



HAL
open science

Energy, knowledge, and demo-economic development in the long run: a unified growth model

Emmanuel Bovari, Victor Court

► **To cite this version:**

Emmanuel Bovari, Victor Court. Energy, knowledge, and demo-economic development in the long run: a unified growth model. 2019. hal-01698755v3

HAL Id: hal-01698755

<https://hal.science/hal-01698755v3>

Preprint submitted on 25 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Energy, knowledge, and demo-economic development in the long run: a unified growth model

Emmanuel Bovari • Victor Court*

This version: March 2019

Working paper: please do not circulate nor quote without authors' permission

Abstract This article provides a knowledge-based and energy-centred unified growth model of the transition from limited to sustained economic growth. We model the transition between: (i) a pre-modern organic regime defined by limited growth in per capita output, high fertility, low levels of human capital, technical change generated by learning-by-doing, and rare general purpose technology (GPT) arrivals; and (ii) a modern fossil regime characterized by sustained growth in per capita output, low fertility, high levels of human capital, technical change generated by profit-motivated R&D, and increasingly frequent GPT arrivals. The associated energy transition results from the endogenous shortage in the availability of renewable resources, and the arrival of new GPTs that, together, redirect technical change towards the exploitation of previously unprofitable exhaustible energy. A calibrated version of the model replicates the historical experience of Britain from 1700 to 1960. Counterfactual simulations are performed to characterize the impact of the energy transition on the timing and magnitude of the British economic take-off. Another simulation exercise compares the different trajectories of Western Europe and Eastern Asia to determine which parameters of our model are the most crucial to reflect the diverging dynamics of these two world regions.

Keywords Unified growth theory • Directed technical change • Useful knowledge • Energy transition • Demographic transition

JEL Classification C68 • J13 • J24 • N10 • O31 • O40

*E. Bovari: University Paris 1 Panthéon Sorbonne, Department of Economics, Centre d'Économie de la Sorbonne, 106-112 boulevard de l'Hôpital, 75013 Paris, France; and Chair Energy & Prosperity, Institut Louis Bachelier, 28 place de la Bourse, 75002 Paris, France (e-mail: Emmanuel.Bovari@univ-paris1.fr). V. Court: Science Policy Research Unit (SPRU), Business School, University of Sussex, Jubilee Building, Brighton, BN1 9SL, United Kingdom; and Chair Energy & Prosperity, Institut Louis Bachelier, 28 place de la Bourse, 75002 Paris, France (e-mail: v.court@sussex.ac.uk). We benefited from the fruitful comments of seminar and conference participants of the 67th AFSE Congress, 6th ISEFI, 18th WEHC, 5th FAERE and the annual conference of the Climate Economics Chair. We thank Joseph Enguehard for his helpful comments on an earlier version of this article, and Noah Ver Beek for correcting the spelling and grammar of this manuscript. All remaining errors are our own. Emmanuel Bovari gratefully acknowledges the financial support from the Agence Nationale de la Recherche (grant DS0207).

1 Introduction

1.1 The need for a Unified Growth Theory and its development so far

Many agree with Galor's arguments to justify the need for a unified analytical framework able to explain both the occurrence of the *Great Divergence* and its persistence over time. Indeed, as long as the economic take-off some countries encountered two hundred years ago remains a mystery, confidence in modern economic growth theories will remain fragile. Moreover, the factors that prompted the take-off of the currently developed economies must be identified to allow a comprehensive understanding of the obstacles preventing developing countries from reaching a state of sustained economic growth (Galor, 2005, p. 176). Traditionally, the many structural changes occurring in an economy transitioning from limited to sustained growth have been studied in two-sectors models (agricultural vs. manufacturing sectors, or more generally traditional vs. modern sectors). Yet, all these models necessarily require exogenous shocks on prices, productivities, mortality rates, or preferences in order to generate a transition from limited to sustained economic growth (e.g., Hansen & Prescott, 2002; Mourmouras & Rangazas, 2009).

Galor & Weil (2000) proposed the first model able to deliver a purely endogenous transition from limited to sustained growth. This benchmark Galor-Weil model fostered further research which now form a *Unified Growth Theory* (UGT). Different mechanisms have been proposed to explain the economic take-off process, such as: the scale effect of population on technical change in Galor & Weil (2000), Yakita (2010), Galindev (2011), Fröling (2011), and Strulik *et al.* (2013); the Darwinian selection of child-quality oriented individuals in Galor & Moav (2002); the Darwinian selection of entrepreneurial-oriented individuals in Galor & Michalopoulos (2012); the improvements in gender equality in Lagerlöf (2003); the decreasing demand for child labor in Doepke (2004); the decreasing child mortality rate and consequent improvement in life expectancy at birth in Cervellati & Sunde (2005); the improvement of health (but not longevity) in Hazan & Zoabi (2006); the increasing productivity of agriculture in Strulik & Weisdorf (2008); the increasing size of markets in Desmet & Parente (2012); and the increase of general knowledge in O'Rourke *et al.* (2013). An important feature of UGT is that the transition from a rather stagnant Malthusian regime towards sustained modern growth appears as the inevitable outcome of the growth process itself. As a corollary, in all these models it is possible to observe differential timing and magnitudes of take-off across countries as the result of deep determinants (of a biogeographical, cultural, institutional, or contingent nature), but a country cannot be locked in a stagnation trap because the take-off is inevitable by construction.¹

1.2 Missing perspective, goal, and organization of the article

There is no doubt that the numerous contributions to Unified Growth Theory have shed new light on the process of long-term economic growth, from both the theoretical and empirical perspective. However, it is precisely the path-breaking significance of the benchmark model of Galor & Weil (2000) that might have obscured other dimensions of the shift towards modern economic growth. In particular, a key issue usually overlooked in economic growth theories is the role of energy. Indeed, to the best of our knowledge, energy is absent from all unified growth models of the literature apart from Fröling (2011). All these models are therefore supposed to explain the Industrial Revolution without considering the role of energy, and in particular the associated energy transition towards fossil fuels. This view contrasts with the work of many economic historians, such as Pomeranz (2000), Fouquet (2008), Allen (2009), Kander *et al.* (2013), and Wrigley (2016), who place a great emphasis on the role of coal to explain the early economic take-off of England towards sustained economic growth (see Subsection 2.3.2).²

¹Nguyen Dao & Dávila (2013) argue that this is mostly because technical losses are not possible in unified growth models, whereas in reality technology must not only be acquired but maintained too. History indeed provides many examples of technology losses for geographic, cultural, or political reasons, see Diamond (1997, pp. 257–258, pp. 312–313), and Morris (2010, pp. 413–417) respectively.

²Several economic historians, such as Debeir *et al.* (1991), Crosby (2007), and Smil (2017), go further and make energy central to their analysis of the entire history of human societies.

In Fröling's (2011) model, the final energy supply is a constant elasticity of substitution (CES) aggregate of final energies made from primary resources of coal and biomass that only require labor for their extraction. Knowledge enhances the conversion efficiency of coal primary energy into final energy, but it cannot augment the conversion efficiency of primary energy from biomass, which contradicts historical evidence (Kander & Stern, 2014). The resulting aggregate final energy supply is then combined with labor and land in a Cobb-Douglas function to produce the final output good. It is assumed that another stock of knowledge enhances total factor productivity (TFP) in the production of the final output, which implies that TFP growth augments final energy, labor, and land at the same rate in the final good production. Such an hypothesis hardly cope with historical facts regarding relative productivities of production factors. Similarly, the model assumes a constant overall allocation to research and development (R&D) over time, and a constant split of this allocation between final and energy research sectors. The addition of these different drawbacks might explain that the global simulations of Fröling's (2011) model have difficulty to fit with historical data.

Taking advantage of recent advances in unified growth theories, the present article aims to provide a unified growth model able to better take into account the crucial role of energy for the transition from limited to sustained growth. It thus helps to reconcile economic growth theory with historical facts regarding the role of energy emphasized by historical economists. Section 2 presents several empirical facts regarding the relation of useful knowledge, demography, and energy with the economic take-off towards modernity. Based on these insights, Section 3 develops a knowledge-based and energy-centered unified growth model. The balanced growth path of this model is analyzed in Section 4. Calibrations of the model to the historical experiences of England, Western Europe and Eastern Asia are performed in Section 5. Finally, a summary of the contributions of this article is given in Section 6, along with recommendations for future research.

2 Demography, useful knowledge, and energy transition

In this section, we briefly survey the literature centered around demography, knowledge, and energy to highlight the key role that these variables play in economic growth. Performing this analysis allows us to identify the main building blocks of the theoretical framework presented in Section 3.

2.1 Demographic transition: choosing among controversial issues

A very important feature of unified growth models is the proposal of an endogenous demographic transition associated with the take-off from limited to sustained economic growth. Most of these models assume that parents consciously evaluate a trade-off between the number of children they want to have and the level of education they choose for their children. However, as detailed in Appendix A.1, the empirical evidence precludes any consensus on the existence and the causes of such a child quantity-quality trade-off. In the model developed in Section 3, our population module is an adaptation of Strulik *et al.*'s (2013) formulation to which we have added an impact of the stock of useful knowledge on the efficiency of the human capital production. As a result, the endogenous demographic transition of our model is triggered by an initial per capita production increase that does not hinge on a conscious quantity-quality trade-off. However, once the economic take-off is established, there is a decrease in fertility and an increase in education expenditure as a result of the changing technical environment. Such an approach seems the most appropriate to comply with historical evidence while remaining as neutral as possible on the unresolved debate surrounding both the existence of a conscious child quantity-quality trade-off and the underlying mechanism of this arbitrage.

2.2 Useful knowledge rather than human capital of general population

In theoretical models of the modern regime, human capital is considered central to sustained growth. This causes many unified growth theories to make human capital also crucial to explain the take-off from limited to sustained growth. However, as detailed in Appendix A.2, Mokyr (2011, p. 486)

emphasizes the importance of schooling, and the resulting improvements of human capital, to explain the *second* phase of the Industrial Revolution (i.e. after 1850), but it is clear that an alternative mechanism is needed to explain the early take-off of Britain associated with the *first* phase of the Industrial Revolution. Recent theoretical and empirical literature seems to identify useful knowledge as a more likely cause of the intellectual changes preceding the Industrial Revolution.

Several scholars, such as Goldstone (2009), Jacob (2014), and Mokyr (2011, 2017), indeed attribute much of the credit for the burst of innovations and accelerated diffusion of best practices after 1750, not to mass education in general, but to the scientific culture that emerged with the European Enlightenment. They argue that Western European societies were particularly dynamic and inclined to see technical breakthroughs in the eighteenth century thanks to the increase, or propagation during the previous two hundred years, of printing books, publishers, scientific societies, university networks, relatively accessible public lectures, and growing day-to-day exchanges between scientists, engineers, and craftsmen. Hence, these authors explain the success of the British Industrial Revolution through changes in the intellectual, social, and institutional background environment that enabled Britain to acquire a modern science culture. This change in the intellectual environment permeated the whole society and was decisive in converting *useful* knowledge — i.e., ideas and inventions that often came from other European countries (such as France, Belgium, the Netherlands or Germany) — into workable innovations that were rapidly transformed into practical technologies necessitating *applied* knowledge — i.e., skills — yielding profits to their developers.

³

The empirical study of Squicciarini & Voigtländer (2015) is the first to provide systematic evidence for Mokyr's hypothesis about the importance of useful knowledge for industrialization. As a proxy for scientific elites, these authors use *Encyclopédie* subscriber density and show that this measure of 'upper-tail knowledge' is strongly associated with other indicators of local scientific activity, both before and after the *Encyclopédie* was printed in the mid-eighteenth century. Squicciarini & Voigtländer (2015) then show that upper-tail knowledge is a strong predictor of city growth after the onset of the French industrialization. Furthermore, by joining data on British patents with a large French firm survey from the 1840s, it appears that scientific elites indeed caused productivity increases in innovative industrial technology which were then associated with economics growth. On the contrary, Squicciarini & Voigtländer (2015) show that literacy levels representing human capital of the general population are associated with development in the cross-section, but they do not predict growth. Furthermore, de Pleijt *et al.* (2019) perform a quantitative assessment of the effect of industrialization, captured by the number of steam engines per person installed in England by 1800, on the average working skills of 2.6 millions workers. They find support for a causal relationship going from the diffusion of steam engines to higher skill-demand. Moreover, de Pleijt *et al.* (2019) show that early industrialization was negatively associated with primary schooling and with the acquisition of literacy skills for women. Overall, de Pleijt *et al.*'s 2019 findings tend to (i) confirm Mokyr's (2011) conclusion that basic education and the associated human capital was not a key ingredient in England's early industrialization; and (ii) show that the causal relationship running from useful knowledge (i.e., scientific ideas and inventions) to industrialization, highlighted by Squicciarini & Voigtländer (2015), was complemented by a causal relationship going from industrialization to applied knowledge (i.e., skills).

Recent unified growth models have taken into account the importance of useful knowledge. Strulik *et al.* (2013) propose a setting where technical change is initially only due to learning-by-doing *prior* to the apparition of an expanding input variety R&D sector that then fosters sustained

³It is important to understand that all these scholars do not denigrate the many scientific breakthroughs that episodically originated in China and Islamic countries. They instead highlight how early Britain was in creating a scientific culture able to transform useful knowledge into applied technical change thanks to a favorable institutional environment. Lipsey *et al.* (2005, pp. 225–289) similarly stress that the roots of mechanistic science in Western Europe lie in the emergence of the appropriate institutions associated with the development of pluralistic societies in the last half of the medieval period that ultimately freed natural philosophers to seek an explanation of the world in terms of mechanical laws. Lipsey *et al.* (2005, pp. 225–289) also assert that the absence of early economic takeoff in China and advanced Islamic countries is explained by the failure of these countries to develop anything like modern science because of inappropriate institutions determined, in part, by their monolithic state structures and religious dogmas.

economic growth. O'Rourke *et al.* (2013) introduce a stock of useful general knowledge whose level impacts the cost of innovation in a Schumpeterian R&D sector. As will be shown in Section 3, we build on these two recent articles and explicitly introduce a stock of useful knowledge when developing our model. More precisely, we rely on the innovative theoretical approach of Schaefer *et al.* (2014) that tracks down the history of technical improvements that accumulate in a stock of useful knowledge. The latter stock then shapes the pattern of arrival of *General Purpose Technology* (GPT), which in turn impact the rate of change of all kinds of technical changes — distinguished between learning-by-doing and profit-motivated R&D as in Strulik *et al.* (2013) and O'Rourke *et al.* (2013).

2.3 Coal and the transition towards sustained economic growth in Great Britain

As already mentioned, apart from Fröling (2011), energy is absent from all unified growth models. In our view, this omission of energy in growth theories is due to (i) the assumed unimportance of energy for the growth process, justified by the relatively small share of energy cost in the national income, and (ii) the inattention of economists to the different stage of energy refining, from primary, to final, to useful. For the sake of space, these problems are respectively addressed in Appendix A.3 and Appendix A.4.

In the specific case of Great Britain, the central role of coal in explaining the early take-off of the economy is obvious for many economic historians. First, it is recognized that, from the sixteenth century onward, the extensive use of slaves to extract natural resources (sugar, tea, tobacco, coffee, fur, and more specifically guano, wood, and cotton) from the New World implied an Atlantic Trade that flooded Western European markets with new exotic products. This expansion of European markets, and the institutional changes that accompanied it, helped lead several Western European countries towards an *Industrious Revolution* characterized by an increasing prevalence of wage labor (i.e., labor offered on a specialized and relatively well-remunerated market instead of household labor) (de Vries, 1994).⁴ In particular, for two Western European proto-industrial nations, Britain and the Netherlands, wages broadly increased from the sixteenth to the eighteenth centuries relatively to other European Nations and development cores in other parts of the world. This so-called *Little Divergence* within Europe implied that incentives for labor-saving technologies were more important in Britain and the Netherlands compared to other European nations, while nonexistent in China, Japan or India where labor remained relatively cheap (Allen, 2011; Allen *et al.*, 2011).⁵

Simultaneously, because proto-industry relied heavily on wood fuel, wood scarcity (leading to price increases) was a frequent occurrence in most of Western Europe, especially in Britain (Pomeranz, 2000, pp. 220–223). At these times of consequent incentives for both labor-saving and woodfuel-saving technologies in Western Europe, Western European countries, and here again most notably Britain, were lucky to be endowed with large and relatively accessible deposits of coal.⁶ But substituting wood and charcoal with coal in heat generation was well known for centuries and did not lead to an industrial revolution. What mattered, and constituted a major breakthrough, was being able to turn the heat from coal combustion into mechanical energy with the steam engine. Allen (2011, 2009) comprehensively argues that the British Industrial Revolution originated in the willingness and ability of its people to (i) tap their favorable coal endowment thanks to financial incentives represented in relative factor prices (of labor, capital, wood, and fossil fuels, see Figure 1), and (ii) apply knowledge brought by science (as already highlighted in Subsection 2.2) to convert coal into useful work (i.e., mechanical energy), and in doing so direct and foster sustained technical change during the Industrial Revolution.

In the same line of thoughts, and after comparing the role of energy in Europe and other parts

⁴The concept of Industrious Revolution has a quite different meaning for several authors, in particular for the inventor of the term, Hayami (2015), to whom it implied the emergence of a labour-intensive agriculture compatible with the rise of a market in Japan.

⁵See Humphries (2013) and Stephenson (2018) for a critique of Allen's real wages estimates, to which the later replies respectively in Allen (2015, 2019).

⁶In China, coal deposits were distant from the major manufacturing regions of the lower Yangzi and the south (Pomeranz, 2000, p. 65).

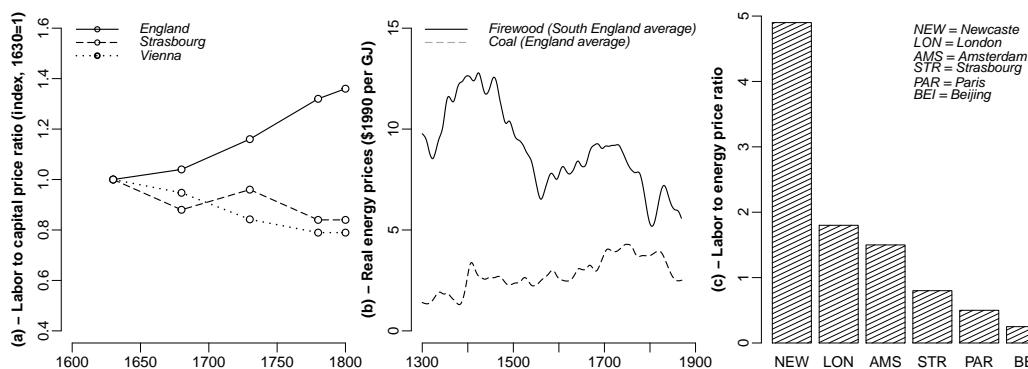


Fig. 1 Panel (a) - Labor to capital real prices in England (*solid line*), Strasbourg (*dashed line*), and Vienna (*dotted line*), 1630–1800; Panel (b) - Real prices of firewood (*solid line*) and coal (*dashed line*) in England, 1300-1870, time series smoothed with Hodrick-Prescott filter (factor 1000); Panel (c) - Labor to energy real prices in six cities, early 1700s. *Data sources*: Allen (2009, p. 139, p. 140 for panels (a) and (c) respectively) and Fouquet (2011) for panel (b)

of the world over the last five centuries, Kander *et al.* (2013, p. 366) conclude that it is hard to imagine anything like modern economic growth occurring without the adoption of fossil fuels, beginning with coal. They further emphasize that they view the transition to fossil fuels both as a necessary condition and an enabling factor *leading* to modern growth (*Ibid.*, italic emphasis present in original). Kander *et al.* (2013, pp. 367–368) then assert that the high complementary between coal, the steam engine, and the iron industry was crucial in delivering unprecedented amounts of mechanical energy that structurally reshaped the industrializing of British society.⁷ They argue that for the first time in history it was possible to reliably and in a controlled form convert heat to motion, equipping people with inanimate ‘energy slaves’ (machines). Steam engines enabled a large concentration of energy in time and space; they were more powerful than previous sources of kinetic energy and much easier to control. Steam engines saved labor, and initiated a capital-deepening growth path where one worker could be in command of ever greater amounts of power.⁸ As Berg (1994, p. 207) notes, it was not the spinning machinery itself (in operation since 1770) that made England a leader in textile production, but rather the application of steam power to spinning, to water and surface transport, to brick-making, grain-threshing, iron-making, shovelling, construction, and all sorts of manufacturing processes that transformed the British economy. Hence, the entire process of capital-deepening growth was almost wholly reliant on fossil fuels and eventually led to both increased incomes and a dynamic that has continued to raise incomes (Kander *et al.*, 2013, p. 368).

