Understanding Growth in Europe, 1700-1870: Theory and Evidence

Joel Mokyr
Depts. of Economics and History
Northwestern University
Berglas School of Economics
Tel Aviv University

Hans-Joachim Voth
Department of Economics
Pompeu Fabra University

In one witty formulation, God gave the easy problems to the physicists (Lebow et al., 2000)

Tentative outline

1. Introduction (incomplete and very preliminary).

a. Can theory help the historians?

In recent years, growth theory has turned to the Industrial Revolution and the emergence of modern economic growth as a topic of research. For economic historians, for whom these issues have been their bread and butter for over a century, this is a cause for jubilation. The issues are complex and difficult, and they need all the help they can get. What better than from the likes of Robert Lucas, Edward Prescott, Oded Galor and their colleagues, who had previously developed a set of novel and highly sophisticated and influential tools to analyze hugely complex phenomena such as the business cycles and the effects of monetary policy in elegant and powerful models?

Theory has to simplify reality and to make assumptions. It tries to establish causal connections between exogenous and endogenous variables, to establish equilibria and the trajectories that a model of the economy will follow on its way there. What is being explained here is, at some level, rather straightforward. In the early nineteenth century, income per capita and a set of related variables, started to increase dramatically in a small number of economies in the northern Atlantic and European offshoots. In his long and detailed survey, Galor (2005, p. 177) raises the main questions that he feels need to be answered, such as why there was so little growth before the great takeoff, why previous technological advances had not resulted in similar growth processes, and what the connection is between demographic changes and the growth spurts. Economic historians have raised other issues that need to be addressed here, and which play a lesser role in the models proposed by growth theorists: what was the role of formal and informal science (propositional knowledge) in triggering the growth spurts? what advantages did Britain possess that awarded it a leadership role, however ephemeral? what role did formal and informal institutions such as government and independent NGO’s play in the process? what importance did colonial ventures and overseas trade, both short- and long-distance have in the process? how did the interaction between traditional sectors (agriculture and domestic industry) and the modern sector matter? Theory, it seems to us, can help us answer all of those questions by focusing on the variables that mattered and by pointing out likely and less likely causal connections, as well as by adding precision to the analysis.

But human history is far more complex than natural phenomena.¹ Theorists, by stripping away parts of the problem, are presenting an important and much improved way to solve the problem, but their models are not the magic wand that will solve the economic historian’s difficulty. There are basically four reasons why unbridled optimism about the use of economic models in

---

¹In one witty formulation, God gave the easy problems to the physicists (Lebow et al., 2000)
explaining the questions are probably out of place. One is the role of unobservables and intangibles. As economists have become increasingly convinced of the importance of culture in explaining economic growth (e.g. Tabellini, 2006; Guiso, Sapienza and Zingales, 2005), it is clear that our ability to understand large-scale historical events in single models involving measurable variables is somewhat limited. Second, economic growth consists of both shifts of the production function (that is higher productivity due to technological progress) and shifts along a production function (capital accumulation and changes in allocation that increase output within a given set of techniques). The composition of growth between those is, however, unknown and the interaction effects between them make the story even more complex. Third, to the extent that the ultimate process is driven by technological progress, our understanding of economic growth will only be as good as our understanding of what really drives technology. To do so, we need a much improved understanding of the social and cognitive aspects of useful knowledge, the institutions generating and disseminating innovation, and the political struggles between innovative entrepreneurs and the incumbents (Mokyr, 2002). While progress has been made on these matters, there is still much to be learned. Finally, there is the lack of observations: the great transition succeeded only once. To be sure, there is a fair amount of cross-sectional variation between economies and regions that has been exploited a great deal, but there was only one Industrial Revolution, and if that phenomenon is to be taken — as it should — as the origin of modern economic growth, it really boils down to just one observation.

b. Unified models: pros and cons

Do we need a single theory to explain everything? Scientists still keep hoping to come up with a realization of Einstein’s dream of a single TOE (Theory Of Everything). This has remained the holy grail of physics, and while it may one day be realized, it has not to date, and not for lack of trying. In biology, too, we do not have anything that resembles a TOE, Darwinian models notwithstanding, and it remains highly questionable if one will ever be found. If so, how likely is it that one will ever be found in economic history?

Perhaps the search is all that matters. Throughout the history of science, the world of knowledge has been enriched by both “hedgehogs” looking for a single TOE and “foxes,” who were content to deal with smaller problems. It is fair to say that economic history has been mostly a land of foxes. The last (and possibly the only) great hedgehog was Karl Marx, whose TOE shaped the field for many decades, but it was not really geared to deal with economic growth, and had next to nothing to say about changes in demography and population dynamics. Yet an argument could be made that there is a fruitful conversation between hedgehogs who see a “big picture” and point to basic trends, and foxes who work on details, listen patiently to the hedgehogs and their grand theories, and then raise hard questions of detail, reminding the theorists of the complexities that are always greater than the models can deal with. Yet without the hedgehogs the work of foxes would possibly drown in a mind-deadening ocean of detail, without motivation or direction. It is this kind of complementarity that makes the scholarly conversation interesting and useful.

Yet we need more than that. Economic growth is the most significant event in modern history, and must be counted as one of the truly significant turning points in history, comparable to the emergence of homo sapiens or the rise of Christianity in the West. Understanding it, in the views of many, holds the key to the economic fate of humanity. As Robert Lucas wrote in 1988, once you
start thinking about this, it is hard to think of anything else. In this search, a single “unified” theory of economic growth and demographic change is unlikely to satisfy the empiricists and details people who are immersed in the fine details of how the patent office really encouraged or discouraged innovation or who worry about the interactions between smallpox vaccination and other infectious diseases.

Unified growth is supposed to “unveil the underlying microfoundations consistent with the entire process of economic development” (Galor, 2005, p. 219). But at what level? Presumably these microfoundations must not only point to the human capital accumulation and the demographic changes that accompanied development, but also point out how they did so. This question is still quite far from being answered. The development of dynamic models with latent state variables in which both the location and the stability of steady-state equilibria changes as a result of population change, while ingenious, seems thus far less than historically verifiable. While they can reproduce a number of historically observed phenomena (such as the discontinuities in output per capita and non-monotonicities in fertility) they are far from the only way to account for these phenomena and they leave little room for institutions and beliefs to determine much except second-order differences in timing. Secondly, unified growth is supposed to account for both pre-modern “stagnation” and the transition into sustained growth. Theories that account for just one of those periods are deemed to be ad hoc. But methodologically we could have had a regime change or “phase transition” as physicists call it, and separate accounts of the different regimes, so long as we could also explain the regime change itself. Above all, unified theories tend to contain an element of inevitability and “hindsight bias” that historians may find somewhat disconcerting. The potential for growth was already present in latent form in traditional society and its emergence was just a matter of time. This element of TOE is probably unattractive to historians who, like the biologists, know the importance of contingency and accidental factors (“path dependence” if one wishes) in determining final outcomes.

At the same time, however, the hedgehogs not only force the foxes to re-think their assumptions, they point to those issues that researchers should be working on. In this particular case, it is the connection between technological progress, the formation of human capital, and demographic change that has been at the center of attention thanks to work in unified growth theory (Lucas, 2002, Galor, 2005). Specifically, a great deal of emphasis has been placed on drawing connections between growth, technology, fertility, and the investment in education. Economic historians would clearly deny that this is all there is to the transformation, but by facing models that imply that it was, they are forced to rethink positions and dig for more data or examine afresh what they already have, and thus our understanding is advanced.

c. The riddles of growth: can we avoid Eurocentricity?

Historians have increasingly developed an antipathy to what they think of as “Eurocentric” history, that is the kind of history that places the West on the pedestal of the successful economic model of economic growth that other parts of the world needed to emulate. Such a history inevitably asks questions such as “why did China (or India of Africa) fail to be like the West.” An entire literature has sprung up debunking this approach. Yet when all is said and done, by 1914 a gap had opened up between a club of economies that had achieved a high income per capita, and the
correlates that came with that: longer life expectancy, more comfortable daily existence, and the kind of military potential that made Western dominance possible.

And yet it is arguable that the question why did China fail to develop is, indeed, illegitimate because what needs to be explained is not what failed to happen in China, India, Africa or the Middle East, but the European Miracle (Jones, 1981). In this interpretation, the evolution of the rich and industrialized economies in the West was a highly unlikely event, the result of a fortunate concatenation of circumstances. Unified growth theory is willing to concede that accident may determine the timing of economic growth but that the event itself was wholly preordained from the day of creation. But how can we be sure? Many things could have wrong in the European experience, starting with military events (e.g., the failure of the Mongols to devastate Europe after the battle of Legnitz in 1241), the fact that the Black Death killed “only” a third of the population but left the rest alive (unlike the demographic devastation of the indigenous populations of America after 1492), the failure of the counter-reformation to suppress the reformation, and the rise of a free and competitive market for ideas in Europe (Mokyr 2006). The origins of economic growth in Europe, in this interpretation, far from being pre-ordained in the inevitabilities of unified growth theory, are a fluke of history. Once it happened, however, its effects on other parts of humanity were ineluctable. Whether one buys this interpretation or not may be a matter of taste; but there are few tests we can bring to bear to discriminate between it and the models of economic growth that imply that the roots of economic growth are to be found in European history long before it actually blossomed.

Either way, it is hard to avoid the fact that the history of economic growth is Eurocentric. This raises, of course, the question what it was about Europe that gave rise to the phenomenon. The answer is hopelessly overidentified: we have but one event, and yet we are facing a huge range of answers, from the silly (Christianity was the only religion that was suitable to economic growth, as in Stark, 2000) to the geographically deterministic (Jones, 1981), to the superiority of European culture (Landes 1997). Yet, to date, growth theory has been of little help in answering that question.

