economic history "wie es wirklich gewesen ist." Jones (2001), for example, produces a model in which working hours have to fall during the transition, while they probably rose in actual fact. Models in the Lucas and Becker tradition emphasize the increasing demand for and returns to human capital, when we find little evidence of this. Acemoglu, Johnson, and Robinson underline the importance of constraints on the executive in early modern Europe, when it is far from clear that the groups producing serious restraints had anything to offer for growth. The list could be extended, but it is mainly meant to be illustrative.

Economic historians have been quick to point out the most glaring contradictions, pointing out that the "Industrial Revolution in most growth models shares few similarities with the economic events unfolding in England in the 18th century" (Voth 2003). We believe that the discussion should not stop here. The logic of many unified growth models makes them less than well suited as an explanatory toolkit for the classic British Industrial Revolution, despite the tendency of growth theorists to apply their models to that case (Galor, 2005; Hansen and Prescott, 2002; Lucas, 2002). When applied to the early modern period as a whole, unified growth models have much more to offer. Conversely, economic historians have been too narrow in many of their criticisms. Once we lengthen the period during which demographic changes are analyzed, examine human capital accumulation over the very long run, and broaden the relevant set of skills beyond literacy, as traditionally emphasized, many of the apparent contradictions are diminished. The Malthusian regime slowly broke down during the years 1500–1800, and many relevant changes in human capital probably began after the Reformation. In the same vein, once economic historians and theorists focus more on non-cognitive skills and cultural features such as patience, prudence, and discipline, it will be easier to build models that are broadly consistent with the historical record. In this light, unified growth theories have substantial explanatory power – especially those that emphasize a transition to modern growth in two phases (such as Galor and Moav 2002), with Malthusian constraints declining first, and human capital becoming more important later.

The time period under consideration is not the only important source of misunderstanding. As the German philosophers Rickert and Windelband emphasized, history is 'ideographic' – seeking to explain what is unique. Theorists are, of necessity, 'nomothetic' – in search of covering laws. While the discussion of the British Industrial Revolution in, for example, Hansen and Prescott (2002) is meant as an illustration of a model that could apply to Europe (or, indeed, the world) as a whole, historians have often focused on empirical accuracy in individual cases. Equally problematic is the tendency to examine the logic of historical papers from a cross-sectional perspective. Economic historians have rarely resisted the temptation to demand predictive power for our own time's distribution of economic development from models designed for the crucial question of explaining "why Britain came first," and not France or China (Crafts, 1995; Broadberry, 2007). The models in Kremer (1993) and Galor and Weil (2000) apply to the world as a whole. Nonetheless, economic historians have criticized endogenous growth models because they fail to offer convincing explanations for income divergence between countries. This is clearly a case of hunting rhino with sharpened kiwis.

However, and by the same token, theorists have not given much attention to the important implications of cross-sectional differences in the timing of growth spurts.[33] For example, Hall and Jones (1999) document large differences in output per capita between rich and poor countries. They conclude that neither capital nor human capital can explain these differences; total factor productivity and "social capacity" must be responsible. The underlying models only make sense if we assume that economies have reached their steady state, or that they should have the means to converge to it rapidly. Most papers in the growth literature using the Summers-Heston dataset share this assumption. Yet when we examine the broad sweep of history, one of the most striking observations is the sheer time scale necessary for the escape from the Malthusian world. Even over the last 200 years the "take-off" into self-sustaining growth (some sad reversals such as Argentina apart) has occurred at different points in time in different countries. These differences in timing have proven hard to explain. A delayed start to economic growth will lead to an inverse U-shaped pattern of productivity differences over time. Relaxing the assumption that economies are in steady state, and focusing on what allows economies to enter the phase of rapid, self-sustaining growth, resolves some key puzzles in the current growth literature (Ngai, 2004). Relatively small inefficiencies can produce large differences in output per head if they *delay* the onset of modern growth. What is needed, then, is a set of theories modelling dynamics: what was behind phenomena such as timing, lags, and the long historical delays between prior historical changes and the onset of modern growth. This implies that theorists may continue to rely on cross-sectional evidence about divergent growth paths as inspiration for their models, but that economic history offers far greater riches. A closer collaboration between those who want to discern general laws and those who have studied the historical facts and data closely may have a high payoff. Only when we understand which inefficiencies and delaying influences produce the time pattern of "take-offs" that we observe during the last two centuries can we begin to claim that our understanding is as complete as the term "unified growth" theory suggests.

[33] One exception is recent work by Voigtländer and Voth (2006, 2008), who offer models designed to explain part of the "First Great Divergence" between Europe and China.