To quantify the importance of coal as a source of both heat and mechanical energy in the transition from limited to sustained economic growth, Malanima (2016, pp. 96–99) followed the seminal contribution of Wrigley (1962) in order to estimate land- and labor-savings due to coal use, respectively as a source of heat and mechanical energy, in England & Wales on the period 1560–1913.⁹ The results presented in Figure 2 exhibit two distinct historical phases. During the first one, which lasted from 1560 until about 1830, the use of coal was mainly land-saving. It is only during the second phase, from 1830 to 1913, that coal was really both land and labor-saving. Covering both

⁷It is because they completely miss the importance of the synergy between coal, the steam engine, and the iron industry that Clark & Jacks (2007), relying solely on the contribution of coal mining rents to national income, can claim that the possession of coal reserves made a negligible contribution to the success of the British Industrial Revolution.

⁸It is often argued that the first steam engines were extremely inefficient, and that the end-use cost of the mechanical energy produced by a steam engine was not lower than that of a windmill, waterwheel, or laborer. This so-called paradox is solved when one considers that, unlike the prime movers that preceded it, a steam engine is both mobile and continuously operational. A calculation of the production cost for the same ‘controllable energy unit’ would clearly show the competitive advantage of coal combined with a steam engine over all competing prime movers of the same period.

⁹As noted by Malanima (2016, p. 96), usual social savings calculations based on relative costs of old and alternative technologies appear to be quite impossible here since they would require the computation of counter-factual wood prices and labor wages in a theoretical British economy where coal was absent.

phases from 1800 to 1900, the land-related (resp. labor-related) savings grew from 1 to 14 times the extent of the entire country of 15 million hectares (resp. from 1 million to almost 290 million workers while the English labor force was about 13–14 million workers in 1900). These estimates strongly support Wrigley's (2016, pp. 2–4) claim that the energy required to produce iron and steel on a large scale or to construct and operate a railway system implied that it was unreasonable to expect that it could be secured from the annual flow of energy derived from plant photosynthesis. As a corollary, for Wrigley, an Industrial Revolution could not have been accomplished as long as mechanical energy continued to be provided principally by human and animal muscle.

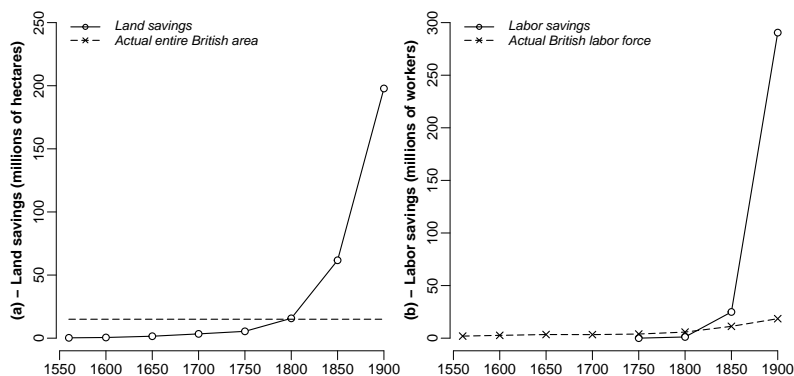


Fig. 2 Panel (a) - Land saved thanks to the use of coal in England & Wales (*solid line*) and actual entire British area (*dashed line*), 1560–1913; Panel (b) - Labor saved thanks to the use of coal in England & Wales (*solid line*) and actual British labor force (*dashed line*), 1560–1913. *Data sources:* Malanima (2016, p. 98) and activity rate of 45% applied to the population level of Fouquet (2014) to compute the actual labor force.

Based on the evidence of the key role of energy during the Industrial Revolution, we explicitly introduce final energy carriers within our modeling setup, the production of which depends on specific primary resources and technologies. These features will allow us to (i) distinguish several energy forms (and their corresponding prices), (ii) describe the technical change associated with fossil energy as a response to pre-modern renewable energy shortage, and (iii) emphasize the key role of the abundant and cheap fossil energy supply during the Industrial Revolution and the whole modern regime.

3 Theoretical model

Building on the literature review of Section 2, we present in this section a unified growth model in which demography, knowledge, and energy interact to explain the transition from limited to sustained economic growth. In our framework, time is discrete and indexed by t .

3.1 Household's preferences and supply of physical and human capital

We consider an economy populated by three overlapping generations: children, young adults, and old adults. In the first period of life (childhood), $t - 1$, we suppose that children do not work at all and thus earn no income. Furthermore, regardless of their education, child-rearing requires a fraction of the parental unit time endowment that is logically increasing with the number of children. In addition, financial expenditure (but not parents' time attention) increase with children's formal education. In the second period of life (parenthood), t , young adults, whose number is denoted by N_t , are endowed with one unit of time that can be converted into one unit of labor force. Young adults allocate their unitary time endowment between child rearing and effective labor force participation to earn a wage. The wage income is either spent through direct consumption, saved to supply the capital investment market, or spent on children's education. In the third period, $t + 1$, retired adults only consume their savings plus interests.

3.1.1 Household's preferences

In order to derive the main results conveniently and to get closed-form solutions, we make a number of simplifying assumptions that are typical for unified growth, overlapping generation models (e.g. [Strulik *et al.*, 2013](#)). In each household: (i) there is one unisex parent to avoid matching issues, (ii) newborns are a continuous number denoted by b_t to avoid indivisibility issues, and (iii) the motive for child expenditure is non-operational to avoid the maximization of dynastic value functions.¹⁰ Adults' preferences are assumed to be represented by a utility function defined over: (i) household's immediate consumption, c_t , (ii) future consumption during retirement, c_{t+1}^r , (iii) births per capita, b_t , determining family size, and (iv) the future level of human capital, h_{t+1} , that each child receives through present education. Thus, the representative household's utility function writes

$$u_t = \log(c_t) + \chi \log(c_{t+1}^r) + \rho \log(h_{t+1}) + \eta \log(b_t), \quad (1)$$

where the positive parameters χ , ρ , and η capture, relative to current consumption, the elasticities of utility with respect to future consumption during retirement, the human capital of children, and the family size.

During adulthood, the revenue of the representative household comes from two sources: labor wages and patenting revenue from the R&D sector. The latter is driven to zero under the free-entry condition in R&D and our market design for capital goods production presented hereafter in [Subsection 3.2.3](#).¹¹ Hence, for adults, revenue only results from labor activities, rewarded by the competitive market wage, w_t , per efficiency labor unit h_t . We can therefore define the potential income, $z_t \equiv w_t h_t$, as the labor earned if the entire time endowment is devoted to labor force participation. As we assume that only time is required to produce children, we define τ as the fraction of the representative household's time endowment that is required to raise one child, regardless of its level of human capital. Hence, raising one child shall be seen as an opportunity cost valued τz_t . In addition, building the human capital of children, h_{t+1} , requires formal education, which is conceptualized as a financial expense expressed in final good units, e_t , per child.¹² Finally, the generation in adulthood might choose to dedicate a fraction of their income in savings, s_t , in order to fund consumption while retired. Thus, in summary, one can write the budget constraint faced by an individual in parenthood as follows,

$$c_t + s_t + b_t e_t \leq z_t [1 - \tau b_t]. \quad (2)$$

Furthermore, we assume that education expenditure e_t is converted into human capital h_{t+1} through a schooling technology that controls for schooling costs, approximated by w_t ,¹³

$$h_{t+1} = A_E(Q_t) \frac{e_t}{w_t} + \bar{h}, \quad (3)$$

where \bar{h} represents informal human capital acquired without formal education,¹⁴ and $A_E(Q_t)$ is an endogenous productivity rising with the level of total applied knowledge of the economy, Q_t . We assume that the education technology shows decreasing returns with respect to this stock of applied knowledge, so $A_E(Q_t)$ satisfies the following conditions: $A_E(0) = 0$, $\partial A_E / \partial Q > 0$ and $\partial^2 A_E / \partial Q^2 < 0$. In other words, we assume that the overall level of applied knowledge, Q_t , is a good proxy for various phenomena that positively affect the efficiency of schooling, such as the

¹⁰This last point means that the parent's motivation to spend on their children's education is not driven by the anticipation of the increase of children's utility caused by this expenditure, but by a general desire for having 'higher quality' children.

¹¹Relaxing these hypotheses would not change qualitatively the behavior of the representative household.

¹²In other words, the provision of aggregate education services $b_t e_t N_t$ has a cost of $b_t e_t N_t$ units of final good. This technical assumption ensures market clearing on the goods market, as illustrated hereafter in Eq. (37).

¹³One can think of w_t as the wage of teachers, hence e_t / w_t represents efficient education expenditure.

¹⁴Such a basic human capital level can be thought as informal knowledge that children acquire through the time τ spends observing and imitating their parents and peers at work. This knowledge (of farming or a particular craft, for example) is useful, i.e. it creates human capital at level \bar{h} , but it comes for free, at no educational cost. On the contrary, e_t is a financial investment that allows the child to receive a formal education through school and the consumption of cultural goods in order to increase their human capital above \bar{h} .

rising spatial density of schools (Boucekkine *et al.*, 2007), the evolution of social norms favoring formal education, or changes in law limiting child's labor (Doepke, 2004). Throughout this paper, we assume $A_E(Q_t) = \bar{A}_E Q_t / (1 + Q_t)$, with $\bar{A}_E > 0$.

Lastly, as is typical in the OLG literature (de La Croix & Michel, 2002), we assume that the unique revenue of the current retired generation consists in the savings they accrued in former young adulthood, which can be invested either in physical capital or financial assets. In the latter case, the retired adult is ultimately rewarded with an interest rate, r_{t+1} . Establishing the usual no arbitrage assumption between the returns of both (physical and financial) assets, this interest rate shall equal the price charged to productive firms for the use of physical capital *minus* the depreciation rate of capital. Moreover, at equilibrium, as each individual agent does not hold any debt, she only invests in the physical capital asset. As a result, and for the rest of this paper, we assume that adulthood generations only invest in physical capital. This assumption allows us to define the budget constraint of the old generation as

$$c_{t+1}^r \leq (1 + r_{t+1})s_t. \quad (4)$$

The behavior of the representative households can be formally introduced through an optimization problem introduced here for clarity.

Problem 1 (HH - Household) *The representative young adult planner seeks to maximize the utility function define in Eq. (1) under five constraints: the inter-temporal budget constraint in Eq. (2); the budget constraint while retired in Eq. (4); the education technology defined in Eq. (3); and two non-negativity constraints (hereafter implicitly assumed) on the number of children and education such that $(b_t, e_t) \geq 0$. Hence, taking factor prices w_t and r_t as given, the household's problems writes*

$$\begin{aligned} \max_{c_t, c_{t+1}^r, s_t, b_t, e_t \geq 0} \quad & u_t = \log(c_t) + \chi \log(c_{t+1}^r) + \rho \log(h_{t+1}(e_t, Q_t)) + \eta \log(b_t) \\ \text{s.t.} \quad & c_t + s_t + b_t e_t \leq z_t [1 - \tau b_t], \\ & c_{t+1}^r \leq (1 + r_{t+1})s_t \end{aligned}$$

thus setting per capita levels of consumption as a young adult, c_t , consumption as a retired adult, c_{t+1}^r , savings, s_t , fertility, b_t , and investment in education, e_t . Combined with the endogenous demographic dynamics described in Eq. (8), the representative household ultimately sets the supply of human capital, H_t .

We immediately assume that the budget constraint of the retired generation, Eq. (4), is saturated due to the strict monotonicity of utility in consumption. The remaining first order conditions of this problem yield the following optimal decisions regarding per capita consumption, savings, and number of children,

$$c_t = \frac{z_t}{1 + \chi + \eta}, \quad s_t = \chi c_t, \quad b_t = \frac{\eta c_t}{e_t + \tau z_t}. \quad (5)$$

It is clear from Eq. (5) that increasing education reduces fertility, which is a direct illustration of a quality-quantity trade-off. Turning now to child quality, there exist an endogenous knowledge threshold, \tilde{Q}_t , such that by monotonicity of $A_E(\cdot)$,

$$e_t = \begin{cases} 0 & \text{if } Q_t < \tilde{Q}_t, \\ \frac{[\rho \tau A_E(Q_t) h_t - \eta \bar{h}] w_t}{A_E(Q_t) (\eta - \rho)} & \text{if } Q_t \geq \tilde{Q}_t. \end{cases} \quad (6)$$

This threshold is defined as the solution of $e_t(\tilde{Q}_t) = 0$, that is $\tilde{Q}_t = (\bar{A}_E \rho \tau h_t / (\eta \bar{h}) - 1)^{-1}$ according to the functional form we gave to $A_E(\cdot)$. For a given h_t , whenever $Q_t > \tilde{Q}_t$, the schooling technology is sufficiently efficient to be worth investing by the representative household. Moreover, one can easily see that $\frac{\partial \tilde{Q}_t}{\partial h_t} < 0$, which means that the per capital level of human capital tends to decrease the worthwhile threshold \tilde{Q}_t of education. This fact can be interpreted as the positive effect

of average per capita knowledge, and concomitantly of teachers' training, on the overall inclination of society toward schooling. Further analyses of the optimal education choice yields the following proposition.

Theorem 1 (Education accumulation threshold) *Provided that wages are increasing, whenever the education threshold \tilde{Q}_t is reached, the economy only accumulates human capital.*

Proof. One can easily check from Eq. (6) that $\partial e_t / \partial h_t > 0$ and $\partial e_t / \partial Q_t > 0$, and from Eq. (3) that $\partial h_{t+1} / \partial e_t > 0$. Hence, given that the applied knowledge index Q_t is cumulative in our model, as soon as $Q_t > \tilde{Q}_t$, a virtuous circle is established between education, human capital, and general knowledge, such that these three variables all tend to grow together. \square

This last comment finally allows us to give a simple interpretation of the transition from limited to sustained growth in human capital. Initially, education expenditure, e_t , is null while human capital is stuck at its minimal level, \bar{h} . In such a state, the knowledge threshold \tilde{Q}_t is fully exogenous and only depends on the parameters of the model, and we have $\tilde{Q}_{h_t=\bar{h}} = (\bar{A}_E \rho \tau / \eta - 1)^{-1}$. Due to the learning-by-doing technical change introduced in Section 3.3, Q_t gradually increases but education expenditure (resp. human capital) remains null (resp. equal to \bar{h}) up to the point where \tilde{Q}_t is crossed. From then, the representative households begins to invest into education, the demographic transition sets in, and both per capita education and per capita human capital begin to rise.

Moreover, combining Eqs. (6) and (3) allows the derivation of the following law of motion for human capital per capita

$$h_{t+1} = \frac{\rho \tau A_E(Q_t)}{\eta - \rho} h_t - \frac{\rho}{\eta - \rho} \bar{h},$$

which highlights a quantity-quality trade-off with respect to fertility and education decisions.

Proposition 1 (Quantity-quality trade-off) *The higher the desire for a large family, η (resp. for an educated family, ρ), the higher (resp. the lower) fertility, b_t , and the lower (resp. the higher) future levels of human capital, h_{t+1} . In addition, the higher the potential income, $z_t \equiv w_t h_t$, the lower fertility, b_t , and the higher future levels of human capital, h_{t+1} .*

Proof. The analysis of optimal fertility in Eq. (5) and optimal education choices in Eq. (6) immediately yields the above results. \square

Lastly, summarizing the previous results reveals two fertility paths. During the pre-modern regime, the representative household does not invest in education and human capital per capita is stuck at its lower value, $h_t = \bar{h}$. Consequently, we obtain an upper bound for fertility, defined as

$$\bar{b} \equiv b_{h_t=\bar{h}} = \frac{\eta}{\tau(1 + \chi + \eta)}.$$

In the modern growth regime, human capital per capita accumulates and is associated with a demographic transition. Fertility ultimately reaches its lower bound, defined as

$$\lim_{h_t \rightarrow \infty} b_t = \frac{\eta - \rho}{\tau(1 + \chi + \eta)}.$$

3.1.2 Aggregate supply of physical and human capital

Two aggregate production factors are inelastically supplied to the productive sector by households: physical capital and human capital. The stock of physical capital is built by the aggregate accumulation of net savings of households following the usual law of accumulation

$$K_{t+1} = I_t + (1 - \delta)K_t, \quad (7)$$

where $\delta \in [0, 1]$ stands for the depreciation rate of capital, and I_t is the net investment which consists in savings of the current generation of young adults, $s_t N_t$, minus the dissavings of the current

generation of retired adults. As we preclude voluntary bequest motives, the retired generation withdraws all the remaining stock of physical capital they hold, that is $(1 - \delta)K_t$.¹⁵ The stock of physical capital for the next period is thus entirely determined by the savings of the current young adult generation, such that Eq. (7) simply write $K_{t+1} = s_t N_t$.

Turning to the supply of human capital, one can first derive the law of motion of the population of young adults, N_t , as

$$N_{t+1} = b_t N_t. \quad (8)$$

Taking child-rearing time into account, the size of the workforce, L_t , is then given by

$$L_t = (1 - \tau b_t) N_t, \quad (9)$$

whereas the aggregate human capital supply, H_t , corresponds to

$$H_t = h_t L_t. \quad (10)$$

3.2 Aggregate production

We turn now to the description of the production side of the economy. Following [Acemoglu *et al.* \(2012\)](#),

we consider a final sector in which a composite good is produced with k imperfectly substitutable final inputs. Each k final input comes from the combination of intermediate capital goods (i.e., machines), human capital, and primary energy. The different k energy types come either from renewable or exhaustible resources. In addition to machines and human capital needed to refine these primary resources into final carriers, each one of them incurs a sector-specific extraction cost.

All sectors are assumed to be perfectly competitive, with the exception of intermediate capital goods. Indeed, technical change occurring in this sector follows a Schumpeterian innovation process where quality improvements are specific to each machine line ([Aghion & Howitt, 1992](#)). Those machines are provided by monopolists owning a patent on their variety, endogenously supplanted by successful innovators in a process of creative destruction. Moreover, as described in [Section 3.3](#), this profit-motivated R&D innovation interacts with *General Purpose Technologies* (GPTs), which also shapes the level of a learning-by-doing knowledge, affecting production.