2. The pre-industrial economy

Many formal models of historical growth assume that before 1800, there was no or negligible long-term growth (e.g. Hansen and Prescott, 2002, p. 1205; Galor and Weil, 1999, p. 150; Galor, 2005, p. 180). Some economic historians share this view (e.g. Clark, 2007, ch. 8-1). At some level, this statement is an oversimplification: there was far more dynamism in the pre-modern economies than is supposed by theorists. There can be little doubt that in Gregory King’s day (1688), income per capita in Britain was substantially higher than it had been at the time of William the Conqueror (Snooks, 1994). The growth rate may not have been high, but compounded over centuries, it changed historical reality. While growth was not yet self-sustaining, living standards in Europe were

---

2Adam Smith had no doubt that “the annual produce of the land and labour of England... is certainly much greater than it was a little more than century ago at the restoration of Charles II (1660)... and [it] was certainly much greater at the restoration than we can suppose it to have been a hundred years before” (Smith, 1776-1976, pp. 365-66).
not at “subsistence levels” for most parts of the population, or even close. Conventionally measured growth rates may be revised upward if they are computed in a more encompassing way. New products became available after 1600, and conventional measures of GDP tend to ignore both the appearance of goods not previously known, the value of variety, and quality improvements.

All the same, growth before 1750 was, if not totally absent, different in nature from what was to occur in the nineteenth century. Degree is everything in economic history, and a rate of growth of 0.1% or 0.2% is a very different phenomenon than one of 1.0%. Moreover, it was, on average, far less steady than it was to become after 1850, with periods of fairly rapid medieval expansion being punctuated with sharp and even disastrous declines caused by epidemics, wars, famines, and climatic events. Just as in the developing world today, slow growth was also volatile growth (Acemoglu and Zilibotti 1997). Furthermore, pre-industrial growth was local or at most regional: some towns and small regions such as London, Antwerp, the maritime provinces of Holland, Southern Germany, Venice, and Tuscany experienced periods of high prosperity reflecting earlier growth; yet growth at the level of “the economy” was rare, and perhaps is a concept that should not apply, Adam Smith notwithstanding.

In pre-industrial economies, even when growth occurred, it typically led to forces that eventually extinguished it. Because of these “negative feedback effects”, output increased in spurts — what Goldstone (2002) calls efflorescences. In contrast to similar episodes in the post-1820 period, they did not last. One reason for negative feedback was the Malthusian mechanism, in which rising per capita income produces population pressure that eventually forces wages back down. An alternative explanation focuses on negative institutional feedback, with growth leading to an upsurge in the influence of rent-seeking that eventually causes stagnation.

a. The Malthusian model: theory and evidence

The economic history of the pre-1750 world is often referred to as the Malthusian epoch. It is characterized by a model that relies on two assumptions: one is that, some exogenous shocks excepted, population growth was governed by movements of income per capita. The other is that income per capita was negatively related to population size. Together, these two assumptions imply that whatever advances might be achieved by technological progress, capital accumulation, or gains from trade, they will inevitably be frittered away through the birth and survival of more babies. In figure 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 shock 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. In the more extreme versions of the model, this meant that little that could happen to these economies made much difference to wages or population size — demographic responses pin down the long-run equilibrium. As Clark (2007, ch. 2, pp. 22, 29) puts it, the Malthusian model implies that in the pre-industrial world tech-

---

1 Fogel (1990, 1994) argues that the bottom 20% of the population in France (and no group of society in Britain) were too short of nutrients to work long hours. The remainder enjoyed ample consumption opportunities.

4In H.G. Wells’s utopian novel, *Men Like Gods* Utopia in the past “spent the great gifts of science as rapidly as it got them in a mere insensate multiplication of the common life. At one time in the Last Age of Confusion the population of Utopia had mounted to over two thousand million…” (ch. 5 part 4).
Lucas (2002, pp. 14-15) describes how he came to see the empirical power of this theory as its ability to account for the similarity in real incomes across different societies and the stability of living standards over time in the face of ongoing technological change. Others who buy into this model, in one version or another, include... 

Galor (pp. 183-184) shows some graphs that indicate that in pre-industrial Britain population and real wages moved roughly in opposite direction and that crude birth rates and crude death rates were negatively correlated.

Figure 1: A Malthusian model

Technological progress produced “people, not wealth” and that “good government could not make people rich.” The influence of this model on modern theorizing on the pre-industrial economy has been substantial. Galor (2005, p. 180) acknowledges that there were substantial improvements in per capita income during the Malthusian period, but that accelerating population growth eventually reversed all gains.

Is the Malthusian model a good characterization of the pre-modern era? In the end, this is a matter of the historical evidence. Over the short run, movements in population before 1750 seem to offer resounding support for the model. The crucial idea here is that some variables like population size are invariably slow-moving – and so will the real wage be. Yet mortality and nuptiality can adjust even over the short run. High-frequency events like famines, wars and epidemics had much smaller long-term effects than the disasters would suggest: 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;  

---

5 Lucas (2002, pp. 14-15) describes how he came to see the empirical power of this theory as its ability to account for the similarity in real incomes across different societies and the stability of living standards over time in the face of ongoing technological change. Others who buy into this model, in one version or another, include...

6 Galor (pp. 183-184) shows some graphs that indicate that in pre-industrial Britain population and real wages moved roughly in opposite direction and that crude birth rates and crude death rates were negatively correlated.
As Wrigley and Schofield point out, even if “for the individual it was better to be wealthy than to be poor if one
wished to live long and be untroubles by infection... for society as a whole the balance of advantage was harder to strike.
But the raw materials underlying these figures are imperfect, and there are serious conceptual and econometric difficulties in testing the model. The real wage data computed by Clark (2005, p. 1311) replaces the traditional wage series computed by Phelps-Brown and Hopkins and is based on
a much broader array of commodities and a more comprehensive set of nominal wages. Both show
the same, rather miraculous sharp decline of wages in Tudor England between 1495 and 1575, a
decline that is accompanied by stable and then rising population as well as an unusually long life
expectancy.

An important problem in confronting the Malthusian model with the data is endogeneity. Wages influence population size and vice versa. Without some source of exogenous identification, there is no hope of pinning down the size and importance of relationships. Recent work by Kelly
(2005) suggests that weather may be a useful instrument for wages – the part of 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 (negatively) to wages 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. However, we would need to estimate a full model, with feedback from population to
wages, to really confirm this finding. Cross-sectional evidence that richer Englishmen had more
surviving offspring also appears to attest to the strength of Malthusian forces in early modern
England (Clark and Hamilton, 2005).7

However, Malthusian constraints probably mattered more over the short- than the long run. Short-run demographic adjustments to real-income shocks may not imply that the “iron law of
wages” held true. It is perfectly possible for the system to respond to high grain prices with
increased mortality and reduced fertility, and yet not to be in a long-term stasis. In figure 1, the
functions may be shifting so much over the longer term. At some point, if the death schedules, birth
schedules, and technology function shift enough, Malthusian factors will no longer be prime
determinants of living standards – even if short-run fluctuations seem to suggest that this is so.
Population was affected by long-term changes in background mortality, driven through mechanisms
we only understand very partially as a result of exogenous variations in disease environment and
climatic changes (Goldstone, 1991). If we are to believe the Maddison figures, all European
economies in 1700 were both more populous and richer per head than they had been in 1500. One
of the most spectacular cases is the Netherlands, where income per capita started to increase after
1500 with a corresponding population increase that lasted till 1650. But the population stabilized
around 1650 with income per capita being at a level that in no way can be regarded as subsistence.
De Vries and V.D. Woude (1997, p. 688) note that “by the end of the [sixteenth] century a

---

7As Wrigley and Schofield point out, even if “for the individual it was better to be wealthy than to be poor if one
wished to live long and be untroubles by infection... for society as a whole the balance of advantage was harder to strike.
Increasing wealth bore an ambiguous relationship to improved mortality” (1997, p. 206). Yet not all cross-sectional evidence is
consistent with a positive relationship between income and rapid population growth. In 1954, G. Utterström showed that poor
and remote areas in Sweden nonetheless had lower death rates than the richer areas.
Oddly enough, the richest economy in Europe is also one of the last to start experiencing the population growth that commenced elsewhere around 1750; in the Netherlands this did not begin until about 1815 (Hofstee, 1978). The conventional Malthusian model is no longer adequate to account for the economic and demographic dynamics... demographic forces interacted with the economy in ways far more complex than can be comprehended within a Malthusian model.\footnote{Oddly enough, the richest economy in Europe is also one of the last to start experiencing the population growth that commenced elsewhere around 1750; in the Netherlands this did not begin until about 1815 (Hofstee, 1978).}

It may well be that, while the Malthusian adjustment mechanisms held in the short run, the interesting shifts were caused by other factors. Rising income has traditionally been associated with increased urbanization, to the point where the proportion urban has been taken to be a proxy for income per capita. Cities, however, experienced very different demographic parameters than the countryside, and had far higher mortality rates. Hence the curve D, which is a composite of rural and urban demographic behavior, could slope upwards over some part of the w-D space because of this composition effect. There could then be multiple equilibria: societies could fluctuate between one state where population was large, wages were low, cities small, and aggregate death rates low, and another one where wages were higher, cities larger, death rates higher, and population smaller. Cities were the locus for much of the inventive activities of the time – the slow, gradual improvements of machinery, of social organizations and the like. All of this may have improved the economy’s ability to sustain more people at the same level of per capita income over the long run. 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. In this case, Malthusian forces could still dominate short-run changes, but the key \textit{explanandum} would no longer follow from its basic tenets.

Some of the unified models also predict (modestly) rising living standards before the Industrial Revolution. This is because of “the inherent delay in the adjustment of population to the rise in income per capita, generated positive but very small growth rates of output per capita” (Galor 2005). In models such as Jones (2001), there is a similar, delayed response of population to technological advances. Given that total fertility rates for females in many pre-modern populations (and especially European ones) were substantially below their biological maximum, this is an unconvincing mechanism to explain why living standards drifted up, albeit slowly, in the centuries before 1800. Where population growth depended on economic conditions, such as in the relatively “low pressure” demographic regimes characterized by the European marriage pattern, it remains unexplained why, over time, fertility would have been curtailed at progressively higher and higher levels of income – and that these restrictions broke down from the 1750s onwards.