3.2.1 Final composite good production

Building on [Acemoglu *et al.* \(2012\)](#), we consider that the aggregate final good, Y_t , comes from the combination of imperfectly substitutable inputs that ultimately rely on k primary energy resources taken in a set \mathcal{K} , which are either: (i) renewable energy resources corresponding to biomass and wind/water/solar flows; or (ii) nonrenewable energy resources corresponding to fossil fuels, such as coal/gas/oil. Considering an elasticity of substitution σ between all inputs that are aggregated into the final good, we have

$$Y_t = \left[\sum_{k \in \mathcal{K}} \vartheta_k Y_{k,t}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (11)$$

where $\sum_{k \in \mathcal{K}} \vartheta_k = 1$, with ϑ_k measuring the relative economic usefulness of the input Y_k . The final good sector can then be formally described through an usual optimization problem.

¹⁵It is worth mentioning that, as exposed hereafter in [Subsection 3.2.3](#), physical capital holders are exactly compensated for the depreciation δ , such that the overall interest paid to the retired generation is the interest rate r plus the depreciation rate δ .

¹⁶Accordingly, the total population, P_t , corresponds to $P_t = 2[(1 + b_t)N_t + N_{t-1}]$.

Problem 2 (FG – Final good producer) *The final good producer is perfectly competitive and uses the technology defined in Eq. (11). Its price is chosen as the numeraire. The representative firm takes the prices of the final inputs $\{p_{k,t}\}_{k \in \mathcal{K}}$ as given to solve*

$$\max_{\{Y_{k,t}\}_{k \in \mathcal{K}}} \left[\sum_{k \in \mathcal{K}} \vartheta_k Y_{k,t}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} - \sum_{k \in \mathcal{K}} p_{k,t} Y_{k,t},$$

thus setting demand schedules in final inputs, $Y_{k,t}(p_{k,t})$ with $k \in \mathcal{K}$.

The resolution of this problem immediately yields the isoelastic demand schedules for each final input, $k \in \mathcal{K}$, that is $Y_k = [p_k / \vartheta_k]^{-\sigma} Y$.

3.2.2 Primary energy extraction and final inputs production

For each energy type, $k \in \mathcal{K}$, the provision of the final input, $\{Y_{k,t}\}_{k \in \mathcal{K}}$, results from the combination of (i) a continuum of machines of measure one, $\{x_{k,i,t}\}_{i \in [0,1]}$ with specific endogenous quality, $\{q_{k,i,t}\}_{i \in [0,1]}$, (ii) human capital, $H_{k,t}$, and (iii) a primary energy flow, $E_{k,t}$. It is worth mentioning that $q_{k,i,t}$ designates the quality level *after* innovation decisions of period t that occur *prior* to production decisions. Those elements are combined according to a Cobb-Douglas production function with constant returns-to-scale,

$$Y_{k,t} = A_{k,t} \left[\int_0^1 q_{k,i,t}^{1-\alpha_k} x_{k,i,t}^{\alpha_k} di \right] H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}, \quad (12)$$

where $\alpha_k + \beta_k + \gamma_k = 1$, and $A_{k,t}$ is the technical level achieved through learning-by-doing in sector k . The provision of the final input $\{Y_{k,t}\}_{k \in \mathcal{K}}$ is perfectly competitive. Each final input is ultimately sold, at a price $p_{k,t}$, to the final composite good sector. In addition to the cost of sector-specific machines, $p_{k,i,t}^x$, and of human capital, w_t , the final input provider must also endure a convex cost, $\Psi(\cdot)$, associated with the extraction of primary energy.¹⁷ This cost is a function of the remaining level of primary energy resource, $\mathcal{R}_{k,t}$, that is still not captured in the case of renewable energy, and still underground in the case of nonrenewable energy. In the case of a renewable resource, we have

$$\mathcal{R}_{k,t} = \mathcal{R}_{k,0} - E_{k,t-1}, \quad (13)$$

while in the case of a nonrenewable resource, we have

$$\mathcal{R}_{k,t} = \mathcal{R}_{k,0} - \sum_{a < t} E_{k,a}. \quad (14)$$

The set $\{\mathcal{R}_{k,0}\}_{k \in \mathcal{K}}$ is determined by the natural environment and corresponds to the levels of (renewable or nonrenewable) primary energy virgin resources.¹⁸ Following Court *et al.* (2018), we suppose that as each stock of resource $\mathcal{R}_{k,t}$ gradually decreases towards zero, it becomes increasingly difficult to extract primary energy. On the other hand, technical improvements can lower the

¹⁷This extraction cost might also be seen as the price charged by a perfectly competitive primary energy extracting firm. Thus, our framework is neutral regarding an integrated or segmented energy sector.

¹⁸Nonrenewable and renewable primary energies are both physically bounded by the finite character of planet Earth. This point is straightforward for nonrenewable energies that come from finite stocks. The untapped level of a nonrenewable primary resource, $\mathcal{R}_{k,0}$, formally corresponds to the Ultimately Recoverable Resource (URR). According to British Petroleum (2015), the URR is an estimate of the total amount of a given resource that will ever be recovered and produced. It is a subjective estimate in the face of only partial information. Whilst some consider URR to be fixed by geology and the laws of physics, in practice estimates of URR continue to be increased as knowledge grows, technology advances and economics change. The URR is typically broken down into three main categories: cumulative production, discovered reserves and undiscovered resource. Renewable energies are also bounded by the ultimate size of their annual flows (as an illustration, one might consider that, for a given year, the maximum solar energy ultimately recoverable cannot exceed the natural sun radiation), which is called the Technical Potential (TP) and corresponds to $\mathcal{R}_{k,0}$ in our framework. For the (IIASA, 2012, chapter 7, p. 434), the renewable Technical Potential is the degree of use that is possible within thermodynamic, geographical, or technical limitations without a full consideration of economic feasibility.

extraction cost of primary energy. We use the stock of applied knowledge, Q_t (introduced thereafter in Eq. (22) of Section 3.3), as a proxy for the technical advancements that decrease extraction costs, so as to extend the amount of economically profitable reserves out of physically bounded resources. We thus assume that all kinds of applied innovations contribute to decrease extraction costs, featuring a technical externality of applied technical change.¹⁹ Normalizing $\mathcal{R}_{k,t}$ and Q_t to their initial values, we then define $\{\Psi_k\}_{k \in \mathcal{K}}$ as follows

$$\Psi_k = \bar{\Psi}_k \left(\frac{\mathcal{R}_{k,t+1}}{\mathcal{R}_{k,0}} \right)^{\psi_{\mathcal{R},k}} \left(\frac{Q_t}{Q_0} \right)^{\psi_{Q,k}}. \quad (15)$$

with $\bar{\Psi}_k > 0$ a scaling parameter, and $\psi_{\mathcal{R},k} < 0$ and $\psi_{Q,k} < 0$ defining the convexity of the extraction cost. The technology of the energy provider being set, we can formally describe its behavior through the usual optimization problem.

Problem 3 (FI – Final input provider) *The final input provider $k \in \mathcal{K}$ is perfectly competitive and uses the technologies defined in Eq. (12) and (15). The representative firm takes prices of machines, human capital, and final inputs ($\{p_{k,i,t}^x\}_{i \in [0,1]}$, w_t , and $\{p_{k,t}\}_{k \in \mathcal{K}}$ respectively), the current level of technologies, $A_{k,t}$, $\{q_{k,i,t}\}_{i \in [0,1]}$ and Q_t , as given, as well as the current level of resource, $\mathcal{R}_{k,t}$, to solve*

$$\begin{aligned} \max_{\{x_{k,i,t}\}_{i \in [0,1]}, H_{k,t}, E_{k,t}} \quad & p_{k,t} A_{k,t} \left[\int_0^1 (q_{k,i,t} x_{k,i,t})^{\alpha_k} di \right] H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k} \\ & - \int_0^1 p_{k,i,t}^x x_{k,i,t} di - w_t H_{k,t} - \Psi_k E_{k,t} \end{aligned}$$

thus, in each sector $k \in \mathcal{K}$, setting demand schedules in machines, $x_{k,i,t}(\hat{v})$, human capital, $H_{k,t}(\hat{v})$ and primary energy, $E_{k,t}(\hat{v})$, where “ \hat{v} ” stands for “ $\{p_{k,t}\}_{k \in \mathcal{K}}$, $\{p_{k,i,t}^x\}_{i \in [0,1]}$, w_t , $\{\mathcal{R}_{k,t}\}_{k \in \mathcal{K}}$, or Q_t ”.

The first order conditions corresponding to this program yield the optimal demand schedules for each factor as follows

$$\frac{p_{k,i,t}^x}{p_{k,t}} = \alpha_k A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k} \left[\frac{q_{k,i,t}}{x_{k,i,t}} \right]^{1-\alpha_k}, \quad (16)$$

$$\frac{w_t}{p_{k,t}} = \beta_k \frac{Y_{k,t}}{H_{k,t}}, \quad (17)$$

$$\frac{\Psi_k}{p_{k,t}} = \gamma_k \frac{Y_{k,t}}{E_{k,t}}. \quad (18)$$

As shown in Appendix B, one can compute the aggregate raw physical capital demand, $K_{k,t} = \int_0^1 x_{k,i,t} di$, and thus the aggregate production technology for each final input k is

$$Y_{k,t} = A_{k,t} [\tilde{Q}_{k,t} Q_{k,t}]^{1-\alpha_k} K_{k,t}^{\alpha_k} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}. \quad (19)$$

$Q_{k,t+1} = \int_0^1 q_{k,i,t} di$ stands for an average quality index in sector k after innovation occurred at the beginning of period t . This quality index captures the technological level of sector k achieved through R&D, and will thus shape innovation and production decisions in period $t + 1$. $\tilde{Q}_{k,t}^{1-\alpha_k} = \left[\alpha_k^{\frac{\alpha_k}{1-\alpha_k}} q_k \lambda_{k,t} + (1 - \lambda_{k,t}) \right] / \left[\alpha_k^{\frac{1}{1-\alpha_k}} q_k \lambda_{k,t} + (1 - \lambda_{k,t}) \right]^{\alpha_k}$ is a scaling parameter for the quality index that capture the hybrid nature of competition in the production intermediate capital goods.

¹⁹For instance, steam engines were first used to pump out water from flooded coal mines. Despite their poor efficiencies, such a use of steam engines decreased extraction costs because older methods using mechanical power from horses were more inefficient and expensive. Moreover, the depth of accessible coal seams increased as a result of this technical breakthrough. Similarly, one can think about the appearance of water wheels (resp. fracking) as a new technology that implicitly decreased the capture and conversion costs of previously unused water flows (respec. shale oil deposits) so as to make them economically profitable.

3.2.3 Capital goods production

Machines used to extract and use energy to produce the final inputs are seen as intermediate capital goods supplied from the stock of available raw capital by intermediate producers. The production regime is contingent on the machine-specific R&D process, which is described in more detail in [Subsection 3.3.3](#) and occurs *prior* to production at the beginning of each period. In case of R&D success, the innovator is endowed with a one-period patent on her machine line. Otherwise, the machine line is produced competitively under the previous design, a situation one shall interpret as patents becoming public. As is customary in the Schumpeterian literature (see [Acemoglu, 2009](#)), R&D is assumed to be performed only by potential entrants that have higher incentives due to the potential monopoly profit they might benefit. This setup, along with the free-entry condition into R&D introduced in Eq. (31), ensures that there are no aggregate monopoly profits.²⁰

For each machine line i in the sector k , the production technology is linear and transforms one unit of raw capital stock, K_t —which is rented from households at the interest rate r_t plus the depreciation rate of capital δ —into one unit of specialized machine $x_{k,i,t}$. Hence, the corresponding operating profit is $\pi_{k,i,t} = (p_{k,i,t}^x - r_t - \delta)x_{k,i,t}$.²¹ This behavior can be formalized through an usual optimization problem.

Problem 4 (CG – Capital good producers) *Each capital good sector $k \in \mathcal{K}$ sustains a hybrid regime of perfect and monopolistic competition, depending on the success of profit-motivated R&D. Whenever the competition is perfect in machine line i (unsuccessful innovation), the price equals the marginal cost of production, that is $p_{k,i,t}^x = r_t + \delta$ (also denoted $p_{k,i,t}^c$). Whenever the competition is monopolistic for the machine line i (successful innovation), the capital good producer takes the specific demand schedule, $x_{k,i,t}$, as given by Eq. (16) as well as the price of capital, r_t , to solve*

$$\max_{p_{k,i,t}^x} \pi_{k,i,t} = (p_{k,i,t}^x - r_t - \delta)x_{k,i,t}(p_{k,i,t}^x)$$

thus setting the price of intermediate capital goods, $p_{k,i,t} = (r_t + \delta) / \alpha_k$ (also denoted $p_{k,i,t}^m$).²² The supply of intermediate capital goods is then given accordingly.

It is worth noting that the price in each regime of competition is common to all segmented machine lines and sectors (no dependency on the specific machine line). This price is thus denominated $p_{k,t}^x$, or respectively $p_{k,t}^c$ and $p_{k,t}^m$ whenever relevant.

3.3 Knowledge and endogenous technical change

Following [Strulik et al. \(2013\)](#), we have considered two kinds of technical changes: non-profit motivated learning-by-doing occurring in the final good sector, and profit-motivated R&D affecting both final and energy-producing sectors. As suggested by [Schaefer et al. \(2014\)](#), all these technical advancements should be interrelated by the evolution of a *General Purpose Technology* (GPT) referred to as G . [Lipsey et al. \(2005, p. 98\)](#) define a GPT as a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects.²³

²⁰These assumptions prevent any issue of inter-temporal patent allocation and pricing, without precluding the set of incentives central to profit-motivated R&D ([Acemoglu, 2002](#); [Aghion & Howitt, 1992, 1998](#)).

²¹To derive this result we assume that, as a clearing condition on the capital market, the rental rate of capital shall equalize its rate of net return.

²²To derive this result, we make a customary assumption of the patent-race literature. Hence, we suppose that the monopoly price can be fully charged without the possibility of the monopolist holding a patent on a previous version of the technology vintage to reap out the technology market. By doing so, we assume technical improvements to be sufficiently large such that a drastic innovation regime occurs (see [Aghion & Howitt, 1992](#), for an analysis of the non-drastring case that yields similar comparative static results when the production function is of Cobb-Douglas type). An alternative view shall be to assume that patent holders on a specific machine line engage in a competing setting *a la Bertrand*, such that the most recent innovator (holding the highest quality) implements a limit price, ensuring that they remain below the marginal cost of the next quality vintage holder, such that they remain alone on the market. The key issue here is that each patent holder realizes a positive profit, yielding incentives to enter profit-motivated R&D activities.

²³[Lipsey et al. \(2005, p. 97\)](#) further stress that GPTs are typically use-radical but not technology-radical, meaning that GPTs do not stand out from other technologies because of a revolutionary technical basis, but rather because of outstanding

3.3.1 Endogenous dynamics of GPTs

Following Schaefer *et al.* (2014), we assume that successive vintages of GPTs, $G_{v,t}$, are developed endogenously as a result of non-profit motivated activities. In the following setting, we suppose that the level of the current GPT vintage positively affects technical change, i.e. the *rate of growth* of technical levels achieved through learning-by-doing or R&D. In a sense, GPTs gather all kinds of technical externalities fostering technical change. We assume that several vintages of GPTs, indexed by v , succeed over time. While it is still active, the *level* of a given vintage of GPT might also evolve over time, featuring learning-effects. The magnitude of the latter phenomenon, along with the evolution of the expected duration before the arrival of a new GPT, has a crucial impact on the degree of complementarity between past and current knowledge. Together, those features allow compliance with two stylized facts regarding the historical arrival of GPTs: (i) the initially slow evolution of the efficiency of new GPTs, and (ii) the decreasing time interval between successive GPTs. We differ from Schaefer *et al.* (2014) in two ways. First, following the endogenous growth literature centered on human capital (Jones, 1995; Strulik *et al.*, 2013), we consider researchers (i.e., human capital allocated to R&D) rather than machines (i.e., lab-equipment purchased through financial expenditures) to be the key driver of innovation processes. Moreover, we consider that all kinds of technical changes are involved in the evolution of GPTs. Finally, we assume that GPT vintages active for period t evolve *prior* to innovation and production decisions; that is, at the beginning of the current period.

To start with, let's assume that new GPT vintages follow a Poisson process with endogenous mean, μ_t , defined as

$$\mu_t = \mu_0 G_{v,t-1}, \quad (20)$$

with $\mu_0 \in [0, 1]$. The level of the active GPT, $G_{v,t-1}$ eases the arrival of the new vintage.²⁴ Once discovered, a new GPT is initialized with a level of

$$G_{v+1,t} = \bar{G} \tilde{q}_t \quad (21)$$

where \bar{G} is a positive scaling parameter, and \tilde{q}_t is an index of applied knowledge available for the GPT in period t . To further characterize the index of available applied knowledge, \tilde{q}_t , let us define two stocks of knowledge, $\tilde{Q}_{v+1,t}$ and $\tilde{Q}_{v,t}$. The former represents the improvement history of the current (and potentially newly introduced) GPT, whereas the latter tracks the improvement history of all previous vintages of GPTs. Introducing the aggregate quality index of the economy

$$Q_t = \sum_{X \in (A, Q), \mu \in (Y, \mathcal{K})} X_{u,t}, \quad (22)$$

that measures the extent of all technical developments through learning-by-doing and R&D within the period t , one can write the following identity

$$\tilde{Q}_{v+1,t} + \tilde{Q}_{v,t} = Q_t. \quad (23)$$

It is worth noting that the total quality index of the economy, Q_t , should be interpreted as the overall stock of *applied* knowledge already introduced in the efficiency of the schooling technology, Eq. (3), and the extraction technology of primary energy inputs, Eq. (15). Following Mokyr (1990, 2011), *applied* knowledge, taking the form of learning-by-doing and R&D technologies, shall be distinguished from *useful* knowledge contained in human capital and in GPT's waves that evolve concomitantly with the development of applied knowledge. The index of applied knowledge then

applications and adaptations to other technologies and sectors of the economy. GPTs are typically not born in their final form, so they often start off as something we would never call a GPT and then develop into something that transforms an entire economy. The considerable scope of improvement of GPTs is explored as their range and variety of use increase, which in the meantime generate knowledge and practical spillovers on other technologies and organizational processes.

²⁴In other words, the time interval between successive GPT vintages, T , is given by the cumulative distribution: $\mathbb{P}(T \leq t)$ where T follows a Poisson process of mean μ , hence the average waiting time corresponds to $\mathbb{E}(T) = 1/\mu$.

only capture the quantity of applied knowledge that can be used to strengthen the current GPT vintage. Depending of the complementarity between past and current applied knowledge, quantified by the parameter $\zeta \in [0, 1]$,²⁵ one can then define this index as

$$\tilde{q}_t = \tilde{Q}_{v+1,t}^\zeta \tilde{Q}_{v,t}^{1-\zeta}. \quad (24)$$

As long as it remains active, the quality increments of GPT vintage evolves over time, that is as a function of the quality index \tilde{q}_t ,

$$G_{v,t} = G_{v,t-1} + \zeta \tilde{q}_t, \quad (25)$$

with ζ a constant controlling the speed of *diffusion* of the current GPT vintage within the economy through the learning-by-doing and R&D performed with this vintage. Thus, the efficiency of each kind of technical knowledge is improved (as specified thereafter), again strengthening the level of the current GPT according to Eq. (23). Moreover, as GPTs are improved, the time interval before the arrival of a new GPT decreases according to Eq. (20).