In short, a substantial amount of evidence points to problems with the blind acceptance of the Malthusian model as the description of preindustrial Europe. While Malthus probably characterized some early part of pre-modern growth correctly, it is unclear how much the model applied by 1700. The theoretical literature that advances this notion must be concerned with the rather obvious cases in which Malthusian models did not hold or ceased to hold before the Industrial Revolution. What Malthusian models involve is negative feedback: growth leads to higher nuptiality and fertility and lower mortality, thus higher population, which leads to lower incomes. There is unanimity that this model is — roughly speaking — consistent with the absence of sustained growth in (almost) all pre-industrial societies. But to convince us to take the Malthusian model as an accurate description, it needs to be tested against alternatives.

\textbf{b. Institutions and growth: an alternative to Malthus?}
One alternative mechanism that generates reversals of fortune is institutional feedback (Mokyr, 2005). The particular idea is part of a broader class of models that sees political, legal and social factors as prime determinants of long-run growth. Rent-seeking can produce negative feedback, with prosperity leading to a host of institutional effects that end up terminating the effects of technological progress or commercial expansion. The pre-modern economies faced a particular dilemma, namely that the accumulation of wealth required on the one hand a strong government to protect it from foreign rent-seekers, such as invasions, strict mercantilist policies, or pirates. Moreover, as Epstein (2000) has argued, centralized nation states solved certain coordination problems that societies with heavily dispersed powers could not resolve, without which modern markets could not have evolved.

Yet a powerful government itself could be the biggest rent-seeker or worse, farm out its taxes to rent-seekers, break contracts, and seize assets. Ideally, a government should be strong enough to protect trade and property from foreign invasion, yet constrained in what it could do to its own citizens. This kind of combination did not emerge until the eighteenth century. Until then, rent-seeking (from abroad or from domestic rulers) time and again reversed the fortunes of those regions or towns that had managed to accumulate significant wealth. Examples of negative institutional feedback are not hard to find, and until we know more about their relative importance, accepting the notion that the economic outcomes before 1750 were mainly governed by Malthusian forces seems at least incomplete. Thus in early modern Europe, less developed but large political units, such as the young nation states of Philip II and Louis XIV, threatened the richer but smaller city states of Italy and the Low Countries. This military imbalance created a basic source of instability and inefficiency in the history of European cities. 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. The only two areas that escape this fate enjoyed unusual geographical advantages for repelling foreign invasions – Britain and the Netherlands.

If the institutional feedback mechanism turns out to have been important in the pre-1750 world, it would shed a very different light on the emergence of modern economic growth and its roots. Much of the modern theoretical and literature assumes that what has to be explained is the transition from a Malthusian to a post-Malthusian regime. In that story, demographic and technological elements are modeled in various ways we shall see below. However, if the constraints on growth before 1800 were as much institutional as they were demographic, the story will have to be amended in important ways.

Most interpretations of early modern Europe do no focus on this negative feedback mechanism to explain the intermittent nature of growth. Instead, the story is one of constraints that falling away first in some parts of Europe, then in others. The single best-known statement in this tradition was formulated by North and Weingast (1989). They argued that the Glorious Revolution and the Bill of Rights in England did more than put government finances on a firm footing. Because of the changes in the role of parliament and the increasing power for common law courts, the monarch’s power had been very effectively curtailed and was widely viewed as such through credible commitment. High-handed breaking of contracts and seizure of property came to an end.⁹

⁹These 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 gold of goldsmiths was seized).
North and Weingast argue that, once property rights had been firmly established, risk premia fell. Capital accumulation took off, and investing in new ideas became much more profitable. Eventually, Britain’s growth rate took off. Many scholars have taken issue with this interpretation. Clark (1988) showed that interest rates on private instruments such as rent charges did not fall after the Glorious Revolution. Stasavage (2002) looks at public interest rates more closely and argues that the new settlement was not stable for a long time, and that interest rates were as much determined by partisan politics as they were by constitutional change. Sussman and Yafeh (2006) argue that wars and the threat of revolution mattered a great deal for British interest rates, and that the new Hanoverian regime was far from firmly established after 1688.

Following the work of North and Weingast, numerous scholars that have tried to use institutional analysis to explain the divergent growth records across early modern Europe. DeLong and Shleifer return to the argument in Montesquieu (1748) who famously argued that growth was likely to be more vigorous in Republican states which did not suffer arbitrary interventions by the authorities. They argue that absolutist rule was harmful because of three reasons – centralized powers run by ambitious, powerful princes fought more wars, taxed more comprehensively, and respected property rights less. They were also further away from the new trade routes to the Americas and Asia. 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. Against their view is the arguments put forth by another institutional approach is taken by Acemoglu, Johnson and Robinson (2005), who argue 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, their measure of “constraints on the executive” in Britain and the Dutch Republic grew, making it easier for these countries to overtake other European powers during the 17th and 18th century. They also demonstrate that this improvement in the quality of institutions mattered for growth – urbanization rates surged wherever geographically-determined “exposure” to Atlantic trade was high.

---

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…. 

[10]
Yet interpretations of Europe’s growth record that rely on institutions face as many difficulties as does the Malthusian Model. To start with, the exact definition of institutions remains a matter of some dispute. 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) wants to include 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 society will set up the rules that make exchange and investment possible. Yet there are few good theories that explain 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 theoretical model that sees better institutions work wonders through capital accumulation, 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 dictator’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 is needed 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 early modern Europe, the “constraints on the executive” variable compiled by Acemoglu et al. successfully predicts urbanization rates. By itself, this variable does not explain whether institutions will be successful. States with extensive checks-and-balances, such as Venice, the Holy Roman Empire and Poland, indeed placed very considerable constraints on a monarch’s freedom of action, and continued to do as long as they continued to exist. Yet they did not become hothouses of economic dynamism. Other states in which the absolutist agenda was successfully carried out, saw a significant reduction in the number of hurdles placed in the path of a prince’s wishes. Even for that epitome of absolutist rule, the Sun King Louis XIV, historians have largely rejected the idea that his rule can meaningfully be described as an implementing a successful, far-reaching absolutist agenda. For a generation, a new consensus inspired by the works of, inter alia, Georges Pagès and 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. Even if revisionism along these lines has gone too far, as some have argued – it seems doubtful that the currently available classification schemes capture enough of what is directly relevant to the argument that institutions and restrictions on executive caused economic growth before 1800.

The question is one of identification: one relationship between power and economic development specifies that the government should commit to respect property and contracts, and that this commitment should be made credible by placing adequate constraints on the government’s power. Centralizing power by absolutist monarchs is what built the (militarily and politically) successful nation states of Europe between 1500 and 1800. War was often an important catalyst of these changes – in Tilly’s memorable phrase, “states made war, and war made states.” Where constraints on the executive were far-reaching, such as in the Italian Republics, military successes could be few and far between and foreign invaders were able to destroy much of the wealth, in the classic pattern of Olson’s “roving bandits.” 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).

For the early modern period, looking just at constraints on the executive may be a misspecified model. As part of their struggle to field ever larger military forces, the modernizing states of Europe had to overcome the interests of powerful groups within society, from local lords to the church and the guilds. The constraints on many monarchs in 1500 were real enough. They probably, on average, did not increase in Europe until 1800, and probably declined significantly in some states. Yet most of the constraints of the executive at that point took the form of rent-seeking groups ensuring that their share of the pie remained constant. None of the groups that offered

---

11 From the sixteenth century onwards, the Polish Sejm had extensive powers over taxes, legislation, foreign policy, and the budget. Since 1505, the king could not pass legislation without support by the Sejm. From the end of the sixteenth century, majority voting declined, and the unanimity principle began to be followed. German Emperors had little direct power outside his own territory. Implementing any kind of new policy required the co-operation of the princes that ran most of the Empire. Getting legislation through the Reichstag, the imperial assembly (consisting of the prince-electors, of dukes and other princes, as well as the imperial cities), was very difficult. In 1648, the Emperor became formally bound to all decisions of the Reichstag. It is hard to think of stricter rules to bind a head of state. Yet German urbanization and income growth rates largely stagnated, at least after 1600.

resistance to the absolutist agendas of rulers in France, Spain, Russia, Sweden, and elsewhere were interested in creating growth – nor did they ensure that sensible, long-term policies were enacted. Put another way, if the Pagès and Mousnier revisionism is right then French absolutism failed in economic terms largely because it did not succeed in implementing much of the absolutist agenda because it was unable to control special interests. Large parts of Europe’s early modern history read like one long tale of gridlock at the hands of interest groups. Constraints on the executive, carried to the Polish extreme, were not conducive to economic development – not least because they could contributed to the disappearance of the state itself at the end of a sequence of grueling military defeats.

The example of the Venetian Republic is also instructive. In terms of its institutional setup, it hard to think of a political entity that would more closely approximate the modern ideal. Property rights were well-protected. Doges were elected officials, theoretically for life; 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 as a military and as an economic power. Few doubt that the events following the League of Cambrai (1516) were directly and indirectly responsible for the demise of Venetian power and the eventual decline of its 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 correlations probably had the opposite sign. The political entities with highly effective constraints on the executive quickly became victims of outside powers whose rulers operated without being hamstrung by domestic opposition. The Thirty Years War, made more brutal and protracted by outside intervention, ravaged the Holy Roman Empire of the German Nation; Poland disappeared at the hands of Austria, Russia and Prussia; and the Northern Italian Republics declined as outside powers – notably France, Spain, and Austria – increasingly intervened. For early modern states, political stability and a chance to escape being a victim of outside aggression may well have depended on pursuing an absolutist strategy. The effect of “constraints on the executive” on growth will hence be a composite of the (positive) effect for property rights, as well as the negative effects through the continued influence of rent-seeking groups, and the (similarly negative) effects of political instability and military defeat.