3.3.2 Technical change through learning-by-doing

We model the technical level achieved through learning-by-doing in input sector $k \in \mathcal{K}$, $A_{k,t}$, as a function of (i) the current human capital stock allocated to the specific production sector, $H_{k,t}$, and (ii) the current GPT's level, $G_{v,t}$, capturing the conventional technical externality (i.e., the so-called standing-on-giants-shoulders effect).²⁶ With $\Omega > 0$ representing the efficiency with which useful knowledge contained in human capital and GPT-related know-how are converted into applied learning-by-doing knowledge for production, we have

$$A_{k,t+1} - A_{k,t} = \Omega H_{k,t}^{\omega_{H,k}} G_{v,t}^{\omega_{G,k}}, \quad k \in \mathcal{K}. \quad (26)$$

We suppose that there are decreasing returns in both human capital and technical externalities, i.e., $\omega_{H,k} \in]0, 1[$ and $\omega_{G,k} \in]0, 1[$, meaning that in the long run there is no more technical change through leaning-by-doing. This assumption calls for another source of technical change, namely, profit-motivated R&D presented below, to sustain growth in the long-run. Finally, we define the growth rate of the technical level (i.e., the technical change) obtained through learning-by-doing as $g_{A_k,t} \equiv \frac{A_{k,t} - A_{k,t-1}}{A_{k,t-1}} = \Omega H_{k,t-1}^{\omega_{H,k}} G_{v,t-1}^{\omega_{G,k}} A_{k,t-1}^{-1}$, with $k \in \mathcal{K}$.

3.3.3 R&D-based technical change

As conventionally assumed in the endogenous growth literature (Acemoglu, 2002; Aghion & Howitt, 1992, 1998), R&D is profit-motivated. As described in Subsection 3.2.3, in each sector, intermediate capital goods (i.e., machines) are produced either in perfect competition whenever the corresponding patent is public, i.e. R&D was unsuccessful, or by monopolists that are former successful innovators. We assume that each machine line follows a specific quality ladder: the quality of the machine line in a specific sector writes $q_{k,i,t} = q_k^{\kappa_{k,i,t}}$, where $\kappa_{k,i,t}$ is the number of successful innovations for machine line i in sector k up to time t and $q_k > 1$ the sector-specific rung of the corresponding quality ladder. At the beginning of each period, successful innovations bring the corresponding machines to a higher rung of the specific quality ladder, that is $\kappa_{k,i,t}$ becomes $\kappa_{k,i,t} + 1$. Otherwise, the quality of machines remains constant. We follow Schaefer *et al.* (2014) and assume that the probability of success of a potential innovator in a specific sector k and machine line i writes

$$\lambda_{k,i,t} = \Phi_{k,i,t} H_{R,k,i,t} G_{v,t}, \quad (27)$$

where (i) $\Phi_{k,i,t}$ captures the increasing difficulty to perform R&D with the current complexity of the corresponding production line $\kappa_{k,i,t}$, (ii) $H_{R,k,i,t}$ stands for the amount of human capital dedicated

²⁵Past (respectively current) knowledge is useless whenever $\zeta = 1$ ($\zeta = 0$).

²⁶Given that our formulation of $G_{v,t}$ depends on technical levels $A_{k,t}$ achieved through learning-by-doing, Eq. (26) is strictly in line with the formulation of Jones (1995).

to research in the machine line i of sector k , and (iii) $G_{v,t}$ is the level of the current GPT vintage. It is worth mentioning that such a modeling choice of the probability of success does not preclude *per se* an upper bound for productivity gains in each sector—in others words this probability is unlikely to reach its zero lower bound. As suggested by Ayres & Warr (2009, pp. 52–53), thermodynamics constraints of real processes necessarily imply an impossibility for infinite technical change. Given that the present paper focuses on economic take-off and the energy transition associated with the Industrial Revolution, we consider that the introduction of such thermodynamic limits is beyond its scope. However, as advocated by Kortum (1993) and Stokey (1995), we explicitly introduce decreasing returns for the cost of R&D through the following functional form

$$\Phi_{k,i,t} = \frac{\bar{\lambda}_k - \lambda_{k,i,t}}{\phi_k} \frac{1}{q_k^{\kappa_{k,i,t}+1}}, \quad (28)$$

where ϕ_k is a parameter capturing the cost of innovation in sector k , and $\bar{\lambda}_k \in (0,1)$ stands for the ultimate level of probability success in R&D. One can thus write the probability of an innovation success as

$$\lambda_{k,i,t} = \bar{\lambda}_k \frac{H_{R,k,i,t} G_{v,t} q_k^{-[\kappa_{k,i,t}+1]}}{\phi_k + H_{R,k,i,t} G_{v,t} q_k^{-[\kappa_{k,i,t}+1]}}. \quad (29)$$

It is clear from Eq. (29) that the probability of successful R&D converges to its limit value, $\bar{\lambda}_k$, whenever the stock of human capital and the level of the GPT grow. This expression allows us to formulate the following proposition.

Proposition 2 (R&D effort) *The level of human capital requirements for R&D decreases with the level of the current GPT, $G_{v,t}$, and increases with the fixed cost, ϕ_k , the level of complexity of the corresponding machine line, $\kappa_{k,i,t}$, as well as the level of innovation success, $\lambda_{k,i,t}$. Moreover, at a certain point, increasing the R&D success probability becomes infinitely costly.*

Proof. The proof of this proposition is immediate after isolating $H_{R,k,i,t}$ from Eq. (29) as

$$H_{R,k,i,t} = \phi_k \times q_k^{\kappa_{k,i,t}+1} \times \frac{1}{G_{v,t}} \times \frac{\lambda_{k,i,t}}{\bar{\lambda}_k - \lambda_{k,i,t}}.$$

□

Innovation decisions (and ultimately the free-entry conditions defined shortly) are driven by the one-period success monopoly profit arising from the improvement of a sector-specific machine line,

$$\pi_{k,i,t}^S = \bar{\pi}_{k,t} \left[\frac{1}{P_{k,t}^m} \right]^{\frac{\alpha_k}{1-\alpha_k}} q_k^{\kappa_{k,i,t}+1}, \quad (30)$$

where $\bar{\pi}_{k,t} = (1 - \alpha_k) \left[p_{k,t} \alpha_k A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k} \right]^{\frac{1}{1-\alpha_k}}$. R&D processes and monopoly profits being defined, we can then describe the behavior of innovator through a free-entry condition.

Problem 5 (R&D – Sectoral innovators) *Each R&D sector $k \in \mathcal{K}$ is viewed as a pool of innovators, willing to enter the capital good production market through a successful innovation. Each potential monopolist takes as given the current level of GPT, $G(v,t)$, the complexity level of the targeted production line, $\kappa_{k,i,t}$, as well as the price of human capital, w_t , to maximize their expected profit, $\lambda_{k,i,t} \pi_{k,i,t} q_k$, such that at the equilibrium the following free entry condition in R&D holds*

$$\lambda_{k,i,t} \pi_{k,i,t}^S = w_t H_{R,k,i,t} \quad (31)$$

This condition sets the level of the probability of successful R&D, which ultimately determines the optimal demand for human capital in each research sector, $H_{R,k,i,t}(w_t, \pi_{k,i,t})$.

In writing this problem, we assume funds to be ultimately lent by households to potential innovators, and then repaid through profits (i.e. dividends) whenever innovation is successful. We can now state the following proposition.

Proposition 3 (Sector-specific R&D success probability) *The R&D success probability only depends on the sector and not on the machine lines that make it up, that is*

$$\lambda_{k,t} = \bar{\lambda}_k - \frac{\phi_k w_t [p_{k,t}^m]^{1-\alpha_k}}{G_{v,t} \bar{\pi}_{k,t}}. \quad (32)$$

This expression is increasing in the level of the current GPT, $G(v,t)$, as well as an indicator of the value of innovation, $\bar{\pi}_{k,t}$, and decreasing in the sectoral research cost, ϕ_k , in the wage level, w_t , as well as in the cost of producing machines, $p_{k,t}^m$.

Proof. The proof is obtained by substituting Eq. (30) and (29) into Eq. (31), which yields an explicit expression for the probability of successful R&D that is independent from the machine line index, i (that can be dropped).²⁷ \square

As a result, human capital allocations, $H_{R,k,i,t}$, are uniformly distributed among all machines lines of a sector and do not depend on the level in the quality ladder, $\kappa_{k,i,t}$. In each sector, one can thus use Eqs. (27) and (28) to derive the aggregate amount of human capital dedicated to research as

$$H_{R,t} = \sum_{k \in \mathcal{K}} \frac{\phi_k}{G_{v,t}} \frac{\lambda_{k,t}}{\bar{\lambda}_k - \lambda_{k,t}} q_k Q_k. \quad (33)$$

In addition, the law of motion of the quality index, $Q_{k,t}$, can be computed in each sector by using the law of large numbers, ensuring that the probability of innovation success, $\lambda_{k,t}$, coincides with the fraction of machine-lines that will experience a success in R&D. This leads to the following quality dynamics

$$Q_{k,t+1} = \lambda_{k,t} q_k Q_{k,t} + (1 - \lambda_{k,t}) Q_{k,t}, \quad (34)$$

and thus the growth rate of innovation in each sector, $g_{Q_{k,t}}$, is

$$g_{Q_{k,t}} = [q_k - 1] \lambda_{k,t}. \quad (35)$$

Finally, we can compute the aggregate technical growth rate of the economy, g_t , as a weighted average of the different technical growth rate obtained, either through learning-by-doing, $\{g_{A_{k,t}}\}_{k \in \{\mathcal{K}\}}$, or through sector-specific R&D, $\{g_{Q_{k,t}}\}_{k \in \{\mathcal{K}\}}$. Formally, we have

$$g_t = \sum_{X \in (A,Q), k \in \mathcal{K}} g_{X_{k,t}} \frac{X_{k,t}}{Q_t} \quad (36)$$

where Q_t has already been defined in Eq. (22).

3.4 Market-clearing and general equilibrium solution

At each time period, real flows must ensure that all markets —namely (i) final good, (ii) physical capital, (iii) human capital, and (iv) financial assets— clear, that is

$$Y_t = c_t N_t + (1 + r_t) s_{t-1} N_{t-1} + K_{t+1} - (1 - \delta) K_t + \sum_{k \in \mathcal{K}} \Psi_k E_{k,t} + b_t e_t N_t, \quad (37)$$

$$H_t = \sum_{k \in \mathcal{K}} H_{k,t} + H_{R,t}, \quad (38)$$

$$K_t = \sum_{k \in \mathcal{K}} K_{k,t}, \quad (39)$$

$$K_{t+1} = s_t N_t. \quad (40)$$

²⁷This property is due to the assumption that rungs of the quality ladder are proportionally distant.

These conditions ensure that the provision of real flows equals their uses. It is worth mentioning that the constant returns to scale assumption for the production technologies, combined with the set of hypotheses conditioning the behavior of the household, ensure that Eq. (37) always holds. Moreover, physical constraints shall hold in the provision of energy inputs, as stated in Eq. (14) and (13), and in the provision of capital goods. Hence, we have

$$\begin{aligned} R_{k,t} &= R_{k,0} - E_{k,t-1} \quad (\text{renewable resource}), \\ R_{k,t} &= R_{k,0} - \sum_{a<t} E_{k,a} \quad (\text{exhaustible resource}), \\ K_{k,t} &= \int_0^1 x_{k,i,t} di, \forall k \in \mathcal{K}. \end{aligned} \quad (41)$$

We now turn to the explicit definition of a decentralized dynamic general equilibrium in our theoretical framework.

Problem 6 (GE – General equilibrium) *An equilibrium is a sequence of: per capita consumption, $\{c_t\}$, savings, $\{s_t\}$, fertility $\{b_t\}$, educational investment $\{e_t\}$, physical capital allocations, $\{\{K_{k,t}\}_{k \in \mathcal{K}}\}$, human capital allocation, $\{\{H_{k,t}\}_{k \in \mathcal{K}}\}$, final input provision and primary energies' extractions flows, $\{\{Y_{k,t}, E_{k,t}\}_{k \in \mathcal{K}}\}$, as well as prices $\{r_t, w_t, \{p_{k,t}\}_{k \in \mathcal{K}}\}, \{p_{k,t}^x\}_{k \in \mathcal{K}}\}$, such that*

- (i) $\{c_t, s_t, b_t, e_t\}$ solve **Problem HH**;
- (ii) $\{\{p_{k,t}\}_{k \in \mathcal{K}}\}$ solve **Problem FG**;
- (iii) $\{\{x_{k,i,t}\}_{k \in \mathcal{K}, i \in [0,1]}, \{H_{k,t}\}_{k \in \mathcal{K}}, \{E_{k,t}\}_{k \in \mathcal{K}}\}$ solve **Problem FI** under primary energy resource constraints defined by Eq. (13) and (14);
- (iv) $\{\{p_{k,t}^x\}_{k \in \mathcal{K}, i \in [0,1]}\}$ solve **Problem CG** along with the free-entry condition of **Problem R&D**;
- (v) $\{r_t\}$ is such that the physical capital market clears, that is, Eq. (40) holds;
- (vi) $\{w_t\}$ is such that the human capital market clears, that is, Eq. (38) holds;
- (vii) $\{\{x_{k,i,t}\}_{k \in \mathcal{K}, i \in [0,1]}\}$ are such that the capital resource constraint is satisfied, that is, Eq. (41) holds;
- (viii) Physical capital follows the accumulation dynamics described in Eq. (39);
- (ix) Population and human capital follow the endogenous dynamics described in Eq. (8) and (10);
- (x) GPTs are generated from a Poisson process of endogenous mean given by Eq. (20), and evolve according to Eq. (23) and (25);
- (xi) Learning-by-doing technical changes endogenously evolve according to Eq. (26);
- (xii) R&D-based technical changes endogenously evolve according to Eq. (29) and (31).

4 Analytical results

In this section, we derive some theoretical properties from our model, focusing on a two-sector case that we will also consider for the numerical exercises of Section 5. The two sectors are respectively labeled *renewable* (indexed by r) and *exhaustible* (indexed by e), that is $\mathcal{K} = \{r, e\}$. In order to derive tractable analytical results, we limit our model by formulating several assumptions.

First, we consider the particular case of symmetric technologies in final inputs production. This assumption ensures that the endogenous directed technical change only results from the endogenous changes in (i) energy input stocks, and (ii) their associated extraction technologies, and not from *ex-ante* biases with regard to production technologies.

Assumption 1 (Symmetry in production) *Production technologies of final inputs are assumed to be symmetric, that is $\forall k \in \mathcal{K}, (\alpha_k, \beta_k, \gamma_k) = (\alpha, \beta, \gamma)$ and $\forall k \in \mathcal{K}, \vartheta_k = \vartheta$.*

Second, to simplify the analytical resolution without altering the driving forces of our framework, we also assume for this section that extraction costs depend on *past* and not *current* resource availability.

Assumption 2 *Extraction costs depend on past resource flows, that is $\forall k \in \mathcal{K}, \Psi_{k,t} = \Psi(\mathcal{R}_{k,t}, Q_t)$.*

Lastly, we assume final inputs to be gross substitutes to cope with historical facts and obvious physical properties of energy carriers.

Assumption 3 *Final inputs are gross substitutes, that is $\sigma > 1$.*

We may now state some results regarding equilibrium relative prices and quantities before turning to the endogenous direction of technical change that is the key transition mechanism discussed in this paper.

4.1 Relative equilibrium prices and quantities

From the FOC of **Problem 2**, we retrieve the usual substitution effect between final inputs, that is, an inverse relationship between their equilibrium relative price and quantities.

$$\frac{p_{r,t}}{p_{e,t}} = \left[\frac{Y_{r,t}}{Y_{e,t}} \right]^{-\frac{1}{\sigma}}. \quad (42)$$

Combining the first order conditions with regards to final inputs and their corresponding production technologies allows the equilibrium relative price of final inputs to be expressed as a function of technical levels and extraction costs given at the beginning of the period

$$\frac{p_{r,t}}{p_{e,t}} = \left[\frac{A_{r,t}}{A_{e,t}} \right]^{-1} \left[\frac{\bar{Q}_{r,t} Q_{r,t}}{\bar{Q}_{e,t} Q_{e,t}} \right]^{-(1-\alpha)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{\gamma}, \quad (43)$$

where $\bar{Q}_{k,t} = 1 + \lambda_{k,t}(\alpha^{\frac{1}{1-\alpha}} q_k - 1)$, a scaling factor for the corresponding sector-specific R&D knowledge. We show in this section some conditions under which the R&D success probability, $\lambda_{k,t}$, is higher in the most technically advanced sector. This expression allows us to state a first equilibrium result.

Proposition 4 (Relative input price) *The higher (resp. the lower) the relative level of learning and R&D knowledge (resp. the relative cost of extraction), the lower the relative price of final inputs.*

Proof. Combining Eqs (16) and (18) allows the resource flow to be isolated as

$$E_{k,t} = \left[\frac{\alpha}{r_t + \delta} \right]^{\frac{\alpha}{\beta}} \left[\frac{\gamma \bar{Q}_{k,t} Q_{k,t}}{\Psi_{k,t}} \right]^{\frac{1-\alpha}{\beta}} [A_{k,t} p_{k,t}]^{\frac{1}{\beta}} H_{k,t}.$$

Substituting in Eq. (12) this expression and the equilibrium demand of the capital good, that is Eq. (16), yields the following expression for the equilibrium final input production

$$Y_{k,t} = \left[\frac{\alpha}{r_t + \delta} \right]^{\frac{\alpha}{\beta}} [\bar{Q}_{k,t} Q_{k,t}]^{\frac{1-\alpha}{\beta}} \left[\frac{\gamma}{\Psi_{k,t}} \right]^{\frac{\gamma}{\beta}} A_{k,t}^{\frac{1}{\beta}} p_{k,t}^{\frac{1-\beta}{\beta}} H_{k,t}.$$

Equalizing the marginal product of labor across sectors, that is Eq. (17) taken for each sector, and using the previous relations leads to Eq. (43) from which the proposition directly results. \square

Additionally, taking the ratios of the previous expressions and substituting the equilibrium relative price of final inputs allows a second equilibrium result to be stated.