Constraints on the executive is an appealing concept for us today because it appears to be the inverse of arbitrary confiscation of property, of the tearing up of valid contracts, etc. In the early modern period, it often went hand-in-hand with the preservation of precisely the wrong incentives for growth. 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 more closely capture this dimension. 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. Instead of relying on the coding by a group of economists whose immediate aim is to assess the importance of institutions, they should be compiled collaboratively, by historians who specialize in the legal history of individual European countries. Ideally, they should provide a differentiated measure of how secure property rights were, how difficult it was to not only have access to legal recourse, but to actually win when in the right in a law suit against individuals or groups in power, how easily constitutions could be changed by brute force, etc. This could be complemented by a similar, collaborative data compilation effort that
produces an index of (legally sanctioned, as well as illegal) rent-seeking in Europe’s early modern states.

c. **Patience capital, natural selection, and other “cultural factors”**

Since Max Weber’s work on the spirit of capitalism, culture is one of the “usual suspects” that may determine wealth and productivity. Modern scholars (e.g. E.L. Jones, 2006, pp. 126-132; Temin 1997) have concurred. The problem, of course, that culture means different things to different scholars. Culture may be subject to evolutionary forces (e.g., Boyd and Richerson, 1985). Galor and Moav (2004) offer a model in which the crucial state variable that changes 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. But disentangling “inherent quality“ from changes resulting from responses to changing incentives seems a formidable challenge. Also, natural selection normally need not increase quality at all; it simply is adaptive to existing circumstances but utterly myopic, so that it is easy to see why it may not result in any improvements. 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. Surprisingly, the Galor-Moav approach has recently received some qualified empirical support Clark and Hamilton (2003) found that the rich and literate in early modern England fathered more surviving children. Whether 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 1700's is still far from an established fact.

One other change in culture that may have had an impact on economic growth as well as on the income distribution is the rate of time preference. In an innovative paper, Doepke and Zilibotti (2006) argue that the rise of a bourgeois elite in industrializing Britain may be regarded 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 of society – the middle classes – had accumulated a larger stock of “patience capital”, that is, a host of cultural practices and norms that make the delay of immediate gratification accepted and expected. Through centuries of careful saving and investing, these groups had 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. In their story, the absence of functioning credit markets is a key element in the story – only when financial markets are segmented do returns to patience (adjusted for risk) differ across groups. They apply their model to the decline of the aristocracy in Britain, generating an impressive fit overall between their predictions and historical fact.

---

13 Given that the earliest data are from the 16th century, there were only approximately 5-6 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 March 12, 2006).
Yet the concept of patience capital arguably holds even greater promise. It may be no accident that the “nation of shopkeepers”, as Adam Smith called it, became the first to industrialize. It offered an environment in which bourgeois values and practices flourished and gained in relative importance. Clark’s and Hamilton’s result that wealthier Englishmen had more surviving children could suggest that, instead of leading to an upward drift in some unmeasured, unnamed indicator of human quality, it simply enlarged the share of those who had learned to save (and invest). Such a change in population composition would also have contributed to the decline in English interest rates since the Middle Ages (Clark 1988), from 10-11 percent in the 13th century to 4 percent by the 18th. 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 the theories that attribute the rise in savings to the “Calvinist Ethic.” The tendency of interest rates to decline in stable, prosperous countries, noted by Adam Smith in the case of Holland, may not necessarily indicate that technology was stagnant and returns to capital diminishing rapidly. It may simply be that patience (and financial intermediation) was growing faster than savings could be usefully recycled into investments.

Compositional change can also help us understand evolving demographic behavior. Behavior often differs across subgroups, as both historians and economists have found.14 Given that many more children could have been fed, and that the constraints on fertility behavior were mostly social and cultural (working through nuptiality rates), it is easy to see how evolving norms could have changed population growth rates. Differential fertility behavior and evolutionary mechanisms might thus explain how and why the Malthusian regime came to an end. In the Galor and Moav (2002) model, for example, this is depicted as the gradual increase of the number of people with a strong preference for “high-quality offspring.” As yet, we know far too little about the relative differences in reproductive behavior and success in early modern Europe. Compositional change may have played a large role, but at the current stage, it is hard to tell. What is needed is much more evidence along the lines of the material gathered by Clark and Hamilton documenting differential fertility over the long run.

If Europe saw a rise of bourgeois values prior to the Industrial Revolution in terms of savings behavior, it was complemented by a rise in work intensity and the length of the working day. DeVries (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. Clark (1987) found evidence that work intensity in the most economically advanced parts of Europe was much higher than elsewhere. Voth (1998, 2001) argues that the workyear in Britain was already a long one by 1750, and that it increased yet further as a result of a decline in festivals, holy days, and the practice of taking Mondays off (“St. Monday”).

d. Long-run changes in the European economy before the Industrial Revolution: underperformance or undermeasurement?

Before the eighteenth century, there are few output measures that could help us reconstruct production per head. What we have instead are wage series as well as prices, plus some scattered infor-

---

14In addition to Clark and Hamilton (2006) cited above, we may mention the work by Herlihy (1997, pp. 56-57) and Galor and Moav (2002).
mation on rents on land and the like. The GDP figures compiled by Maddison and others are essentially based on these wage series. Even the better ones rest on distinctly shaky foundations. Allen (2003) is one of the more recent attempts to offer new and better series. Yet wage information is still overwhelmingly skewed towards urban workers – a small part of the population at best. Price indices are dominated by food items, and grain in particular. We have very little information on the price or quality of clothing, of rents, and other essential items. As a result, some new revisions imply changes in living standards over a century that are only half as big as earlier series suggested (Clark 2005).

Yet even if we had perfectly measured wages and prices for consumption bundles of early modern populations, we may be underestimating changes in living standards by an order of magnitude. The main reasons are unmeasured quality improvements and the value of variety. As Nordhaus (1997) has demonstrated in the case of lighting, big efficiency gains as a result of new technology are often missed by traditional ways of measuring price changes. Increasing variety may have brought benefits every bit as big. Trade, both long-distance and local as a result of higher population densities and better communications, broadened the range of goods available to the vast majority of Europeans. Between 1500 and 1750, Europeans of all social classes began to consume large amounts of sugar, tea, cod, tobacco, cocoa, porcelain, and cotton goods. Potatoes and maize made important changes in the staple diets. The effects of the range and variety of goods available in the present-day US has been estimated to have boosted living standards by the equivalent of a fall in prices by approximately 30% over the last 30 years (Broda and Weinstein 2004). Even the comparatively modest trade before 1800 expanded the range of available goods greatly, boosting living standards far more than GDP indicators show.

3. The first transition: from Malthus to a new economy.

Whether Europe in 1700 was still in the grips of a Malthusian regime, properly defined, is open to question. In many regions, and over longer periods, population and living standards had started to move up in parallel. To be sure, the changes were slow, sometimes temporary, and subject to reversals and negative shocks. Nonetheless, it is questionable if all of European growth after 1500 could be regarded as yet another passing “efflorescence”. The timing and location of the first transition, from a “traditional, slow-growth” society to a more dynamic society is hard to understand based on modern unified growth models alone. Exogenous growth models treat the transition as basically pre-determined by technology. In Galor-Moav, Jones and others that use endogenous growth based on population size, technological change boosts population size, which then produces ideas with greater frequency.

The rapid rise in population after 1750 in many European countries was to mark a far more dramatic departure from earlier patterns. At the same time, output started to grow – living standards remained constant or even increased slightly, where earlier patterns would have suggested a collapse. This is the phase that Galor has termed the “post-Malthusian” phase. We implicitly follow his approach of dividing the Industrial Revolution as conventionally defined into two phases – an early phase when population size no longer determines living standards, and a later, “Solow”-phase when
technological change becomes rapid and is largely translated into higher living standards.

a. The meaning of the Industrial Revolution

Leaving long-term economic stagnation behind can be modelled either as a sudden discontinuity with the past, or as a gradual acceleration over the very long run. In part, the difference is one of framing the time span appropriately – compared to what happened in the millions of years since homo sapiens left the African plain, developments after 1750 were very rapid indeed. If we focus on changes over decades, the pace no longer seems very impressive – at least compared to, say, growth rates after 1945. The discontinuity par excellence is the Industrial Revolution, but how should modern growth economists interpret it?

The Industrial Revolution remains a pivotal event in the historiography of the beginning of economic growth but its interpretation has changed considerably since the term was first coined two centuries ago. IT was not the beginning of economic growth as such: It is now quite clearly accepted that economic growth did not start during the Industrial Revolution and that its macroeconomic effects did not become discernable until the second third of the nineteenth century. In part this is to be explained by the fact that at first the Industrial Revolution affected only a small sector of the British economy and that most economic activity still took place in slowly-changing “traditional sector” even if the weight of the rapidly-growing modern sector was increasingly rapidly and the exact boundaries of these two sectors remain in dispute. To this arithmetical truism we may add the fact that the first Industrial Revolution was accompanied by three unrelated phenomena that depressed economic growth, namely almost continuous war from 1776 to 1815 (with a brief respite in the decade after 1783), the rapid growth of population that, in addition to increasing overall population size, also increased the dependency ratio, and finally the succession of poor harvests and high grain prices that struck the British economy between 1760 and 1816.

Despite the absence of growth itself, the Industrial Revolution remains the transition from the slow-growing economy of the early modern period to the faster growth of the post 1830 period. In one classic formulation, it was neither the age of steam, nor of cotton, nor of iron, it was the age of progress. Technological progress took place over too many fronts to be simply waved off as a local fluke in the cotton industry, as some historians (Clark) have done. What mattered above all was not so much the famous “wave of gadgets” of the 1760s and 1770s, but the fact that the process did not fizzle out once it had built a number of successful designs. It is perfectly possible to envisage a world of mules and stationary low-pressure steam engines crystallizing and settling down at a new, somewhat higher, technological equilibrium. That this did not happen is the key to the emergence of modern growth, even if that growth came properly speaking after the Industrial Revolution was complete.

A full explanation of this phenomenon cannot be attempted here (but see Mokyr 2006, 2007). The summary is something as follows. The Industrial Revolution, though it took place first in Britain, was really a Western European “joint project” in which British inventors collaborated with inventors and scientists from the European Continent and North America. In doing so, the Western World depended on changes not only in economic circumstances such as a higher living standard and a more commercially integrated environment in 1750 than in 1500, but also on changes in the mind-set of its elites. Among those changes in mind-set, the Enlightenment of the eighteenth century is central. The Enlightenment advocated explicitly the changing the agenda of scientific research to
suit the needs of the economy. In this endeavor, the inspiration came from the Baconian program, implemented first by the Royal Society and subsequently by scores of organizations, academies, and informal groups. Moreover, the Enlightenment movement explicitly advocated the dissemination of knowledge to those who might be able to use it best, through books as well as through informal contacts in which tacit knowledge could be transferred. As a result, the connections between people of science (“natural philosophers”) on the one hand and inventors, mechanics, and engineers on the other was much closer in 1820 than it had been in 1680.

Moreover, the Enlightenment brought about a slow but indisputable fading of the rent-seeking institutions of the ancien régime. By 1820 rent-seeking had been sufficiently weakened that the institutional feedback discussed above was made inoperative. European nations that were getting wealthy were no longer in fear of being invaded by a predatory power, and within most western countries, governments were constrained in the amount of resources they could expropriate from their citizens. Moreover, exclusionary organizations such as guilds, or rent-creating obstacles such as internal (and increasingly external) tariffs were weakened during the first half of the nineteenth century. This triumph of liberal politics inspired by classical political economy — the Enlightenment’s proudest offspring — no matter how temporary, was necessary to lift the West over “the hump”.

It is a somewhat separate literature why this process started in and was most successful in Britain. Britain was peaceful (at least in the sense that no fighting took place on its soil) and that it had a government that had discovered a way of changing institutions peacefully through the emergence of meta-institutions that could change the rules of the economic game legitimately (in the sense of them being accepted even by those at the losing end of the stick). To that we can add the fortunate resource situation in Britain such as the availability of high-quality coal, and the presence of a substantial cadre of highly skilled craftsmen on whose technical dexterity the implementation of new techniques depended. Its more conducive political climate and the existence of complementary resources ensured that Britain could take advantage of its own inventions as well as those of others.  

Yet this advantage was temporary. By 1830, when this process had run its course, the gap between Britain and the Continent began to show signs of narrowing, yet instead of Britain being pulled down to the slower rate of development of the Continent, the growth of income began to spread to the rest of Western Europe. By that time modern economic growth, rapid and sustainable, could truly commence.

b. Endogenous growth: did size matter?

Early models such Kremer’s (1993) paper modeled the transition from “Malthus to Solow” as one long, gradual acceleration of growth rates. Kremer’s model assumes that more people spell

---

15 The Swiss visitor de Saussure had noticed already in the 1720s that “English workmen are everywhere renowned, and justly. They work to perfection, and though not inventive, are capable of improving and of finishing most admirably what the French and Germans have invented” (de Saussure, 1902, p. 218, letter dated May 29, 1727). David Hume [1758, (1985)], p. 328 pointed out (with some exaggeration) that “every improvement which we have made [in the past two centuries] has arisen from our imitation of foreigners... Notwithstanding the advanced state of our manufacturers, we daily adopt, in every art, the inventions and improvements of our neighbours.”
faster technological change since the probability to have a bright idea is more or less constant. Kremer showed that some of the basic predictions derived from such an endogenous growth model driven by population size hold both over time and in cross-sections – since one million BC, 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. Demographic transitions, with fertility responding negatively to higher incomes above some threshold level, avoid the model from exploding. This assumption is also critical in the dynamics of Galor and Weil (2000) where any technological change in the Malthusian world leads just to population growth, but once population exceeds a critical level, it begins to change the model’s equilibrium. Yet a direct nexus between population size and technological progress is historically problematic.

As Crafts (1995) has pointed out, the implications for the cross-section of growth in Europe are simply not borne out by the facts – bigger countries did not grow faster. Even if we substitute population with factors like market size, which might have influenced the demand for innovation, the contrasting growth records of Britain and France are hard to square with endogenous growth models. It is, of course, quite disturbing that in 1750, on the eve of the Industrial Revolution, Britain had just experienced half a century of virtual demographic stagnation. One might also point out that if population size is critical and that it is population growth that “liberates” Malthusian economies from their stagnation, China, where population grew from 130 million in 1650 to 420 million in 1850 yet where no Industrial Revolution could be discerned, should be of some concern. Also, while the Kremer model can account for global growth rates prior to 1700, it strongly underpredicts growth thereafter. Perhaps more seriously, there is a problem of identification here: growth may have accelerated in some parts and period as a result of factors unrelated to the hypothesized population-size-innovation mechanism. As more resources became available, demographic growth accelerated. The simple correlation of population size and growth rates is not a proof of the underlying endogenous growth model; indeed it could just as well be taken to be a confirmation of the Malthusian model. Modeling alternatives are needed if we want to account for the speed of growth after 1700, as well as for the variation between countries.

A second class of models in which size matters takes technological change to be exogenous, and models a set of conditions under which new techniques will be adopted. Early models in the tradition of Murphy, Shleifer and Vishny (1989) also relied on demand effects, and hence the size of economies, to explain when a “big push” might occur. In order to pay the fixed cost necessary for adopting modern production, demand needs to be sufficiently high. In typical “big push” models, this is only the case if a whole range of industries industrializes. The chances of this occurring

---

16 It is indeed striking that prior to the rise of the British economy to the fore, Europe’s most successful economies tended to be city states (Hicks, 1969, p. 42). City states, with high density but relatively overall 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 these institutions up 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.

17 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.
increase with total output. One implication of these models is that industrialization may have been feasible long before it got underway – if only everybody had decided to invest earlier in fixed-cost technology, profits would have been high enough to justify the expense. Simple co-ordination failure can thus undermine the transition to modern technology. Possible modifications and extensions of this approach also assign a role to the income distribution and the structure of demand.

Models in the “big push” tradition run into similar problems as population-based endogenous growth; the European experience after 1700 does not suggest that absolute size of economies is a good predictor of the timing of industrialization. Moreover, most of the technologies adopted during the Industrial Revolution required only limited up-front investment, and were often financed by retained earnings or informal credit networks. Before the late nineteenth century, fixed costs were typically small in manufacturing. When it comes to production technology with high fixed costs, adoption decisions after 1870 could possibly be explained by the 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. Yet these infrastructural investments—canals, turnpikes, harbors—do not appear to have suffered a great deal from capital scarcity, the Bubble Act notwithstanding (Michie, 2001).

Indivisibilities also play a crucial role for models that put risk diversification at the heart of adoption decisions. Acemoglu-Zilibotti (1997) argue that at low levels of development, the volatility of growth rates is high. Households need to diversify their investments. Productive projects require substantial setup costs. In order to invest in them, households need to rich enough—otherwise, they would end up putting “all of their eggs in one basket”. Here, industrialization effectively depends on a number of lucky draws. It also has the feature that, since households do not take into account the effect of their investment decisions on aggregate productivity, industrialization may be delayed because of a co-ordination failure. The model is attractive in that it incorporates a stochastic component—industrialization may partly be the result of chance. Not every aspect of actual industrial transformations is fraught with meaning – and the country that actually went first could simply have been lucky. Yet the size of most industrialization projects was relatively small – even the largest textile mills, had they been financed by a single person, hardly represented a large risk concentration for modestly wealthy individuals.

In exogenous growth models technology just happens and adoption decisions no longer are modelled explicitly and size itself no longer affects technology or productivity change. In one example, Hansen and Prescott (2002) argue that technological change in both the land-using (diminishing returns) and the non-land-using mode of production was exogenously given and constant. They assume that over the course of each generation of 35 years duration, productivity increased 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). Initially, only the Malthus technology is used. 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 growth rates of productivity chosen for the calibration. It is also difficult to square with European economic history as a whole, as well as with the differences between countries. 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
The only obvious alternative is to posit differential rates of productivity increase in the Solow sector, which would rather be a way of assuming the result.\(^{18}\)

c. Changes in demographic dynamics before 1830

One of the biggest challenges in interpreting the history of growth in Europe before 1850 comes is the rather sudden increase in population growth in Europe in the second half of the eighteenth century, after a period of stagnation in the first half. The latest revisions of the Wrigley-Schofield (1997) English population estimates reinforce the impression that fertility increases dominated as a cause of more rapid growth; mortality played a role, but it was responsible for only about one third of the acceleration.\(^{19}\) Regardless of whether one accepts Wrigley and Schofield’s interpretation of the Malthusian model, 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 very responsive to changes in prices. This suggests that in contrast with the arguments of growth theorists, the Malthusian regime was falling apart before the Industrial Revolution and not as a response to it.\(^{20}\) The unsolved question to date is why.

The rise of population that took place without a collapse in per capita incomes may have generated positive externalities of another kind. Regardless if size mattered to the generation or adoption of new technology, as the endogenous growth models suggest, the oldest and in some ways most elementary of mechanisms – a simple increase in the division of labor as a result of greater population size and density – could also have contributed to an acceleration in 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 Sung dynasty China. Similarly, in Europe, higher population densities and greater economies generated the scope for positive externalities, partly through improvements in turnpikes and canals (Bogart 2005a, 2005b).

Regardless of what the sudden population spurt did, it leaves as yet unexplained how and why it happened. Models that link population dynamics to technological progress itself, such as Galor and Weil (2000), run into timing problems, namely that population growth started in the mid-eighteenth century, before any serious impact of technological change on output per capita can be discerned. It remains therefore unexplained itself. If there is any challenge left to economic theory in explaining big events in history, this development is a prime candidate. The challenge is compounded by three complicating factors. One is that even in Britain, the demographic transition

---

\(^{18}\) The only obvious alternative is to posit differential rates of productivity increase in the Solow sector, which would rather be a way of assuming the result.

\(^{19}\) Wrigley (1983) showed that without mortality decline, 18C 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 and Schofield (1997) qualify these conclusions to some extent, finding a faster decline in mortality, but the relative rankings are unlikely to change significantly.