Proposition 5 (Relative factor use) *Under Assumption 1, 2 and 3, the higher (resp. the lower) the relative level of learning and R&D knowledge (resp. the relative cost of extraction), the higher the relative allocation of labor, physical capital, and energy resources.*

Proof. The combination of Eqs. (42) and (43) with equilibrium final input production allows the isolation of the ratio of relative employment in intermediary sectors as

$$\frac{H_{r,t}}{H_{e,t}} = \left[\frac{A_{r,t}}{A_{e,t}} \right]^{\sigma-1} \left[\frac{\bar{Q}_{r,t} Q_{r,t}}{\bar{Q}_{e,t} Q_{e,t}} \right]^{(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)}.$$

A direct assessment of this expression then yields the discussed property. A similar analysis can be performed for the energy resource, substituting the previous expression within the ratio of equilibrium energy resource use. Lastly, computing sector-specific capital demands according to Eq. (41) by substituting the equilibrium quantity of provided machines, that is Eq. (16), gives

$$K_{k,t} = \left[\frac{\alpha}{r_t + \delta} \right]^{\frac{\alpha+\beta}{\beta}} \bar{Q}_{k,t}^{\frac{\gamma}{\beta}} \hat{Q}_{k,t} \bar{Q}_{k,t}^{\frac{1-\alpha}{\beta}} \left[\frac{\gamma}{\Psi_{k,t}} \right]^{\frac{\gamma}{\beta}} [A_{k,t} p_{k,t}]^{\frac{1}{\beta}} H_{k,t},$$

with $\hat{Q}_{k,t} = 1 + \lambda_{k,t}(\alpha^{\frac{1}{1-\alpha}} q_k - 1)$. This expression can be used to infer the relative physical capital allocation. \square

4.2 Endogenous direction of technical change

We focus now on the endogenous direction of technical change, which is the key transition mechanism within our framework. It is obvious that whenever only one of the two sectors experiences R&D, a case that will be discussed within our numerical analysis, technical change is biased against the other sector. Here, we restrict to a situation where both sectors experience R&D to assess the dynamics of the underlying forces that shape technical biases. The endogenous direction of technical change can be assessed using the relative equilibrium growth rates of R&D knowledge that, according to Eq. (35), writes

$$\frac{g_{Q_{r,t}}}{g_{Q_{e,t}}} = \frac{q_r - 1}{q_e - 1} \frac{\lambda_{r,t}}{\lambda_{e,t}}. \quad (44)$$

Sector-specific quality improvements, q_k , and research success probabilities, λ_k , can be respectively interpreted as the sector-specific *magnitude* and the *scale* (i.e. mass of successful innovators) of R&D. Compared to the quite complex parametric form for the R&D success probability introduced in Eq. (29), in this section we restrict to a simpler formulation to derive intuitive and tractable results. We show in Appendix C how to retrieve these results with our original specification in a simplified case, that is assuming R&D is specified in terms of final good consumption (lab-equipment approach). The simplified functional form discussed in this section is

$$\lambda_{k,i,t} = \bar{\lambda} G_{v,t} \left[\frac{H_{R,k,i,t}}{q_k^{\kappa_{k,i,t}+1}} \right]^{\epsilon}, \quad (45)$$

where $\bar{\lambda} \in (0,1)$ and $\epsilon \in (0,1)$ are parameters capturing the efficiency and the convexity of R&D efforts. Note that we assume these parameters to be identical for both sectors in the spirit of Assumption 1. The functional form of Eq. (45) is similar to the one introduced in Eq. (29) in the sense that the R&D success probability (i) increases with R&D efforts, $H_{R,k,i,t}$, (ii) exhibits decreasing marginal returns in R&D efforts, and (iii) decreases with the complexity of the machine line, $q_k^{\kappa_{k,i,t}}$.

Moreover, combined with the free-entry condition in R&D in Eq. (31), the R&D success probability considered here also appears to be independent from the specific machine line as

$$\lambda_{k,t} = \bar{\lambda}^{\frac{1}{1-\epsilon}} \left[\frac{G_{v,t} \bar{\pi}_{k,t}}{w_t [p_{k,t}^m]^{\frac{\alpha}{1-\alpha}}} \right]^{\frac{\epsilon}{1-\epsilon}}. \quad (46)$$

The comparative statics of the equilibrium R&D success probability is also similar for both parametric forms: the higher the sector-specific profit perspectives, $\bar{\pi}_{k,t}$, and the level of the current GPT, $G_{v,t}$, (resp. the lower the production cost of machines, $p_{k,t}^m$, and the cost of R&D, w_t), the higher the R&D efforts and consequently the higher the sector-specific R&D success probability.²⁸

Substituting the previous equilibrium relations into innovator profits defined by Eq. (30) yields

$$\frac{\mathcal{G}_{Q,r,t}}{\mathcal{G}_{Q,e,t}} = \underbrace{\frac{q_r - 1}{q_e - 1}}_{\text{R\&D efficiency}} \times \underbrace{\left[\frac{p_{r,t}}{p_{e,t}} \right]^{\frac{\epsilon}{(1-\alpha)(1-\epsilon)}}}_{\text{price effect}} \times \underbrace{\left[\frac{H_{r,t}^\beta E_{r,t}^\gamma}{H_{e,t}^\beta E_{e,t}^\gamma} \right]^{\frac{\epsilon}{(\beta+\gamma)(1-\epsilon)}}}_{\text{market-size effect}} \times \underbrace{\left[\frac{A_{r,t}}{A_{e,t}} \right]^{\frac{\epsilon}{(1-\alpha)(1-\epsilon)}}}_{\text{productivity effect}}. \quad (47)$$

In a generalization of [Acemoglu et al. \(2012\)](#), the relative sectoral intensity in R&D then appears to be driven by: (i) an *R&D efficiency effect*, (ii) a *price effect* favoring innovation toward the sector with the higher price, that is the least advanced sector and/or the sector where the energy resource is the most scarce according to [Proposition 4](#), (iii) a *market-size effect* favoring innovation in the larger sector in terms of employment and energy use (i.e., the larger potential market for complementary machines), that is, the most advanced sector and/or the sector where the resource is the most abundant according to [Proposition 5](#), and (iv) a *direct productivity effect* favoring innovation in the more advanced sector in terms of learning-by-doing. It is worth mentioning that there is now a direct effect from the sector specific level of knowledge, $Q_{k,t}$, which is due to the specific parametric form of Eq. (45) where the normalization of R&D efforts by $q_k^{\kappa_{k,i,t}+1}$ exactly compensates for the effect from the machine-specific quality level in $\bar{\pi}_{k,t}$. However, this parametric form is usual in the Shumpeterian growth literature (e.g., see [Aghion & Howitt, 2009](#)) and allows the derivation of tractable analytical results by ensuring that the R&D success probability is independent from the machine line. However, as already seen in [Proposition 4](#) and [5](#), relative R&D knowledge shapes equilibrium relative final input prices and factor use, indirectly featuring a *build-on-the-shoulders-of-giants* effect in R&D. To demonstrate this result, we consider an additional technical assumption regarding the competition regime for capital good provision.

Assumption 4 (CG production) *Capital goods are produced within a monopolistic competition regime.*

This assumption simplifies the analysis by ensuring that the scaling factor for the sector-specific R&D knowledge is constant, that is $\bar{Q}_{k,t} = \alpha^{\frac{\alpha}{1-\alpha}}$ and $\hat{Q}_{k,t} = \alpha^{\frac{1}{1-\alpha}}$. However, it does not alter the fundamental drivers of innovation. Indeed, the hybrid competition regime presented in [Subsection 3.3.3](#) where only machines' lines that experience successful R&D are supplied under monopolistic competition while others are supplied competitively, was introduced to ensure zero aggregate profits in order to not interfere with the fertility choices of the representative household. Having an upper limit for the scaling factor $\bar{Q}_{k,t}$, $\hat{Q}_{k,t}$ and $\check{Q}_{k,t}$ which is also bounded, only magnifies the impact of the pre-existing R&D knowledge stocks $Q_{k,t}$.

Substituting the ratios of equilibrium relative prices and quantities within Eq. (47) allows the full characterization of the endogenous direction of technical change by

²⁸It can also be shown that the calibration of the two parametric forms, proposed respectively in Eq. (29) and (45), result in similar *behaviors*, in the sense of close numerical values for R&D success probability derived from the free-entry condition, i.e., from the underlying innovation incentives.

$$\frac{g_{Q,r,t}}{g_{Q,e,t}} = \frac{q_r - 1}{q_e - 1} \left[\frac{A_{r,t}}{A_{e,t}} \right]^{(\sigma-1)\frac{\epsilon}{1-\epsilon}} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)\frac{\epsilon}{1-\epsilon}} \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{((1-\alpha)(\sigma-1)-1)\frac{\epsilon}{1-\epsilon}}. \quad (48)$$

Theorem 2 (Direction of technical change) *Under Assumption 1, 2, 3 and 4, and assuming that final inputs are sufficiently substitutable (i.e., that $(1 - \alpha)(\sigma - 1) \geq 1$), R&D technical change tends to be biased against the least advanced sector, both in terms of learning-by-doing and R&D knowledge, and/or the sector relying on the energy resource that is the most scarce.*

Proof. A direct interpretation of Eq. (48) yields this result. \square

This theorem illustrates the transition mechanism that is central within this paper. At first, the renewable sector is technically more advanced thanks to learning-by-doing ($A_{r,0} > A_{e,0}$), the corresponding extraction cost is slightly less expensive or similar ($\Psi_{r,0} \leq \Psi_{e,0}$), and R&D (if any) is more intense in this sector ($Q_{r,0} \geq Q_{e,0}$). Thus, the renewable resource is used at a larger scale than the exhaustible one, according to the equilibrium consumption ratio

$$\frac{E_{r,t}}{E_{e,t}} = \left[\frac{A_{r,t}}{A_{e,t}} \right]^{\sigma-1} \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)-1}.$$

However, as the renewable resource flow gets closer its ultimate maximum potential, its corresponding extraction cost, $\Psi_{r,t}$, increases. In the meantime the exhaustible extraction cost, $\Psi_{e,t}$, remains steady as the corresponding stock is initially nearly untapped. Moreover, due to the decreasing returns of learning-by-doing, the stock of knowledge in the renewable sector, $A_{r,t}$, ultimately grows less than in the exhaustible sector, $A_{e,t}$. These conflicting forces ultimately reverse the incentives to perform R&D, that becomes biased toward the exhaustible sector. Up until a certain point, and provided extraction costs do not reverse due to the depletion of the exhaustible resource, the direction of R&D toward the exhaustible sector become self-sustained (in the sense that a decrease in the $Q_{r,t}/Q_{e,t}$ ratio only triggers more R&D towards the exhaustible sector). This mechanism also triggers a reversal in the use of energy input, that is in the $E_{r,t}/E_{e,t}$ ratio, which materializes in an energy transition toward the exhaustible resource. In the Section 5, we illustrate this mechanism through numerical trajectories of a calibrated version of our model. Before that, we derive a closed-form solution and some theoretical results in a tailored-made simplification of our model that holds only at the vicinity of our period of interest, namely the energy shift and economic take-off.

4.3 Closed-form solution in a stylized case

In this subsection we maintain Assumptions 1, 2, 3 and 4, and we formulate two additional simplifying hypotheses. First, we assume an exogenous human capital supply.

Assumption 5 (Exogenous labor supply) *The supply of human capital, H_t , is exogenous.*

This assumption shall not affect the results obtained in what follows because, as demonstrated in Section 3.1, the supply of human capital is ultimately determined by the level of knowledge, Q_t , that is inherited from previous periods, and by the wage rate, w_t , that will be shown to be set by market sectoral knowledge levels and extraction costs.

Second, we assume that there is only Schumpeterian R&D, or equivalently that changes in learning-by-doing are negligible, which will be the case under our calibration in Section 5. In addition, we assume that modern R&D efforts involve consumption from the final good and not from human capital (lab-equipment approach as in Aghion & Howitt, 2009).

Assumption 6 (Stylized innovation) *Innovation only consists of modern R&D, that is $A_{k,t} = 1$ with $k \in \mathcal{K}$. R&D efforts involve expenditures in final good, with $H_{R,k,i,t}$ now replaced by $X_{R,k,i,t}$.*

The latter specification of R&D does not affect the incentives to perform R&D. In particular, Eq (46) still holds as it results from the free-entry condition in R&D. However, market clearing

conditions change in the sense that aggregate R&D expenditures, $X_t = \sum_{k \in \mathcal{K}} \int_0^1 X_{k,i,t} di$, are now inputs to the final good market and not the labor market (hence, there is no longer a general equilibrium link between R&D intensity and wages). We thus have the following labor market-clearing condition

$$H_t = H_{r,t} + H_{e,t}. \quad (49)$$

This new market clearing condition considerably simplifies the resolution and allows the closed form solution to be computed from this stylized model. At first, using (43) and the choice of the final good as the *numeraire*, that is $1 = p_{r,t}^{1-\sigma} + p_{e,t}^{1-\sigma}$, give the solution for equilibrium prices as

$$\begin{aligned} p_{r,t} &= \left[1 + \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{-(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{\gamma(\sigma-1)} \right]^{\frac{1}{\sigma-1}}, \\ p_{e,t} &= \left[1 + \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)} \right]^{\frac{1}{\sigma-1}}. \end{aligned}$$

Then, using Eq. (49), the equilibrium allocation of factors can be computed as

$$H_{k,t} = p_{k,t}^{-(\sigma-1)} H_t, \text{ and } E_{k,t} = \left[\frac{\gamma Q_{k,t}}{\Psi_{k,t}} \right]^{\frac{1-\alpha}{\beta}} \left[\frac{\alpha^2}{r_t + \delta} \right]^{\frac{\alpha}{\beta}} [p_{k,t}]^{-(\sigma-1) + \frac{1}{\beta}} H_t,$$

as well as the equilibrium input production as

$$Y_{k,t} = [Q_{k,t}]^{\frac{1-\alpha}{\beta}} \left[\frac{\gamma}{\Psi_{k,t}} \right]^{\frac{\gamma}{\beta}} \left[\frac{\alpha^2}{r_t + \delta} \right]^{\frac{\alpha}{\beta}} [p_{k,t}]^{-(\sigma-1) + \frac{1-\beta}{\beta}} H_t.$$

In order to close the model, we solve for r_t such that the physical capital market clears according to Eq. (39).

Substituting equilibrium physical capital use ultimately gives

$$\frac{r_t + \delta}{\alpha^2} = \left[\frac{H_t}{K_t} \right]^{\frac{\alpha}{\alpha+\beta}} \left\{ \left[\frac{\gamma}{\Psi_{e,t}} \right]^{\frac{\gamma}{\beta}} Q_{e,t}^{\frac{1-\alpha}{\beta}} p_{e,t}^{-(\sigma-1) + \frac{1}{\beta}} + \left[\frac{\gamma}{\Psi_{r,t}} \right]^{\frac{\gamma}{\beta}} Q_{r,t}^{\frac{1-\alpha}{\beta}} p_{r,t}^{-(\sigma-1) + \frac{1}{\beta}} \right\}^{\frac{\alpha}{\alpha+\beta}}.$$

5 Numerical analysis of the transition dynamics

In this section, we present the results of our numerical simulations. First, we focus on the historical experience of Great Britain for the period from 1700 to 1960. After introducing our baseline calibration, we discuss three counterfactual simulations that either assume (i) a higher extraction cost or (ii) a lower initial stock for the exhaustible resource, (iii) a slower diffusion of GPTs, and (iv) a slower pace for learning-by-doing productivity gains. We then turn to a comparison of two regions of the world that must be roughly comparable in terms of territory, population, and resources. [Pomeranz \(2000, pp. 7–10\)](#) has argued extensively that what must be compared to Great Britain is not China as a whole, but only its most advanced development centre at the time of the Great Divergence, i.e. the Yangtze Delta. As a corollary, if the entirety of China is chosen for reasons of data availability, it must be compared to Western Europe. In the following, we will use the term of "Eastern Asia" to refer to China, and "Western Europe" will be restricted to the aggregation of Great Britain, Sweden, France, Germany, Italy, Spain, Portugal, and the Netherlands. This exercise allows us to identify in our model the most crucial parameters that account for the lag in the transition from limited growth to sustained growth between these two regions of the world. Before analyzing these simulation results, we describe the strategy employed to calibrate the model.

5.1 General calibration strategy

Our numerical simulations begin in 1700 and follow a 20 year step. Due to the high dimensionality of our parameter and space of initial conditions, we rely first on several calibration assumptions, and then on the literature, to pin down the values of some structural and unambiguous parameters. In a third stage, we develop a best-fit calibration process to determine our baseline settings.

5.1.1 Assumptions

As for the theoretical analysis, we only consider two final input sectors respectively exploiting a *renewable* (indexed by r) and an *exhaustible* (indexed by e) energy resource. We chose the initial values of exhaustible energy and renewable potential, $\mathcal{R}_{r,0}$ and $\mathcal{R}_{e,0}$, so that both would be half depleted by 2000 in the British and Western European cases; and in the case of Eastern Asia, we assume that the remaining exhaustible stock and renewable potential in 2000 would equal the Western European stock and potential for comparison purposes.

Regarding the functional form for the probability of R&D success, we rely on Eq. (29) for our simulations. This choice is motivated by the numerical tractability of this functional form, which is intrinsically bounded by $\bar{\lambda}_k$ and admits a bounded derivative whenever R&D efforts tend toward zero (e.g. $H_{R,k} \rightarrow 0$). The alternative functional form, introduced in Eq. (45) to derive analytical results, would admit an infinite derivative, such that an equilibrium without R&D would not be possible.

Taking into account several constraints on parameters (namely $\forall k \in \mathcal{K}, \alpha_k + \beta_k + \gamma_k = 1$ and $\vartheta_r + \vartheta_e = 1$), the numerical version of our model accounts for 39 parameters and 10 initial values. To prevent any bias between final input sectors, we generalize **Assumption 1** and impose a symmetric calibration for both sectors, except for extraction cost functions of primary energy and R&D productivity gains. More specifically, we assume that the scale, $\bar{\Psi}_k$, and the convexity, $\psi_{\mathcal{R},k}$, of the extraction cost functions may differ, and that one sector may generate greater quality improvements, q_k , than the other (which could however be the case provided coal energy is for instance more concentrated than wood). Thus, we do not expect one R&D technology to be cheaper or to have a higher frontier success probability. In what follows, we consequently drop the sectoral indexes k whenever their is no ambiguity due to this symmetry assumption.

Furthermore, to ease the calibration process and the presentation of results, the relevant time-series that will be fitted (i.e. GDP, population, human capital per capita, total energy consumption) are normalized to unity in 1700.²⁹ We consistently assume that initial stocks of learning-by-doing (e.g., A_k), R&D (e.g., Q_k), and general purpose (e.g., G_t) knowledge is normalized to unity. These assumptions ensure that, apart from endogenous dynamics, directed technical change and relative use of energy factors (i.e., the energy transition) will only result from clearly identified structural differences in the extraction cost functions of the two sectors. We also consistently assume that the level of human capital per capita absent education, \bar{h} , is normalized to unity. The dimensionality of the calibration problem is thus reduced to 28 parameters and 4 initial values.

Lastly, due to the stochastic nature of the endogenous GPT generation, we adopt a Monte Carlo approach to present our numerical results. We will perform 10,000 runs for a given scenario with constant parameters, so that the resulting differences between runs only come from the stochastic generation of GPTs.³⁰ When presenting our simulation results, we will focus on the truncated average of these runs, but we will also display a 90% [0.05;0.95] probability interval.³¹

²⁹Consequently, we introduce a level parameter, Y_0 , in the final good production technology, such that $Y_t = Y_0 \left[\sum_{k \in \{r,e\}} Y_{k,t}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$ under **Assumption 1** (notably $\forall k \in \mathcal{K}, \vartheta_k = \vartheta$).

³⁰We also check that probability intervals in simulation periods do not significantly vary when the number of simulation increases.

³¹We compute truncated averages, that is averages within the [0.05;0.95] probability interval of each simulated period, to rule out numerical errors that appears on a few runs at the upper tail of these probability distributions. In what follows, we refer to “truncated mean” whenever we mention “mean”.