\(^{20}\) It also suggests that separating institutional factors from the Malthusian model is a serious simplification, because the Poor Law played an important role in cushioning the impact of high food prices on demographic behavior (Post, 1990).
In Cervellati and Sunde (2005) the relationship between mortality and human capital investment is explicitly modelled. This is a little explored aspect of modernization, but one that was historically of some importance. All other things equal, longer life expectancy would encourage investment in human capital, although it is important to emphasize that a reduction in infant mortality would not directly bring this about, because decisions about human capital are made later in life. Increases in life expectancy at age 10 or so are more relevant here. A second relationship in this model is that life expectancy itself depends on the level of education of the previous generation: better educated parents will be better situated to help their children survive.

started before the Industrial Revolution could have possibly have affected living standards. Second, the demographic revolution seems to have spread through many parts of Europe, regardless of their level of industrialization and that the speed and nature of the industrialization process seem to have been unaffected by the decline in mortality. Third, the mechanism through which population increased seems to have differed from country to country, in some such as Britain increases in fertility and mortality accounting for about half the growth each, whereas elsewhere birth rates stayed constant or even declined from an early date, as they did in France. Models that link birth and death rates in the economy are hard to construct, in part because they not only respond quite differently to economic and environment shocks, but also because they affect one another in complex manners. Significant progress in this area has been made in two path-breaking papers by Sunde and Cervellati (2005, 2006). 21 It seems plausible to link declines in mortality to human capacity to fight disease through more educated people who led more hygienic lives, breastfed their children, and had access to medical care. Yet the historians will point out that mortality in this period was probably dominated by a single foruitous event, the discovery of smallpox vaccination (e.g., Mercer, 1990), and most of the impact of education on mortality was offset by the urbanization than accompanied industrialization, leading in fact to heightened mortality rates (Huck, 1995).

d. Institutions and the Industrial Revolution:

Much of the modern debate about growth centers on the relative importance of institutions versus human capital (Glaeser et al. 2004, Acemoglu and Johnson 2004). In cross-sections of countries from the late 20th century, constraints on the executive tend to be positively correlated with higher per capita output. Because of the potential for reverse causation – with higher output improving institutional quality -- work on modern data has principally focused on finding an exogenous factor that affect institutions, but not economic outcomes. One such factor that has been used with great success is historical settler mortality. In a series of papers, Acemoglu, Johnson and Robinson (2001) show that countries in which white settlers had a high chance of survival ended up with more desirable institutional arrangements. They are also markedly richer, making it much more likely that the link between institutions and efficiency is causal. Despite some challenges to the quality of the data, the basic relationship appears to be robust. What is subject to vigorous debate is its meaning. By far the best research on the importance of institutions in European economic development has been carried out in the context of medieval and early modern Europe (Greif, 2005), and surprisingly no detailed work to date has been carried out to explain the British Industrial Revolution.

Three 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 are associated with the mercantilist ancien régime (Mokyr,

21 In Cervellati and Sunde (2005) the relationship between mortality and human capital investment is explicitly modelled. This is a little explored aspect of modernization, but one that was historically of some importance. All other things equal, longer life expectancy would encourage investment in human capital, although it is important to emphasize that a reduction in infant mortality would not directly bring this about, because decisions about human capital are made later in life. Increases in life expectancy at age 10 or so are more relevant here. A second relationship in this model is that life expectancy itself depends on the level of education of the previous generation: better educated parents will be better situated to help their children survive.
2006). The roots of this reaction involve some combination of the changing political influence of economic elites and the influence of a more liberal ideology. Second, these changes in most cases had to be imposed through political violence (in the United States and France). The only apparent exception is Britain, where the meta-institutions such as Parliament could adapt to changing circumstances and beliefs and reform the system peacefully and without major upheavals. 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. Third, as we have seen earlier, the precise mechanisms through which “better” institutions affect economic growth are yet to be specified. The leading candidates are “law and order” and the efficient protection of property rights, the provision of public goods, acting as a coordinating and standard-setting agent, and the direct encouragement of the knowledge-producing sector.

One of the more interesting institutional aspects of the Industrial Revolution is the importance of IPR’s in the early stages of technological progress. Jones (2001) is the only growth paper to date that models this important institutional parameter directly, and it turns out to play a pivotal role in his model in whether the Industrial Revolution was “inevitable.” Modeling the production of “new ideas” is of course one of the hardest things to do in growth models, and macroeconomic models have had to simplify away much of this historical richness. Thus the literature cannot deal 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. Perhaps more interesting is the failure of these model to 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 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 David and Dasgupta, 1994). Finally, it would be a mistake to identify IPR’s during the Industrial Revolution with the patent system as such. 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 prizes were as important as the rents earned by inventors. As Khan and Sokoloff (1998) have shown, the British patent system was far less accessible and user-friendly than the corresponding U.S. system, and yet this does not seem to have affected British technological leadership before 1850.

d. Human and physical capital in the Industrial Revolution: theory and evidence

The most influential class of formal models created in the past decade sees the Industrial Revolution as largely synonymous with the demographic transition and an increasing role for human capital. Becker, Murphy and Tamura model an economy without a fixed factor of production.

---

22This parameter \( \pi \) 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.

23The literature on the operation of the patent system in Britain is quite large, for an introduction see Dutton (1984) and MacLeod (1988).
Improvements in human capital in turn 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 basically identical. Lucas (2002) extends the Becker et al. approach by adding a land-using sector with diminishing returns, and a modern sector where human capital enters linearly. In the Malthusian trap, there is no investment in human capital. One of the main strengths of the “unified theory” of Galor and Weil is that their model actually predicts that in the first stages of the Industrial Revolution the role of human capital is modest, because the rate of technological progress initially is driven by size, not by quality (Galor, 2005). Yet, while it is consistent with the facts, it is not the only model and tale that can be told that predicts this.

Once the Industrial Revolution has taken place, there is ample investment; growth surges, and fertility rates decline. Both types of model offer no good explanation for how the switch from one regime to the other might happen. What they share is the assumption that investment in human capital through the rearing of “higher-quality” (i.e., better-educated) children is an input into the process of technological change. This, in and of itself, is not an obvious truism. The technology of the Industrial Revolution was surely not the single creation of a few mechanical giants in the hero-worshipping traditions of Victorian writers such as Samuel Smiles, but neither was it a popular mass-movement. It was brought about by a technological elite of inventors, engineers, mechanics and skilled craftsmen, whose dexterity and ingenuity was critical. Not only new ideas, but also the ability to implement them, turn blueprints into functioning designs and maintain and operate them effectively were central. Yet none of those involved more than a small minority of the labor force. For the rest of the population, as we shall see below, the links between human capital, skills, and growth are thus far a matter more of convenient modelling and speculation than of historical fact. For the period before 1850, there is little evidence to support human-capital based approaches and an increase in the rate of return on human capital due to the acceleration of technological progress. Skill premia are flat or declining (Clark 2003). The models by Lucas and by Becker et al. seem geared towards the developments after 1850, when fertility began to decline in earnest in some European countries. They have little to say about developments before the middle of the eighteenth century, when a number of countries appear to have cast off Malthusian constraints.

At least some of the assumptions of the model emphasizing the quality-quantity trade-off and the rather abrupt transition to quality need to be reevaluated in view of the historical evidence. One doubt that arises is whether investment in quality before 1800, too, was intensive in parent-time. The main form that training took was apprenticeship, in which a contract between the trainee and the master involved an indenture, a commitment by the trainee to work during his learning period, and at times cash payments (Humphries, 2003). Secondly, there is no doubt that some forms of human capital (such as literacy and numeracy) were on the rise long before the Industrial Revolution. In part this was due to the Reformation, in part due to slowly rising incomes (children’s quality was a normal good), and possibly to a rising demand for literacy in the service sector during an age in which commerce and finance were growing rapidly. To complicate matters, during the Industrial Revolution, literacy rates were probably largely stagnant; there is little evidence of an increase in the returns to education before 1850 (Schofield 1973, Clark 2003). Measuring literacy rates in a consistent and comparable fashion is no minor matter, especially with the kind of pre-1800 sources
available for this matter. Baten and van Zanden recently examined 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 tenfold and a hundredfold. The Netherlands and the UK are far ahead of other countries – the richest areas consumed the largest number of books. A recent literature survey, focusing on the ability to sign one’s name in 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). However, Britain was considerably richer than those countries, and if we allow for the fact that literacy was in part a desirable good that people consumed more of when they became richer, Britain’s lack of advantage in literacy is all the more striking (Mitch, 1992, 1999). Its ability or willingness to educate its young did not appreciably improve during the years of the Industrial Revolution. School enrollment rates did not increase much before the 1870s (Flora et al 1993).

The main conclusion has 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 the negative-feedback low growth regime that survives contact with the most basic facts in economic history. Exogenous growth models, where they emphasize institutional constraints, hold greater promise, even if the messy detail of historical reality is proving hard to press into the Procrustian structure of some modelling approaches.

f. **Factories and the re-organization of production [To be completed]**

[ What is not quite clear is how we should interpret the rise of the factory system in the context of changing relative costs for moving ideas, products, and people. The greater value and complexity of machinery, and a consequent need to supervise workers, were probably important factors in the rise of the factory system. Clark (1994) argued that the strict discipline of the factory offered a way to overcome workers’ own inertia – to pre-commitment to working hard and long when they would otherwise be tempted to take time off. Even if this factor explained the workers’ willingness to submit to factory discipline, it is not clear if it helps us to explain why the factory system arose when it did. Mokyr (2001) tells a story about the relative benefits of cottage-shop production, manufactories, and factories from a transaction cost perspective. He argues that the rise of factories came about as a result of more interdependent production modes, where piece rates could not work as well because the marginal product of workers was harder to assess, and because of the flow-character of production (with the pace being set centrally by the manager). This was the case in some industries; in others, notably textiles, factories rose along with piece rate payments, making it all the more puzzling why workers could not rent equipment from factory owners.]

g. **Was an Industrial Revolution Inevitable?**

Growth before 1850 was, more often than not, a fleeting phenomenon. There were many “false
starts” and stops - periods of rapid growth followed by stagnation and decline (Braudel 1973; Goldstone, 2001). It is at least doubtful that there was much about Britain in 1700 that guaranteed that the next episode of rapid growth would not end like earlier ones. Both economic historians and growth theorists are ambiguous about the inevitability of the Industrial Revolution and the role of chance. One group emphasizes the role of historical accident both in terms of timing and location. Crafts (1977) argued that accidental factors, and not systematic advantages, were crucial -- that France, for example, could have easily industrialized first had it not been for a number of random factors. Acemoglu and Zilibotti (1997) build a probabilistic model of the Industrial Revolution where technology adoption depends on the realization of shocks. Lucas (1998) adds a sudden, one-off increase in returns to human capital to his model to produce an Industrial Revolution, which could be interpreted as reflecting historical accident. Jones (2001) presents a unified growth model where change is driven by the interaction between ideas and population size. He argues that property rights improved in a major way in the 20th century, and adds this to his simulations. This has the result of accelerating the Industrial Revolution by centuries. While some kind of transition to self-sustaining growth is inevitable in his model, its timing is heavily influenced by exogenous variations in the institutions protecting property rights.