5.1.2 Use of related literature

In the next calibration step, we borrow the parameters related to production technologies from the literature. Precisely, α is set to $1/3$, so as to match the share of capital as in [Acemoglu et al. \(2012\)](#), and γ to $1/6$, which is an intermediate value between the modern and preindustrial cost-share of energy in England (respectively approximately 5% and 25% according to [Gentvilaite et al. \(2015\)](#)). This leaves a share of labor, β , set to $1/2$. Due to our symmetry assumption for final inputs, we assume $\vartheta = 1$.³² We also set the elasticity of substitution between final inputs, σ , to 4.4 as in [Kander & Stern \(2014\)](#). The usual assumption of a 5% annual depreciation rate for physical capital translates to setting δ to 0.64. We also assume that young adults save 20% of their revenue for an interior solution of [Problem 1](#), that is $s_t/z_t = 0.2$, which allows us to express η as a function of χ according to Eq. (5).

5.1.3 Best-fit calibration process

In a last step, we calibrate the 22 remaining parameters and 2 initial values so as to fit as best as possible some variables of our model with data time-series of interest, namely energy resource flows, GDP, population, human capital per capita and frequency of GPTs. The details of the calibration process are given in [Appendix D](#), so we only highlight the most important processes here.

Due to the high dimensionality of our calibration problem, we rely on sampling methods (design experiments protocols) to cover a wide range of calibration values that we discriminate using a scoring metric that assesses the quality of the fit. More precisely, departing from a compact calibration set that is mapped to an hypercube, we use a Latin Hypercube Sampling design with a *maximin* criteria to draw a set of calibration designs ([Morris & Mitchell, 1995](#)). Each of these designs is then used to perform a simulation, and each simulation receives squared-error based metrics (such as the sum of squared-error differences with regards to historical). The sampling technique allows us to reduce the dimensionality of a direct scoring approach on the whole (discretized) calibration set, by only considering a few well selected designs. The LHS feature ensures good projective properties in each dimension of the calibration set, and the *maximin* criteria (maximizing the minimal distance across designs) ensures the diversity of the selected designs.

5.2 The British Industrial Revolution

To analyse the British industrial revolution, we first propose a baseline calibration and then we discuss four counterfactual scenarios. We start the simulation in year 1700 and restrain the time horizon to 1960 because we consider that after that date the energy flows of our closed-economy model would depart too much from reality (in particular because of oil). To calibrate our model, we construct a database relying on the time-series developed by [Warde \(2007\)](#) and [Kander et al. \(2013\)](#) for the energy consumption,³³ [Broadberry et al. \(2015\)](#) and [Bolt et al. \(2018\)](#) for the real gross domestic product (GDP) per capita,³⁴ [Fouquet \(2014\)](#) for the population, [Lee & Lee \(2016\)](#) for the human capital per capita, and [Lipsey et al. \(2005, p. 132\)](#) for the frequency of GPTs.³⁵ Values of

³²To be precise, ϑ is set to 0.5 but the level parameter Y_0 that pre-multiplies the production technology in the final good sector allows to set the distribution parameters to 1.

³³The renewable primary energy resource of our model aggregates food, fodder, woodfuel, water and wind flows, whereas the exhaustible primary energy resource is an aggregate of coal, oil, gas, and nuclear.

³⁴Since we do not take into account international trade, our theoretical framework corresponds to a closed economy. It logically follows that the energy resources on which the economy is based do not match domestic natural assets, but rather correspond to the world energy resources necessary for the functioning of the British economy over the simulation period. Similarly, the final composite good of our model should correspond to the gross national product, and not the gross domestic product, of the United Kingdom. For obvious reasons regarding data availability, we will nevertheless use GDP as a first approximation of the macroeconomic output.

³⁵We acknowledge that across (and even within) these different data sources there are several inconsistencies in terms of geographical boundaries (e.g., [Warde's \(2007\)](#) energy consumption estimates concern England and Wales, whereas [Lee & Lee's \(2016\)](#) estimates of human capital stock are for the United Kingdom). Nevertheless, it can be assumed that if the *absolute* values of historical variables change significantly with the geographical boundary under consideration, this is less true for their *relative* dynamics over time. Hence, normalizing all variables to unity in 1700 minimizes the impact of this geographical border mismatch on our simulation results.

parameters and initial variables obtained through our calibration procedure are given in [Appendix E](#). It is worth mentioning that we retrieve the following intuitive features at the end of the calibration process: (i) for the same level of remaining resource and technology, it is relatively more costly to exploit the exhaustible resource, that is $\bar{\Psi}_r < \bar{\Psi}_e$ (e.g., cutting wood is *ceteris paribus* easier than mining for coal), (ii) the exhaustible extraction cost is relatively more convex with respect to the level of remaining stock, that is $\psi_{\mathcal{R},r} \leq \psi_{\mathcal{R},e}$ (e.g., mining for the last 10% of the remaining stock of coal is more expensive than to cut the analogous last 10% of the remaining wood resource), (iii) the overall extraction cost is higher in the exhaustible sector in the beginning of our simulations, that is $\Psi_{e,1700} > \Psi_{r,1700}$, and (iv) R&D quality improvements appears to be equivalent in both sectors, that is $q_r \approx q_e$.

5.2.1 Set of scenarios

We distinguish five scenarios for the numerical analysis of our model:

- *Baseline* : follows the calibration procedure given above to fit the model as best as possible to the British historical data.
- *High cost*: is a counterfactual simulation where the level parameter, $\bar{\Psi}_e$, in the extraction cost function of the exhaustible sector is ten times as large as in the baseline scenario. This second scenario investigates the possible impact of a similar size but less concentrated exhaustible energy stock on the timing of the British industrial revolution.
- *Low stock*: is a counterfactual simulation where the initial stock of the exhaustible resource, $\mathcal{R}_{e,0}$, is ten times lower than in the baseline scenario. This third scenario also investigates the impact of the exhaustible resource scarcity on the economic take-off, but instead of assuming a less concentrated resource as in the *high cost* simulation, here the concentration of the resource is the same but its overall size is ten times lower.
- *Low diffusion*: is a counterfactual simulation where the diffusion speed of GPT, ξ , is ten times lower than in the baseline calibration. In this fourth scenario, we investigate the impact of knowledge diffusion on demography and its ultimate consequences on the transition from limited to sustained economic growth.
- *Low learning*: is a counterfactual simulation where the learning-by-doing productivity gains common to each final input sector, Ω , are ten times lower than in the baseline calibration. In this fifth scenario we assess the impacts on economic growth of lower efficiency of institutions in gathering and accumulating knowledge before the advent of formal R&D.

5.2.2 Simulations results

[Figure 3](#) presents the numerical results of our *baseline* scenario, that is, the mean and the [0.05;0.95] probability interval of the 10,000 runs, against the British historical data. The model fits quite well the time-series of GDP, population, GDP/capita, human capital per capita, and energy resource flows. The advent of the Great Depression during the 1930s, and then World War II (WWII), induce a drop in all data time series in 1940 and 1960. Our model is of course not designed to replicate such exceptional events, and, as a consequence, our simulations naturally tend to overshoot the 1940 and 1960 historical points. To take this fact into account, we perform (exponential or quadratic) data interpolations from 1900 to 2000 to produce "extreme-events-adjusted data" for 1940 and 1960, which are displayed in crosses in [Figure 3](#).³⁶ The fit of the *baseline* scenario is consistent with these adjusted data, including the GPT's frequency as observed in [Figure 9](#) of [Appendix H](#), which illustrate that our model is a good representation of an economy in the absence of major external shocks.

³⁶A more rigorous statistical inference of data trends is far beyond the scope of this paper. The "extreme-events-adjusted data" points are only provided here for the sake of illustration.

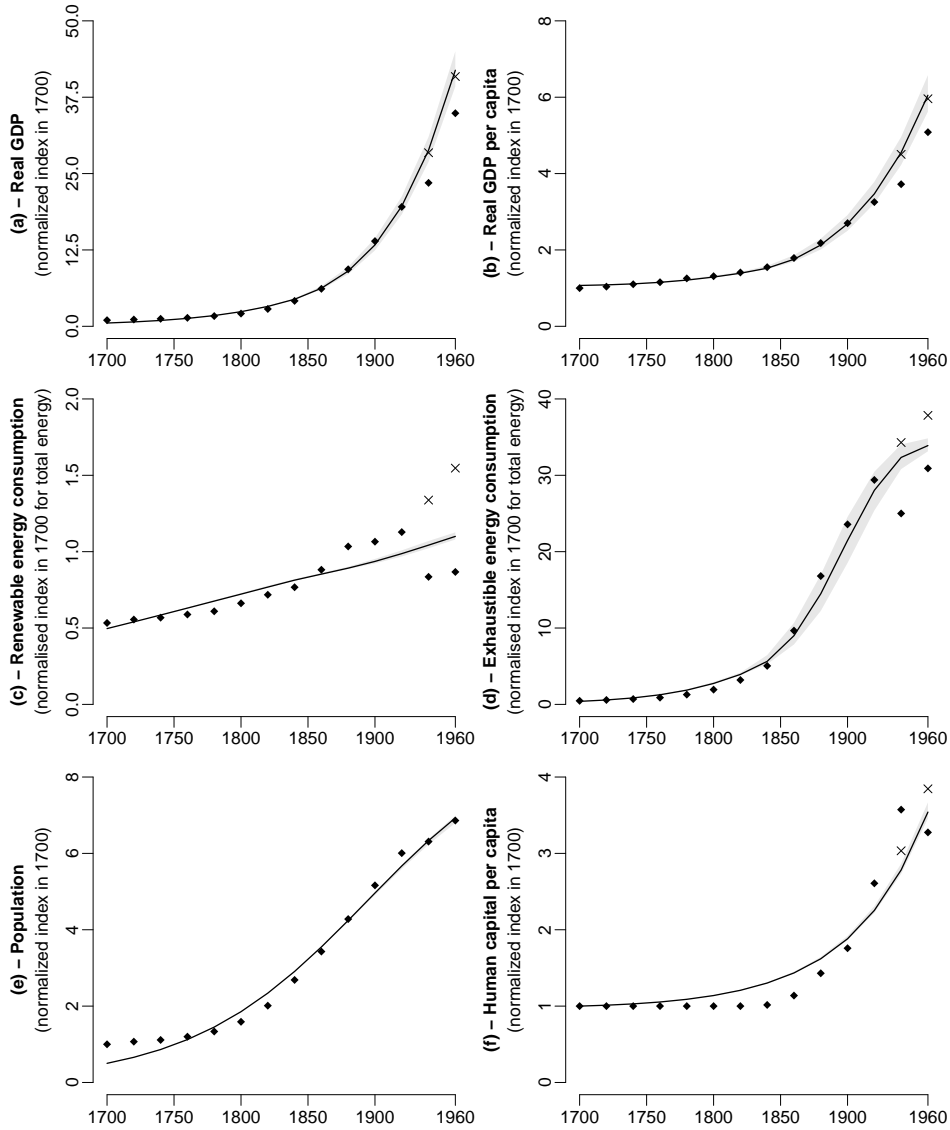


Fig. 3 Comparison of baseline calibration (solid lines are averages, shaded areas represent 90% probability intervals) with historical data (diamonds) and extreme-events-adjusted data (crosses), 1700–1960

One can notice that the model does not perfectly fit the changing convexity of population. This issue cannot be corrected insofar as fertility exhibits a monotonous decreasing evolution from an upper to a lower bound. The introduction within **Problem HH** of a survival consumption threshold, c_0 , associated with an additional constraint $c_t \geq c_0$ as in the seminal paper of [Galor & Weil \(2000\)](#), would solve this issue by generating an initial ramp-up period for fertility. However, this solution significantly increases the difficulty of the calibration procedure. Therefore, since our focus is on the energy transition in this section, we limit ourselves to the formulation introduced so far to keep the calibration manageable.

We now turn to the comparison given in [Figure 4](#) and [Table 1 to 4](#) between our baseline calibration and the four counterfactual simulations. Considering the *high cost* scenario, one can see that a factor of ten increase in the extraction cost of the exhaustible resource implies that, by 1900, nonrenewable energy consumption is 91.1% lower than in the baseline calibration. Relative to the baseline calibration, the renewable energy resource is 8.8% more heavily exploited and the gradual

transition towards exhaustible energy is delayed by almost a century (substitution effect). However, as the renewable resource use increases more rapidly towards its ultimate natural limit in this counterfactual simulation, its extraction cost also increases more rapidly, and ultimately it is not abundant enough to fuel the Industrial Revolution. Accordingly, in 1900 this alternative energy transition dynamic results in an 86.9% total energy consumption decrease and a 34.6% loss in GDP per capita when average trajectories of the *high cost* and *baseline* scenarios are compared.

Relative to the *high cost* counterfactual, the detrimental impacts of the *low stock* scenario are initially less significant, but by 1900 they are equivalent, and after that they are even higher. The persistent fall in output suggests that economic growth is hindered by the lack of available exhaustible resource compared to the *baseline* and *high-cost* scenarios. In the *low stock* scenario, as in the *high cost* counterfactual, the lack of availability of the exhaustible stock is only slightly compensated by a higher use of the renewable resource (around 6-9% both in 1900 and 1960). However, the consumption patterns of nonrenewable energy are very different in both counterfactual scenarios. In the *high cost* simulation, the exploitation of the exhaustible stock is delayed but it broadly follows the baseline S-shape. In the *low stock* simulation the nonrenewable energy consumption is initially higher than in *high cost* scenario, but it does not feature an exponential take-off and instead merely increases. As a result, the decrease in GDP/capita is approximately the same in 1900 (around 30-35%), but in 1960, the GDP/capita loss is still about 35% in the *low stock* simulation and only 21.2% in the *high cost* simulation.

Interestingly, the impact on demography is quite moderate in both the *high cost* and *low stock* counterfactual simulations. This is not surprising insofar as the link between population and energy use goes through the channel of accumulated knowledge. Indeed, all kind of research, whether through learning-by-doing or R&D, is boosted by the arrival of GPTs, and is aggregated within the stock of global useful knowledge that drives the educational choices operated by the representative household. Hence, within both the *baseline* calibration and the *high cost* and *low stock* counterfactual simulations, learning-by-doing and R&D, which start smoothly in 1800 in the exhaustible sector and fifty years later in the renewable sector, are significant enough to make education worthwhile by the beginning of the XIXth century, thus triggering the demographic transition.

A factor of ten decrease in the diffusion speed of GPTs results in a slightly lower consumption of renewable energy, but the decrease in exhaustible energy consumption is far more significant. In 1900, the energy transition dynamics of the counterfactual *low diffusion* scenario are characterised by an overall energy supply that is 31% lower than in the *baseline* calibration. In addition, the slower diffusion of GPTs implies a lower efficiency for both the learning-by-doing and the R&D processes, which together with the lower energy supply, result in a GDP per capita loss of 17.5% in 1900 and 21.8% in 1950, when mean trajectories from the *low diffusion* and *baseline* simulations are compared. Moreover, the hampered knowledge accumulation of the *low diffusion* simulation also affects the child quantity-quality trade-off of the representative household, which in 1960 generate a 3.2% higher population level and a 7.5% loss in human capital per capita between average trajectories.

Next, the *low learning* scenario shows that reducing the efficiency of learning-by-doing in production by a factor of ten has significant impacts and emphasizes the importance of the knowledge-demography-energy nexus. Since global learning accumulates more slowly, the shift from high fertility/low education towards low fertility/high education (the so-called quantity-quality trade-off) is postponed. The resulting lower supply in human capital penalizes energy extraction processes and delays R&D technical change, especially in the exhaustible sector, which overall generates a negative impact on economic growth. As a result, compared to the *baseline* calibration, the energy transition is delayed as both renewable and exhaustible energy consumption experience a decrease of 11.8% and 61% respectively in 1900. Among all counterfactual simulations, the delayed energy transition of the *low learning* scenario generates the strongest lag in economic take-off, with relative GDP per capita losses of 44.9% and 51.3% in 1900 and 1950 respectively. Furthermore, as a consequence of the delayed energy transition and economic take-off, the demographic transition is also postponed, which results in a larger population (+9.2% in 1900 and +18.7% in 1960) with an average lower human capital per capita (-14.6% in 1900 and -26.6% in 1960).

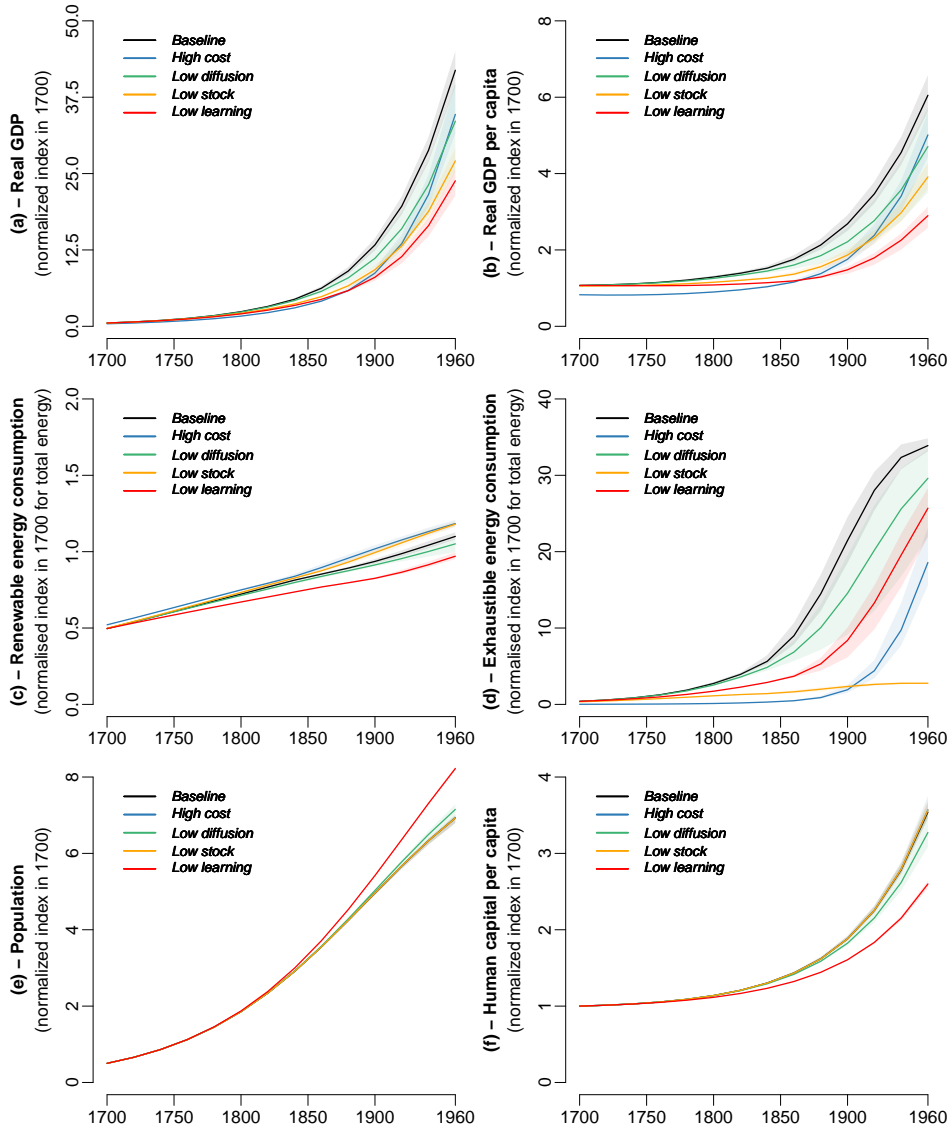


Fig. 4 Comparison of *baseline calibration* (black) with counterfactual simulations *high cost* (blue), *low stock* (green), *low diffusion* (orange) and *low learning* (red), 1700–1960 (solid lines are the averages, shaded areas represent 90% probability intervals)

Finally, we turn to the investigation of the key transition mechanism advocated in this paper: the direction of technical change toward the exhaustible resource sector. **Figure 5** presents a phase diagram of the growth rate of R&D-based technical change in each sector as given by Eq. (35). In this figure, technical change is biased towards the exhaustible sector in the space above the first bisector, because in that upper half of the graph R&D-based knowledge grows faster in the exhaustible sector than in the renewable one; and reciprocally, there is a directed technical change (DTC) toward renewable energy in the lower half of the graph. First, let us remark that R&D-based technical change only starts in the XIXth century (1820 in the *baseline* calibration). Indeed, according to Eq. (32), R&D is not profitable enough in previous periods, notably because of insufficiently mature GPTs and a lack of accumulation in learning-by-doing that together penalize potential monopoly profits.³⁷ Second, in all but the *high-cost* scenario, technical change is biased toward the exhaustible sector

³⁷In other words, the productivity threshold implied by Eq. (32) is initially not met.

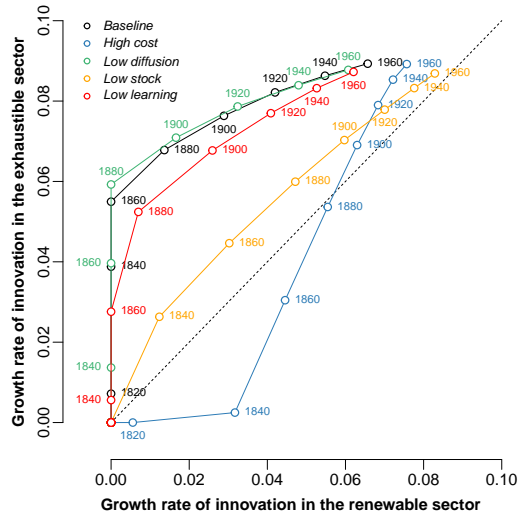


Fig. 5 Phase diagram of the growth rate of R&D-based knowledge in the *baseline* (black), *high cost* (blue), *low stock* (green), *low diffusion* (orange) and *low learning* (red) calibrations (points are averages, dashed line represents the first bisector)

whenever R&D starts. This event coincides with the acceleration of both the exhaustible resource consumption and GDP growth. In particular, in the *baseline*, *low diffusion* and *low learning* cases, R&D even exclusively occurs in the exhaustible sector until 1860, 1880, and 1860 respectively. This is in line with our analytical results, which emphasizing that the rising extraction cost in the renewable sector, due to a shortage in the available resource (in 1800, about 50% of the stock of the renewable resource is extracted in the *baseline* scenario against a mere 2% of the exhaustible one), causes R&D-based technical change to be redirected towards the exhaustible sector. The *high cost* scenario also supports this analysis: a ten-times higher extraction cost in the exhaustible sector reverses this pattern and initially drives innovation toward the renewable sector. However, in this scenario 63% of the renewable resource is extracted by 1880, and the associated rising extraction costs in the renewable sector ultimately drive R&D towards the exhaustible sector. This reversal of incentives again coincides with a delay in the acceleration of economic growth. Third, one can notice that at the turn of the XXth century, R&D in the renewable sector starts to gain momentum in the *baseline*, *high cost*, *low diffusion*, and *low learning* scenarios, which result in a more balanced technical change at the end of the simulation horizon (1960). Indeed, due to large accumulation of knowledge, GPTs arrival and diffusion, and rising costs in the exhaustible sector (about 30% of the exhaustible stock has been extracted by 1940 in the *baseline scenario*), R&D also becomes competitive in the renewable sector, thus reducing (but not offsetting) the initial technical bias towards the exhaustible sector. Finally, the *low stock* scenario yields the most balanced pattern for directed technical change because of the impact of the significantly lower stock of exhaustible energy. Indeed, whenever the remaining energy potentials are more comparable in the two sectors, the incentives to perform more intensive R&D in specific sector vanish because the cost differential is weaker. However, it is worth mentioning that, despite a tenfold reduction in the exhaustible resource size in this scenario, R&D is still biased towards the exhaustible sector. A comparison between the *high cost* and *low stock* scenarios highlight that the bias in DTC is not unilateral in our approach, as analyzed in Section 4.2. It is indeed the strength of the relative extraction cost differential that shapes the bias of DTC, initially towards the renewable sector in the *high cost* scenario. However, the depletion of stocks might trigger a reversal in the initial direction. Indeed, the renewable sector only has a limited potential (defined by the annual solar inflow) that is already significantly used relatively to the exhaustible one at the beginning of the XVIIIth century. This is why all scenarios ultimately end in a DTC bias towards the exhaustible sector, concomitant with the economic take-off.

Table 1 Percentage change in renewable (E_r), exhaustible (E_d), and total ($E = E_r + E_d$) energy consumption, between the average trajectories of counterfactual scenarios and the *baseline* calibration

Chg. w.r.t. Baseline	ΔE_r in %			ΔE_d in %			ΔE in %		
	1800	1900	1960	1800	1900	1960	1800	1900	1960
<i>High cost</i>	+3.8	+8.8	+7.6	-95.7	-91.1	-45.2	-75.0	-86.9	-43.5
<i>Low stock</i>	+1.6	+6.2	+7.3	-59.1	-89.1	-91.8	-46.4	-85.1	-88.7
<i>Low diffusion</i>	-1.2	-2.5	-4.4	-7.4	-32.2	-12.6	-6.1	-31.0	-12.4
<i>Low learning</i>	-7.1	-11.8	-11.8	-37.1	-61.0	-24.2	-30.8	-58.9	-23.8

Note: For any variable $X \in \{E_r, E_d, E\}$ of any counterfactual scenario (c), we denote the percentage change with respect to (w.r.t.) the baseline (b) as ΔX_c , with $\Delta X_c = (X_c - X_b)/X_b$.

Table 2 Percentage change in GDP (Y) and GDP per capita (Y/P), between the median trajectories of counterfactual scenarios and the *baseline* calibration

Chg. w.r.t. Baseline	ΔY in %				$\Delta(Y/P)$ in %			
	1800	1850	1900	1950	1800	1850	1900	1950
<i>High cost</i>	-30.4	-32.9	-34.6	-21.3	-30.4	-32.9	-34.6	-21.2
<i>Low stock</i>	-10.6	-19.7	-30.4	-35.0	-10.6	-19.7	-30.5	-35.0
<i>Low diffusion</i>	-2.6	-6.3	-16.5	-19.8	-2.7	-6.7	-17.5	-21.8
<i>Low learning</i>	-15.2	-26.0	-39.8	-43.0	-16.1	-28.7	-44.9	-51.3

Note: For any variable $X \in \{Y, Y/P\}$ of any counterfactual scenario (c), we denote the percentage change with respect to (w.r.t.) the baseline (b) as ΔX_c , with $\Delta X_c = (X_c - X_b)/X_b$. Given that the time step of the model is 20 years, starting in 1700, the results for 1850 and 1950 are obtained through exponential interpolations of our simulated results.

Table 3 Percentage change in the compound annual growth rate of GDP ($CAGR_Y$) and exhaustible energy ($CAGR_E$) between the average trajectories of counterfactual scenarios and the *baseline* calibration

Chg. w.r.t. Baseline	$CAGR_Y$ in p.p.			$CAGR_E$ in p.p.		
	1800-1850	1850-1900	1900-1950	1800-1850	1850-1900	1900-1950
<i>High cost</i>	-0.08	-0.05	+0.38	-0.94	-0.37	+2.42
<i>Low stock</i>	-0.22	-0.29	-0.14	-1.18	-1.41	-0.53
<i>Low diffusion</i>	-0.08	-0.23	-0.08	-0.26	-0.37	+0.39
<i>Low learning</i>	-0.28	-0.42	-0.11	-0.64	-0.41	1.03

Note: For any variable $X \in \{Y, Y/P\}$ of any counterfactual scenario (c), we denote the percentage change with respect to (w.r.t.) the baseline (b) as ΔX_c , with $\Delta X_c = (X_c - X_b)/X_b$. Given that the time step of the model is 20 years, starting in 1700, the results for 1850 and 1950 are obtained through exponential interpolations of our simulated results.

5.3 Comparative analysis of Western Europe and Eastern Asia

To analyse the comparative development of Western Europe (i.e., the aggregation of Great Britain, Sweden, France, Germany, Italy, Spain, Portugal, and the Netherlands) against Eastern Asia (i.e., China), hereafter respectively labelled WE and EA, we first propose a baseline calibration for each geographical area, and then discuss a counterfactual scenario for Eastern Asia. As for the United Kingdom, we start the simulation in year 1700, but we extend the time horizon to 2000 to better observe the beginning of the economic take-off of Eastern Asia. To calibrate our model, we construct a database relying on the time-series developed by [Kander et al. \(2013\)](#) for the energy consumption,

Table 4 Delay (years) to reach the average levels of GDP ($Delay_Y$), renewable ($Delay_{E,c}$), exhaustible ($Delay_{E,d}$) and total energy ($Delay_E$) of the *baseline* scenario for the counterfactual scenarios

Chg. w.r.t. Baseline	Delay _Y in yr.		Delay _{E,c} in yr.		Delay _{E,d} in yr.	
	1850	1900	1850	1900	1850	1900
<i>High cost</i>	+23.1	+19.1	-12.4	-26.9	+79.2	n.r.
<i>Low stock</i>	+14.5	+20.6	-7.6	-18.5	n.r.	n.r.
<i>Low diffusion</i>	+3.6	+9.0	+7.5	+11.1	+11.7	+24.7
<i>Low learning</i>	+21.7	+27.5	+53.7	+47.8	+41.8	+46.2

Note: “n.r.” means “not reached within the simulated horizon (1700-1960)”. The top-left square reads “in the *high-cost* scenario, it take 23.1 more years to reach the GDP level of the *baseline* scenario in 1850”. Given that the time step of the model is 20 years, starting in 1700, the results for 1850 and 1950 are obtained through exponential interpolations of our simulated results.

Malanima (2011), Fouquet & Broadberry (2015), Broadberry *et al.* (2018) and Bolt *et al.* (2018) for the real gross domestic product (GDP) per capita, (Maddison, 2007, p. 168), (Malanima, 2009, p. 9), Kander *et al.* (2013) and World Bank (2018) for the population, Lee & Lee (2016) for the human capital per capita.³⁸ Values of initial variables and parameters obtained through our general calibration procedure introduced in Section 5.1 are given in Appendix F for Western Europe, and in Appendix G for Eastern Asia. We must also mention that our simulated trajectory might suffer from two biases that appears in our data-sets: (i) the extreme events of WWI and WWII, which significantly affect our reconstructed time series for Western Europe and explain why the model tends to over-fit historical data points from 1940 to 1980, and (ii) the high degree of openness of the economy after the 1970s (and even before that date), which might explain why the fit of our model is worse after this date, especially for energy consumption flows.

5.3.1 Set of scenarios

We distinguish three scenarios for the numerical analysis of Western Europe versus Eastern Asia:

- *Baseline WE* : follows the general calibration procedure of Section 5.1 to fit the model as best as possible to the Western European historical data.
- *Baseline EA* : follows the general calibration procedure of Section 5.1 to fit the model as best as possible to the Eastern Asian historical data.
- *Alternative EA* : is a counterfactual simulation for Eastern Asia investigating the order of magnitude changes that are needed in specific parameters to obtain an Eastern Asian trajectory that more closely matches the Western European dynamics. In this experiment, we maintain all parameters (in particular the initial levels of energy resources, $\mathcal{R}_{k,0}$) except for four of them whose values are set to be of the same order of magnitude as in the *baseline WE* calibration, namely (i) the general productivity, Y_0 , is multiplied by 1.3, (ii) the renewable resource extraction cost parameters, $\bar{\Psi}_r$ and $\psi_{\mathcal{R},r}$, are multiplied by 3 and 1.5 respectively, and (iii) the learning-by-doing technical change, Ω , is multiplied by 10.

5.3.2 Simulation results

Figure 6 (and Figure 9 in Appendix H for GPT’s frequency) presents the numerical results of our three scenarios, that is, the mean and the [0.05;0.95] probability interval of the 10,000 runs against historical data. As a general comment, one can see that the overall fit of the model is very satisfactory for both *baseline WE* and *baseline EA*, except for the population and human capital variables

³⁸Of course, the time series retrieved from these different references are not on a year-by-year basis during the pre-modern period. Hence, we performed (exponential or polynomial) interpolations for these time series to reconstruct estimates every twenty years so as to fit our modeling setup.

of the later. As previously mentioned in [Section 5.2.2](#), this later aspect should not be surprising because fertility choices are monotonic in our framework, while the Eastern Asia case clearly exhibits a ramp-up period for fertility. Again, the introduction of a survival consumption threshold, $c_t \geq c_0$ as in the seminal paper of [Galor & Weil \(2000\)](#), would solve this issue at the cost of a significantly more difficult calibration process, which we refrain from doing in this paper. Moreover, in the *baseline WE* simulation, one shall observe an over-consumption of exhaustible energy during the WWII and the following recovery periods, obviously not modelled within our framework. The resulting additional depletion of the exhaustible resource (with regards to historical data) coupled with a strengthening globalization – associated with increasingly intense international energy flows, notably oil-related – since the 1970s, might explain why the model tends to then under-fit the historical exhaustible energy consumption in the last period of simulation. To take into account these event, we again produce “extreme-event-adjusted data” from 1940 to 1980 which are displayed in [Figure 6](#) with crosses. The fit of the two *baseline* scenarios, especially in the case of Eastern Asia, is consistent with these adjusted data points, which illustrates that our model is here again a good representation of an economy in the absence of major external shocks.

In terms of comparative dynamics, given that all historical time series are normalized to unity in the first simulation period (i.e., 1700), and considering that we used our general calibration procedure to fit the model separately for WE and EA, the discrepancies obtained in parameters’ values may be interpreted as differences in deep-rooted determinants of bio-geographical, cultural, or institutional origins that explain the different dynamics of the two world regions. Several significant differences appear when parameter values of baseline WE and baseline EA are compared. First, while extraction costs of the exhaustible resource are rather similar for both geographical areas,³⁹ renewable energy appears to be easier to extract in EA than in WE, as illustrated by the -67% and -35% respective differences in $\bar{\Psi}_r$ and $\psi_{\mathcal{R},r}$ between the two world regions. This reflects a lower accessibility (i.e., a higher difficulty of extraction) of the renewable energy resource in WE compared to EA, which explains in part why this energy source is more extensively used in EA. Second, learning-by-doing technical change — meaning the efficiency to convert useful knowledge contained in human capital and GPT-related know-how into applied knowledge in the pre-industrial regime — is ten times slower in EA compared to WE, as captured by the -93% difference in the Ω parameter.⁴⁰ This important difference in technical efficiency is strengthened by the initial general productivity, Y_0 , that is found to be 37% higher in WE compared to EA. Third, profit-motivated R&D is quite different in EA compared to WE, in the sense that it is a bit more costly (3% higher ϕ), it has a higher maximum R&D success probability (13% higher $\bar{\lambda}$), and it yields lower productivity gains in the renewable sector (15% lower q_e) but higher productivity gains in the exhaustible one (12% higher q_e). This result could be interpreted as EA having a relatively more efficient fossil-related R&D sector. However, this last point should be tempered by the fact that the Schumpeterian fossil-fueled regime only really begins in the middle of the XXth century in Eastern Asia, while it is in place for a significant longer period in WE. If the two world regions spent similar times under the fossil energy regime, the impact of this calibration bias would be lowered and differences in R&D-related parameters would more closely reflect discrepancies in actual productivity.

When parameters are changed according to the *alternative EA* scenario to create a counterfactual for EA with higher renewable energy scarcity, higher learning-by-doing technical change and higher initial general knowledge, it is striking to observe how the pattern of the simulation becomes similar to the *baseline WE* scenario. This is mostly visible in terms of the timing of the real GDP take-off and trajectories of exhaustible energy consumption flows, as well as in the pattern of R&D-bias displayed in [Figure 7](#). This figure shows that in the *baseline EA* scenario, only the renewable sector experiences significant R&D until the end of the XIXth century. Thus, in line with the lower accumulation of learning-by-doing technical change and more concentrated energy in the renewable

³⁹To be precise, the $\bar{\Psi}_e$ and $\psi_{\mathcal{R},e}$ parameters in the *baseline EA* calibration are respectively slightly higher and lower than their counterparts in the *baseline WE* scenario, which results in comparable extraction cost functions, at least during the pre-industrial era.

⁴⁰The productivity parameter, A_E , of the schooling technology is 16% lower in EA compared to WE, which suggests that the overall productivity of the schooling technology, $\frac{Q_t}{1+Q_t} A_E$, is lower in EA compared to WE.

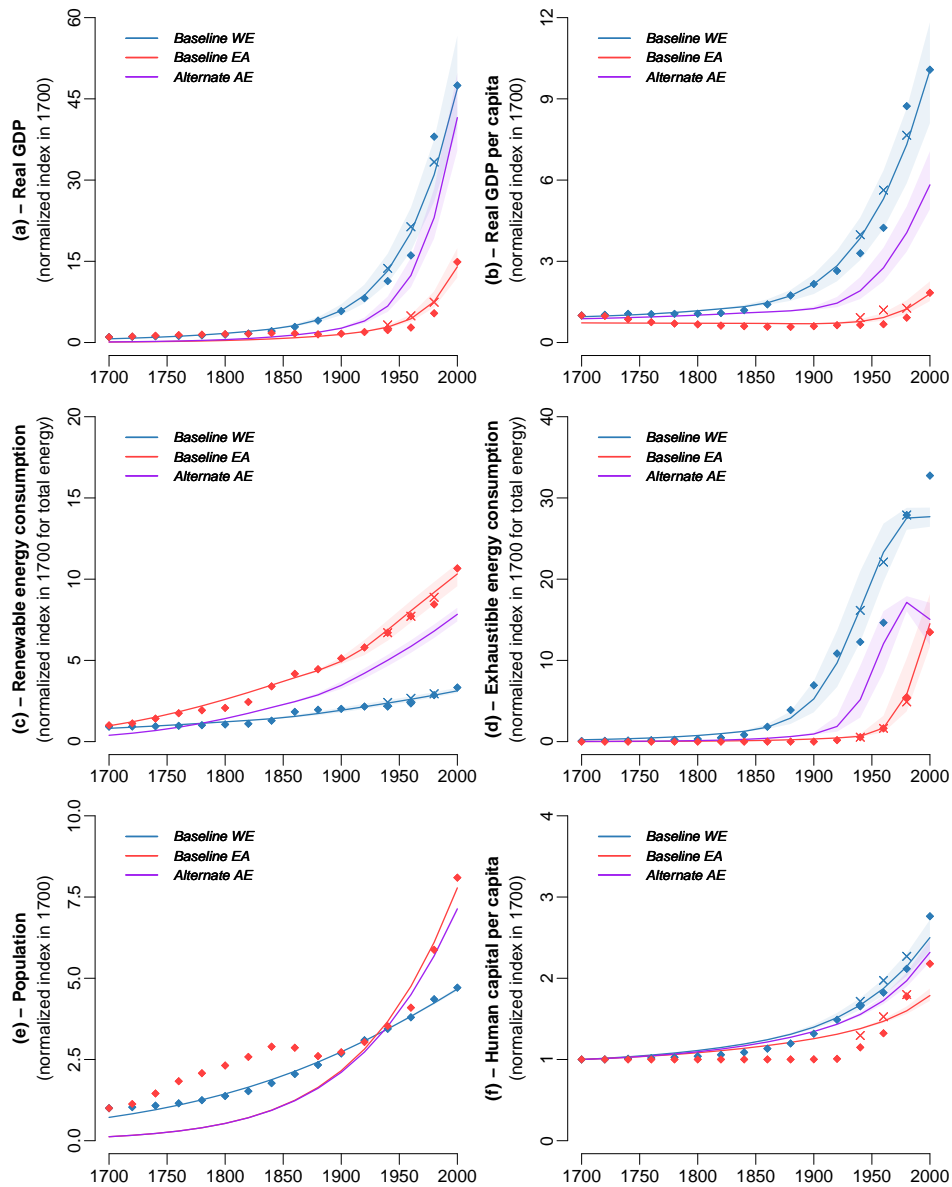


Fig. 6 Comparison of *baseline WE* calibration (*blue*) with *baseline EA* calibration (*red*) and *alternative EA* calibration (*purple*), 1700–2000, (*solid lines are averages, color shade are 90% probability intervals, diamonds are historical data and crosses are extreme-event-adjusted data*)

resource emphasized above, R&D is initially strongly biased against the exhaustible sector that only become exploitable by the middle of the XXth century, concomitantly to the observed economic take-off. By contrast, in the *alternative EA* scenario, directed technical change is more rapidly biased toward the exhaustible sector, notably because knowledge accumulation is more efficient (which also implies that GPTs become more frequent and diffuse more widely) and because the renewable resource is more costly. Again, the shift towards a strong exhaustible sector R&D-bias in both the *baseline WE* and *alternative EA* scenarios coincides with the large-scale use of the exhaustible resource and the beginning of the economic take-off, which is the key mechanism advocated for in this paper.