In contrast, both economic historians and growth theorists have argued that there is little role for historical accident. In models such as those by Galor-Weil (1998) and Hansen-Prescott (2002), given a certain set of starting values, an Industrial Revolution is inevitable. The dynamics governing growth from the earliest date ensure that there has to be an eventual acceleration. While many economic historians will be sceptical about this particular feature because of its post hoc ergo propter hoc qualities, there is a school of thought that has defended such an approach. David Landes, (1994), responding to Crafts’s argument that Britain’s role as the First Industrial Nation may have arisen by chance, has argued that big events need big causes. In his view, both the Industrial Revolution and Britain’s role in it were determined by that country’s starting conditions. Yet there is no contradiction between the argument that big events need big causes and that the Industrial Revolution was not pre-ordained from the start. Some accidents, such as the destruction of the Spanish Armada in 1588, may have been fairly “big.” At some critical junctures of history, a point of bifurcation may have occurred, and that accidental factors may have steered economies on one path or another; once on that path, their direction was clearly no longer accidental.

As noted earlier, of all the possible historical paths, we only observe one. It is easy enough to imagine that it could have looked different, but how different? What appears to underpin theorists and historians’ interest in the role of accident are the wide differences in the fate of individual countries, and the many promising starts that never led to a “take-off” into self-sustaining growth – more often than not, before the transition “from Malthus to Solow” occurred, there were many episodes of a long and stony path “from Malthus to Malthus.” The inability of many modern growth models to explain cross-sectional variations in economic performance has left some economic historians sceptical (Crafts 1995). If progress is to be made, theorists should aim to offer models that have something to say about the factors that influenced adoption probabilities for new technologies in the past, and they ought to assign an explicit role to chance; ideally, they should also offer quantitative answers to question how much the individual features of a country influenced industrialization probabilities.

4. Growth and the emergence of the “new economy”
Models in the Lucas, Becker, and Galor et al. tradition offer a unified and consistent account of the transition into self-sustaining, rapid growth of per capita incomes. They are ultimately driven by human capital accumulation. More rapid technological change (either exogenously given, or as a result of larger population size) increases the returns to human capital. Parents respond by producing higher-quality offspring. Additional human capital input in the economy boosts growth and accelerates technological change yet further. Because parents respond to the changes in the quantity-quality-tradeoff, fertility declines. Hence growth becomes intensive rather than extensive, boosting incomes and not population size. In this section, we examine each of these facets in turn – the importance of human capital for growth in the emergence of the 19th century’s “new economy”, and the importance of quantity-quality tradeoffs for fertility change. We then discuss what determined the accumulation of useful knowledge.

a. Education and the first “new economy”

One factor that is common to models in the Lucas tradition is that they predict an increase in the demand for human capital during the transition to self-sustaining growth – and that technological change should be heavily skill-biased. This is historically problematic. There seems to be little correlation between widespread literacy and schooling and the onset of technological progress. Mitch’s (1998) view has been that, if anything, nineteenth century Britain was overeducated. By this he means that the amount of human capital exceeded that which was needed by the demand for production. Reading and writing were desirable in their own right, that is, as consumption goods, and were not just parts of an investment process in which the rate of return on the margin would equal to interest rate. It is far from clear that the main developments in manufacturing during the Industrial Revolution, or even developments in its aftermath, depended heavily on an increase in human capital, as conventionally measured. Possibly, administrative tasks became more important, and 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 in the normal sense of using large quantities of literate workers with many years of formal schooling. Research on the output of textile workers in New England showed that only experience was a good predictor of output; the ability to read or write was neither useful in its own right, nor did it go hand-in-hand with other, unobserved characteristics that would have raised output (Leunig 2001).

Many contemporaries commented on the de-skilling that accompanied the Industrial Revolution. In the textile industry, the cotton mules, spinning jennies and Arkwright frames replaced skilled labor with a mixture of capital and unskilled labor. In some cases, innovation deliberately sought to replace the skilled “labor aristocracy” whose bargaining power the cotton masters resented (especially in the case of mule-spinners). High wages of skilled mule spinners outside Britain apparently acted as a strong incentive to adopt ring spinning, which could be

---

25Adam 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.”
performed by unskilled labor (Leunig 2003). In Britain itself, and despite some inherent technical limitations, highly skilled mule spinners were cheap enough to make the industry competitive all the way up to 1914. All this suggests that neither formal-education based indicators of skills nor the nature of technological change offer decisive support to the predictions of unified growth theory. However, the current state of knowledge on skill-bias of 19th century technological change is far from complete. Perhaps if one could construct input-output tables with human capital as a separate input, when linked with more data on the schooling and skills of the workforce, these could help to shed more light on many of the crucial issues. Such a project, conducted for a number of European countries and the US in the 19th century, could help scholars construct more detailed estimates of the embodied skill content of production and the extent to which this changed over time.

One indicator of the changing role of skills is the skill premium. Unfortunately, our knowledge of the behavior of the skill premium over time is very 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. Williamson (1985) claimed to show that the skill premium surged until 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. This is problematic for authors such as Doepke and Zilibotti (2005), who argue that child labor laws were introduced in England after wage inequality surged. In their model, the political equilibrium that sustains restrictions on child labor require a substantial premium for well-educated members of the workforce. In this way, there is an institutional response to the voracious demand for human capital. Unfortunately, Doepke and Zilibotti reliance on Williamson’s flawed data undermines the credibility of their results. Galor offers an explanation why skill premia failed to rise in 19th century Britain. He argues that, after the introduction of compulsory schooling, supply was so ample that premia remained flat. 26 This would offer a plausible interpretation if skill premia had increased prior to the introduction of the Factory Acts. Yet this is exactly the key piece of the puzzle that is missing.

Most of the skills that this elite of workers 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 group. 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 to the rest, and on the supply response that these differences in pay engendered. The failure of traditionally measured skill premia to show a rise may well mask an increasing polarization within the workforce, with industrialization raising the returns to supervisory and advanced mechanical skills, and reducing those for standard ones (such as blacksmithing and weaving). If there is one area that shows promise for future work, it is the acquisition of factory- and task-specific skills. Steep experience-based earnings profiles in the textile industry apparently made acquiring skills attractive. During their early years, when unskilled workers such as brickmakers were better paid, workers were effectively investing their own human capital of a highly specific kind; by age 35, they could look forward to earning 2.3 times the wages

---

26 Galor 2006, private communication.
of a brickmaker, and still more than a coal miner (Boot 1995). However, the total number of workers investing in their own skills was sufficient to keep overall premia for, say, skilled mule spinners, relatively low in England at the end of the 19th century (Leunig 2003). One valid test of the human-capital approach would focus on highly skilled workers such as the textile operatives examined by Boot and Leunig, and to ask – did they receive greater rewards for investing in their skills (by accepting years of poorly paid on-the-job training) than, say, apprentices in traditional sectors?

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 reflects “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 increase social control over the lower classes. Since compulsory schooling played a crucial role in raising human capital, examining the history of their introduction becomes crucial. Galor and Moav offer a first, innovative step in this direction. They analyse voting records in the House of Commons, and argue that the educational reforms that came into force as a result of the Balfour Act of 1902 largely reflected the interests of capitalists in improving workers’ education.

b. The human capital that mattered: Factory discipline, supervision, and the skills for microinventions

Standard growth regressions for the 20th century measure human capital as some (earnings-weighted) average of years of schooling. Like all measurements, this is nothing but a convenient shortcut. It is entirely possible that it works better for the present than the past, and that the skills that mattered in the 19th century are poorly captured by calculating school attendance rates. The rise of the factory system required plenty of highly specialized skills that are not necessarily transmitted through formal schooling – discipline, punctuality, and technical prowess.

The equipment and materials used by workers belonged to the capitalist and were costly. Factory owners needed to install into workers a culture of loyalty, punctuality, and sobriety. They wanted to foster a culture of loyalty and respect, a willingness to take instructions from and cooperate with other workers. This is a direct result of the expensive equipment in factories. Wage premia for disciplined work in the factories were high vis-a-vis other, more self-determined forms of

27 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.

28 The latter effect would be in the spirit of Acemoglu and Robinson’s (2001) 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.

29 This insight is hardly indebted to modern theory: Karl 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, Vol. I, pp. 405–06).
employment, and the factory system’s profitability relied crucially on work intensity (Pollard, 1965; Clark 1994). For the same reasons, monitoring workers was a highly important task. If "discipline capital" mattered more for the first Industrial Revolution than education, conventionally measured, economic historians should compile more comprehensive wage measures that capture the rewards for workers who successfully 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 19th century factories. The most obvious testable implication of this idea that early factory owners should have a preference for the employment of comparatively more pliable workers, even if they were of low skill — i.e., women and children. This was very much the case in the early stages of the textile mills.

In an important sense, the Industrial Revolution was carried not by the skills of the average or modal worker, but by the ingenuity and technical ability of a relative few. This is not a return to the "heroic inventors" of the Victorian era. We should recognize, however, that the new technology was generated by a small army of highly skilled men. A few were great engineers, but these could hardly have succeeded without a much larger contingent of skilled artisans and mechanics, upon whose dexterity and adroitness the top inventors and thus Britain's technological success relied. These were the craftsmen, highly skilled 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 the essential complement to the inventors, since they were the ones that turned models and designs into working machinery, maintained and operated it, and produced a cumulative torrent of small, incremental, but cumulatively indispensable microinventions, without which Britain would not have become the "workshop of the world." They comprised perhaps five percent of the labor force. This would put their counts in the tens of thousands but not in the millions.

c. The quality-quantity tradeoff and the demographic transition

Human-capital based models of long-run growth assign a crucial role to the fertility transition. It is normally modelled as a response to changing economic incentives. The predictions are that (i) skill premia surged, (ii) parents responded to this change in the trade-off between child quantity and quality by limiting fertility. As we saw, the first part of this argument is problematic – returns to human capital, conventionally measured, probably did not increase significantly before 1870. The second one will probably also have to be modified. Since the economic benefits of formal education were probably minor for working class employment, any model of parental fertility choice based on quality-quantity tradeoffs faces problems. What is more plausible is to argue that the costs of child quantity increased in the second half of the nineteenth century as a result of compulsory schooling laws. Doepke (2004) argues that other government policies (such as education subsidies) could not have had a similarly large influence. Yet we do not know that government intervention was crucial in moving children out of the factories and into the classrooms. For the US, there is some evidence that state schooling laws had little influence on child labor (Moehling 1999). In the UK, Kirby (1999) argued that child labor laws came in at the same time when technological change made the use of children in mining much less useful, and that there was not much of a causal role for
government legislation in reducing employment rates.

If the importance of government intervention is confirmed, examining the economic and other factors behind the adoption of child labor laws becomes crucial. What is also missing is convincing evidence that higher net costs of child rearing (principally through lower employment opportunities) were important in reducing fertility. 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. We have little information on what determined completed fertility rates, educational investment, age at marriage and the like in the industrializing cities of the North. More detailed demographic analysis of the fertility choices of the working class – combined with information on rates of school attendance etc. prior to and after the introduction of the compulsory schooling laws – could do much to further our understanding of how robust the empirical foundations of human-capital led interpretations of the Industrial Revolution are.

The history of fertility decline represents a further challenge to unified growth theory, and to much of economic demography. Little if any of the changes predicted by standard models such as in Becker and Barro (1987) are in line with historical facts. Most of the fertility decline was concentrated in a few decades, starting in 1870 and accelerating after 1890. In general – with important exceptions such as in the case of France – mortality declines preceded the fall in fertility by decades (Lee 2003, Coale and Watkins 1986). In some countries, such as the UK, 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 declines in infant mortality before fertility started to decline (Sweden, Belgium, Denmark); in others, both series show a concurrent downward movement (France, Germany, Netherlands). Differences in levels are equally puzzling: Swiss, Belgian and Swedish birth rates around 1850 were on the order of 30 per 1000, whereas in the Netherlands, Austria and Germany these were around 35 per thousand.

Finding an economic reason for fertility decline has not been easy, and there is currently no consensus on the principal contributing factors (Alter 1992). The biggest project on the fertility transition, the Princeton European Fertility Project (EFL), concluded that there was no clear link between socio-economic factors and fertility change. Instead, ethnic, religious, linguistic and cultural factors appeared to be dominant (Coale and Watson 1986). The economic value of children, as far as it can be determined, did not change in such a way as to help in explaining the decline in fertility rates (Knodel and van deWalle 1986). 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 EFL is the remarkable similarity in the timing

---

of the transition.31

Studies that go beyond the broad aggregates and look at regional data from, for example, a single European state such as Bavaria, have sometimes reached different conclusions (Brown and Guinnane 2002), assigning a greater role to the opportunity cost of women’s time, while at the same time also documenting the effect of other factors such as religion and political affiliation. The statistical basis for some of the EFP’s conclusions may be less robust than had previously been assumed.32 Independent of whether new, more disaggregated studies can find a role for economic factors in fertility change, the very simultaneity of the drop in reproduction rates across Europe makes it unlikely that economic factors can ever be assigned a dominant role. 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, it is a challenge that there are so few economic factors that have a clear bearing on fertility outcomes.

d. What sustained growth: science, technology, and “useful knowledge”

One organizing concept that has proven hard to model formally but without which no historically accurate picture of modern growth can be formed is the connection between science and technology in the Industrial Revolution and beyond. Historical scholarship has bifurcated here into a minority view, which continue to view science and scientific culture as crucial to the Industrial Revolution (Musson and Robinson, 1969; Jacob, 1997) and a majority, which has dismissed the role of science as epiphenomenal and marginal (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 most the prominent breakthroughs in manufacturing, especially in the mechanical processing of textiles were not based on much more science than Archimedes knew, and that in other areas of progress, such as steam power and animal breeding, progress occurred 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 those who do not is more than a hackneyed dispute between a glass that is half full or half empty, because the glass started from almost empty and slowly filled in the century and 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, 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, it is hard to argue that it would have transformed into a continent-wide process of 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

31 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”.

32 A much larger research project on German fertility decline is now under way (Ogilvie et al. 2005), using that country’s extraordinarily rich data sources.
been discerned in the mid eighteenth century. It is equally clear, however, that in the 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 telegraph and the growing understanding of fatty acids in chemicals take place in the final decades of the classical Industrial Revolution. Trial and error, serendipity, and sheer intuition never quite disappear from the scene, but the ability to know more about how and why a technique works makes it far easier to refine and debug a new technique quickly, adapt it to other uses, and come up with variations and recombinations that would not have occurred otherwise. In chemicals, steel, electricity, food processing, power engineering, agriculture, and shipbuilding technology, to name but a few, the ties between formally educated who tried to understand the natural phenomena and regularities they observed and the people whose livelihood depended

The underlying institutions that made this growing collaboration possible have been investigated at great length. Although IPR’s sure were of some importance, they cannot possibly explain the entire process, as we have seen. Instead a deeper and more encompassing social phenomenon was at play here, namely a growing interaction flows of information and interaction between people who made things (entrepreneurs and engineers) and people who knew things. Not only that this interaction meant that the best that 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 looking for improved bleaches or more efficient engines.

By 1815, the need for this kind of collaboration had become a consensus, and the European economies competed with one another in encouraging them. 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 would 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 kits most prominent entrepreneurs and engineers. Less well known but equally significant were the Spitalfields Mathematical Society, founded in 1717, 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 (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. The decades after 1815, then, were the ultimate triumph of the Baconian vision, that formed the basis for the founding of the Royal Society in 1660. The hopes and aspirations of men like Thomas Sprat and Robert Hooke were slowly

---

33In 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 in 1804,
becoming reality. To achieve this triumph, Europe had to undergo changes in its institutional set-up of 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 it about.

e. Capital and the rise of new forms of business organization

Growth in the 19th century relied heavily on capital investment. Savings as a share of national income increased in the UK after 1750. They had to – to equip the ever larger cohorts entering the workforce with machines and buildings, with a roof over their heads and shovels in their hands, savings had to increase. Eventually, as population growth rates declined, these high savings rates started to increase capital available per member of the workforce. Much of the post 1850 growth was driven by capital deepening. Calculated as a simple residual, not much tended to be left for the role of TFP, of output change not driven by more capital or more labor. That is why Abramovitz and David (1973), in their review of US economic history during the 19th and 20th centuries, spoke of the “rise of TFP” over the period.

There is substantial agreement that the standard assumptions about the separability of capital and technological change are too facile. One way to go beyond the standard Cobb-Douglas function in growth accounting was explored by Crafts (2004), who examined the change in the relative cost of capital that made adoption easier. In the standard TFP framework, we would attribute the economic effect of this demand response entirely to "more capital", which to some extent will be "better and cheaper capital". The element that Crafts captures is that "more capital" only became affordable because of technological change. The method, pioneered by Oliner and Sichel (2000), simply adds the capital deepening in the new "revolutionizing" sector to the TFP estimate, thus deriving the total contribution of technological change. In many subperiods, the results are dramatic. For 1830-60, for example, adding capital deepening in railways implies that the role of technological change was eight times larger than previously calculated. It should be noted that there is now an upward-instead of downward-bias in the TFP calculations, since the underlying assumption is that without technological change, there would have been no investment in transport whatsoever – which is unlikely to be true, imperfect as early models were. Long after the railroad were introduced, horses remained the staple source of transportation power.

Allen (2005) goes beyond the Cobb-Douglas production function altogether, and uses a translog production function for industrializing Britain instead. He calibrates a model for the UK in the eighteenth century, and finds it to have very low elasticities of substitution in factor inputs; the production function is almost Leontief. The implication then is that almost all capital accumulation is driven by technological change – without the chance to put new machinery in place, Britain would not have invested more in additional capital goods. Allen couples this setup with one where technological change drives up profits, which are then reinvested. Somewhat unsurprisingly, given the extreme nature of the production function’s parameters, technological change now accounts for basically all of the output gains after 1700. It is difficult to understand why more capital would have

---

34Abramovitz and David 1973; Feinstein 1978.
been "useless" in the absence of technological progress – few observers would agree that the marginal return to more capital in the UK economy in 1700 would have been essentially zero. While there is no agreement yet on how to model and calculate the interaction between capital deepening and technological progress, it is clear that new capital goods put in place partly to equip the larger cohorts were also better. Technological progress raised the marginal product of capital, while capital goods embodied much of the new technology. This complementarity between capital and technological progress is central to the historical development of the British economy.

Hand-in-hand with a greater role for capital in the economy at large went changes in business organizations. Textile firms were typically small in size; Lancashire had thousands of spinning and weaving firms, all competing vigorously with each other (Leunig 2003). Large-scale production units had not been unknown altogether before 1800 (Mokyr 2001). Factories started to replace the putting-out system of home production gradually, and both forms of organization existed side-by-side for a long time. Yet the minimum efficient size of steel production, of chemicals plants and electrical engineering units in during the “Second Industrial Revolution” dwarfed anything that went before it. One school of business historians has written about industrial development as the history of the rise of big business (Chandler 1990). Lamoreaux, Raff and Temin (2003) have argued that we should understand changes in business organization as a response to changing transport and information costs, in the tradition of Coase and Oliver Williamson.

h. Institutions and the new economy: factory acts, child labor acts, public education
i. Political reform, and complementarity of human and physical capital
j. Summary
References (incomplete)


Cervellati, Matteo and Sunde, Uwe. 2006. “Health, Development and the Demographic Transition.”
Unpub. ms., January.


Delong, Bradford, and Andrei Shleifer. 1993. “Princes and Merchants: City Growth Before the


Stark, Rodney. 2005. *The Victory of Reason: How Christianity led to Freedom, Capitalism, and


____, Time Use in Eighteenth Century Britain, 1750-1830, Oxford: OUP.