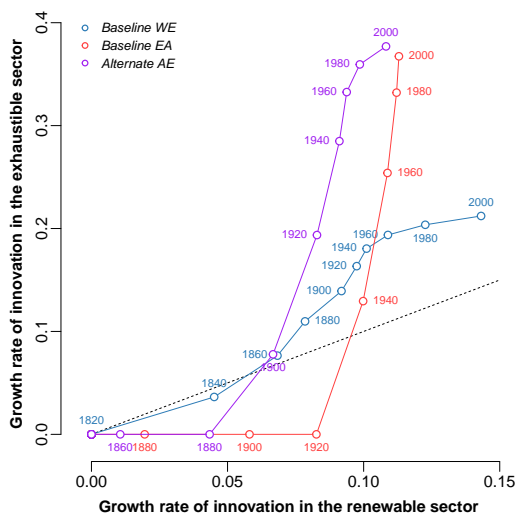


Fig. 7 Phase diagram of the growth rate of R&D-based knowledge in the *baseline WE* (blue), *baseline EA* (red) and *alternative EA* (purple) calibrations (points are averages, dashed line represents the first bisector)

6 Conclusion

In this paper, we develop a knowledge-based and energy-centered unified growth model to investigate the transition from limited to sustained growth. Our results support the idea that energy systems should be more consistently analyzed as a central feature of long-term patterns of technical change and economic development, especially during major transitions such as the industrial revolution.

To varying degrees, all growth regimes before industrialization can commonly be characterized as *organic-Smithian*, i.e., limited and non-sustained growth supported by an increasing division and specialization of labour and extension of the market fueled by renewable organic resources derived from solar energy (including hydropower, wood, agricultural wastes and muscle energy from human and draft animals, which all ultimately derive from solar energy).⁴¹ On the contrary, modern economic growth can be defined as *mineral-Schumpeterian*, i.e., substantial and sustained growth through a massive use of new fossil energy sources and raw materials, new technologies driven by continuing innovation, and new institutions.⁴² As stated by Wrigley (2013, p. 2, 9–10), in all organic societies, growth prospects were necessarily limited because the size of the accessible energy supply was too constrained to allow for sustained growth. A necessary condition for the move from a world where Smithian growth was at best asymptotic, to one in which it could be Schumpeterian-exponential in the long run was dependent upon the discovery and exploitation of a vast reservoir of energy that had remained untapped in all organic economies.

The numerical results we derived from counterfactual scenarios of the British experience emphasize the role that the energy bonanza from coal used in steam engines had in enabling the British economic ‘efflorescence’ of the seventeenth and eighteenth centuries to become a genuine industrial revolution. In our model, if energy *per se* is not a strict necessary condition for a Schumpeterian regime, it certainly appears as a catalyst for sustained economic growth. Our numerical simulations indeed suggest that lowering the stock of exhaustible energy by one order of magnitude delays the economic take-off, which result in a 35% lower GDP per capita in 1950. Another way of inter-

⁴¹ At specific times of minor technical changes and limited labour specialization, ‘mere extensive’ growth would better characterise societies in which total output grows only as the result of population and/or territory increase. In extensive growth, new resources in the form of added labour or land are added to the economy, but population and total output grow at roughly the same rate, so that per capita income is rather stagnant (Goldstone, 2002, p. 324).

⁴² Another important distinction is that improved long run economic performance associated with the progressive transition from *organic-Smithian* to *mineral-Schumpeterian* regimes occurs not so much through an increase in the rate of growth, but through a decline in the rate and frequency of shrinking episodes (Broadberry & Wallis, 2017).

preting the same result is to say that in this counterfactual, nearly 20 additional years are needed to achieve the level of GDP per capita reached in 1900 in the baseline scenario. In addition, in our model the channel of knowledge accumulation is key in unlocking the industrial potential of the exhaustible energy resource: lowering learning-by-doing technical change by one order of magnitude has the consequence of dividing the exhaustible energy consumption by a factor of more than two in 1900, which results in a 45% drop in output per capita at the same time. In this case, the resulting delayed demographic transition magnifies the downside impact of knowledge accumulation, hindering R&D technical change and energy resource extraction. The knowledge-energy-demography nexus that we analytically analyze thus sheds some light on both the timing and the magnitude of the industrial revolution.

Goldstone (2002, p. 362) rightly points out that even though China certainly had all the raw materials and all the necessary technology, it did not achieve an efflorescence from adapting water power to the mechanical manufacturing of cotton textiles. The explanation for the absence of such a water power-based efflorescence is still controversial, but in any case, Goldstone (2002, p. 363) notes that even if China had made the step of deploying water-powered spinning, that by itself would not have launched an industrial revolution. To do so would have required a systematization of knowledge regarding air pressure, power, and its application to constructing heat engines as occurred earlier in England. Unlike water-powered spinning, there is no evidence of such an independent trajectory in China. Through the simulations we undertook to compare Western Europe and Eastern Asia, the model developed in this paper was also able to support these ideas. Indeed, the calibration obtained from our best-fit procedure suggested both a lower initial global efficiency for Eastern Asia compared to Western Europe, and a slower pace of learning-by-doing technical change. This finding suggests that, according to our model, Eastern Asia had a lower efficiency in accumulating knowledge in the pre-industrial regime. Moreover, when we changed some of the most important parameter values in Eastern Asia to make them comparable to the Western European setting (that is, comparable initial general knowledge, learning-by-doing accumulation, and renewable resource accessibility), the counterfactual patterns were logically more in line, both in terms of timing and magnitude of the industrial revolution.

Finally, the significant relationship between economic growth, technical change, demography, and energy consumption suggested throughout this paper calls for further refinements, especially in the context of the transition towards renewable energy that indisputably needs to accelerate in the coming decades to tackle the issues of climate change IPCC (2013). The investigation of this nexus appears to be key to shape policy recommendations dealing with technical change, energy use and demography that would be consistent between each other and remain in line with inclusive growth issues. The investigation of such optimal policies and the resulting inter-generational equity concerns are left for further research.

Appendix A Historical evidences: child quantity-quality trade-off, human capital, and energy

Appendix A.1 The controversial child quantity-quality trade-off

If some studies, such as Cáceres-Delphiano (2006) for the USA, Li *et al.* (2008) for China, and Becker *et al.* (2010) for Prussia, find the expected negative family-size/child-quality relationship, other empirical studies, such as Angrist *et al.* (2010) for Israel, Black *et al.* (2005) for Norway, and Clark & Cummins (2016) for England, find no evidence of such a quantity-quality trade-off. Regarding the emblematic case of Britain on the period 1780–1880, Clark & Cummins (2016) find that family size did not affect education, occupation, longevity, or even wealth. On the wider 1580–1830 period, Wrigley *et al.* (1997, p. 461) suggest that natural fertility was the norm in England, so that small groups may have been practising family limitation, but the reconstitution evidence suggests that such behavior was restricted to a small minority of the population, if present at all.

Clark & Cummins (2016) conclude that modern growth consequently cannot be explained by a switch to smaller family sizes accompanied by more investment in child quality. Modern growth in England had begun 100 years before there were significant reductions in average family sizes. This one hundred year delay between the ignition of the accelerated economic growth and the onset of the demographic transition in England suggests that alternative mechanisms should be called on to better relate the initiation of the transition towards sustained economic growth to the apparently subsequent (and not simultaneous) demographic transition. However, a more recent study performed by Klemp & Weisdorf (2018) finds that parental fecundity positively affected the number of siblings and that children of parents with lower fecundity were more likely to become literate and employed in skilled and high-income professions. In summary, there is no systematic evidence that the child quantity-quality trade-off exist.

Moreover, there is also no consensus on the idea that the quantity-quality trade-off, if it exists, is the main driver of the demographic transition. Different social scientists have suggested that social norms, the large declines in mortality starting in the nineteenth century, and the reduced need for child labor are potential factors contributing to the demographic transition. Becker (1981) was the first to formalize a theory relating the quantity-quality trade-off of households to the rise in demand for human capital. But twenty years before, Becker (1960) advanced the much simpler argument that the decline in fertility was a by-product of the increase in income and the associated rise in the opportunity cost of raising children. This theory hinges on the supposition that individual preferences reflect an innate bias against child quantity beyond a certain level of income. This mechanism was recently modeled by Strulik *et al.* (2013). Before them, Jones (2001) used a simplified version of the same approach with a formal representation of the mortality rate, which allowed him to reproduce the fact that mortality rates decrease before fertility rates in countries experiencing a demographic transition.

Appendix A.2 The exaggerated role of human capital for take-off

One can cast some doubts on the central role that unified growth models accord to human capital of the general population in fostering economic take-off. Indeed, Mokyr (2011, p. 232), notices the weak accomplishment of schooling to build human capital that would be useful to reach a modern regime. According to him, even in the eighteenth-century, public education in Britain was primarily destined to educate gentlemen in the traditional sense of the word, that is, men without a well-defined occupation whose curricula consisted of the classics, languages, and other humanities. Besides, Mokyr (2011, p. 239) shows that adult literacy rates in Britain *circa* 1800 were equivalent to those of France and Belgium, and were even lower than those of the Netherlands. Moreover, Mokyr (2011, p. 239) asserts that even if Britain rapidly became richer than other countries thanks to its early economic take-off, its ability or willingness to educate its young did not appreciably improve during the first phase of the Industrial Revolution. At the end of the nineteenth century, school enrollment was indeed not higher in Britain relative to countries that experienced delayed takeoffs such as Prussia or France.

Finally, as an unequivocal criticism of the crucial role that most unified growth models assign to general human capital, Mokyr (2011, p. 240) adds that, at the time of the British economic take-off, human capital was surely not the result of an investment process in which the human capital rate of return on the margin would be equal to the interest rate. Rather, for Mokyr, it might well be that the causal direction was reversed and that many people decided for non-economic reasons to educate their children and then discovered that this education imparted economically useful capabilities. He then concludes that in any event, to the extent that the available data allow to make inferences, the notion that the Industrial Revolution depended a great deal on human capital as customarily defined is not sustained. On the contrary, Mokyr (2011, p. 233) asserts that the great English engineers of the Industrial Revolution learned their skills by being apprenticed to able masters, and otherwise were largely self-taught. The latter observation suggests that learning-by-doing used to play a prominent role in the pre-modern growth regime.

Appendix A.3 The misguided reasons for the omission of energy

Assigning a modest importance to energy in explaining growth is conventionally justified by its small share in national income. Indeed, the so-called ‘cost share theorem’ implies that, if the aggregate production function is homogeneous of degree one, the output elasticities of production factors equal their income allocation in total GDP. Consequently, GDP elasticities with respect to labor and capital are generally set to 0.7 and 0.3 according to their respective empirical shares of GDP, while energy is usually neglected because its cost usually represents around 5% of the national income. Even when it is considered as a production factor, the output elasticity of energy is set to 0.05, such that labor and capital remain the most important production factors (Denison, 1979). However, it can be argued that this ‘cost share theorem’ is fallacious for several reasons.

First of all, the cost share theorem results from a Lagrange optimization assuming that all perfectly competitive markets are at equilibrium for an economy only composed of small price-taking firms. Consequently, the cost share theorem is only true at the margin (for a fictive economy), such that output elasticities with respect to a given input follow the income cost share of those inputs only for small shocks. Moreover, by construction, GDP is allocated exclusively to capital and labor payments. Accordingly, energy expenditure is itself only made of capital and labor payments (plus temporary market powers).⁴³ But the fact that energy expenditures are relatively low in developed economies does not imply that energy *per se* is of no importance for economic growth. This fact was well illustrated by the first energy crisis of 1973, during which a 5% decrease in oil availability induced a 3% loss of GDP in the US, which is much higher than the mere 0.25% that the cost-share theorem predicted. In a review on how energy price shocks affect the US economy, Kilian (2008) asserts that rising energy prices cause both a reduction in aggregate demand, and a shift in consumer expenditures, which in turn create a ripple effect throughout the economy. The effects of energy price shocks on economic output are hence larger than the share of energy in income would suggest. This means, that the output elasticity of energy of 0.05 generally presupposed in standard macroeconomics is underestimated, whereas the output elasticities of capital and labor of 0.3 and 0.7 respectively, are overestimated.

Furthermore, energy expenditures used to account for up to 70% of national income in pre-industrial, low-growth economies. It is likely only the use of previously untapped concentrated, and consequently cheap, fossil fuels that this value gradually declined below 10% (Fizaine & Court, 2016). Kander *et al.* (2013, p. 7) assert that the decrease in the cost of energy, concurrent with an increase in supply, allowed vast reserves of capital to be employed, delivering other kinds of goods and services rather than covering only basic energetic needs as was the case during pre-modern times. Hence, the small cost share of energy in modern economies is not a sign of its worthlessness, but on the contrary, it might indicate the crucial importance that concentrated fossil energy has on modern economic growth.

Finally, the ground-breaking work of Kümmel & Lindenberg (2014) shows that, whenever *hard* technical constraints – corresponding to ‘limits to automation’ and ‘limits to capacity utilization’ – are taken into account, shadow prices add up to usual factor costs, implying that the cost share theorem simply no longer holds.⁴⁴ In summary, pure financial expenditure accounting downplays the role of energy because it does not take into account the interrelation between energy and specific technical developments that have been crucial to generating an expansion of many sectors of the economy. For instance, the design of modern transport systems and the associated suburban habitat have been wholly dependent on the Internal Combustion Engine (ICE) fueled by gasoline (and similarly, electric or gas-fired heating and cooling systems have made domestic and office life bearable in a variety of climates).

⁴³For instance, the price of gasoline is constituted of capital interest, labor payments, and various taxes that are required to extract and refine the crude oil provided free-of-charge by nature.

⁴⁴Besides, Ayres *et al.* (2013) argue that there are also some *soft* constraints – corresponding to social, financial, organizational, or legal restrictions – that determine additional limits to substitution between inputs over time.

Appendix A.4 Distinguishing several ‘kinds’ of energy

In order to understand the importance of energy for the economic process, it is crucial to distinguish between primary, final, and useful energy.⁴⁵ Primary energy is present in the environment in the form of natural stocks (coal, oil, gas, uranium) or flows (sun, water, wind, geothermal, wave and tide) that must be converted into secondary energy carriers in order to be usable. Such final energy vectors consist of heat flows, electricity, and solid, liquid or gaseous refined products. Finally, end-use devices allow the conversion of final carriers into useful energy in the form of motion (i.e., mechanical drive), heat, and light. Energy services (transport of passengers and goods, space heating, and illumination) are the outcomes of the interaction of useful energies with passive devices/infrastructures. Hence, all useful energy flows are measured in joules, whereas energy services take different units of measurement such as passenger.km or tonne.km for transport, and lumen for illumination. Because technical change effects each conversion step of energy systems with different magnitudes, the prices of primary, final, and useful energies do not evolve similarly. An example of such difference is given in Figure 8, where the average price of primary energy is compared to the average price of energy services in Great Britain from 1700 to 2000. As Fouquet (2011) argues, focusing on the price of primary energy rather than the price of useful energy or energy services can lead to flawed reasoning because the former ignore major technical improvements developed to provide the latter.

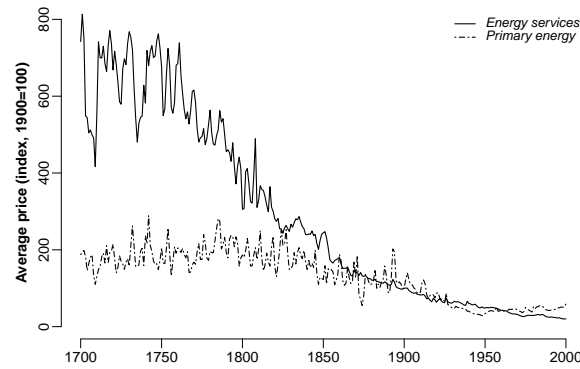


Fig. 8 Average prices of primary energy (*dashed line*) and energy services (*solid line*) in the United Kingdom, 1700–2008 Data source: Fouquet (2011).

Appendix B Aggregation

In this section, we derive the aggregation for final input sectors, $k \in \mathcal{K}$. First, remember that the production technology equation writes

$$Y_{k,t} = A_{k,t} \left[\int_0^1 q_{k,i,t}^{1-\alpha_k} x_{k,i,t}^{\alpha_k} di \right] H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}.$$

⁴⁵As repeatedly stressed by scholars such as Ayres & Warr (2009) and Kümmel (2011), what is commonly called energy in economic studies and models is in fact *exergy*. Exergy is the valuable part, or more formally the potentially useful part, of energy (precisely, it is the maximum work that can be done by a system that reversibly approaching its thermodynamic equilibrium). As required by the first law of thermodynamics, energy is conserved in the economic process. On the other hand, the second law of thermodynamics stipulates that exergy is destroyed through the functioning of the economic system, since it is composed of multiple irreversible processes that imply entropy creation. Energy enters the economy as a high quality (high exergy content) input in the forms of concentrated solar energy (biomass and water/wind flows), geothermal and tidal potential, fossils fuels, and nuclear energy. These energy forms are ultimately dissipated into a lower-quality (lower exergy content) heat output that potentially contains zero exergy (and thus zero ability to generate useful work) if its temperature is the same as the broader environment. Hence, it is the exergy content of energy that constitutes a production factor used up in the economic process and not energy *per se*. In the remainder of this article we will stick to the familiar term of energy, even if, strictly speaking, we refer to exergy.

For each machine line $i \in [0, 1]$, the optimal demand schedule is given by the first order condition of Eq. (16) with $p_{k,i,t}^x$ the price of the machine. As discussed in [Subsection 3.2.3](#), this price depends on the status under which the machine is supplied. Whenever R&D was successful in the beginning of the period for the machine line, the machine is supplied under monopolistic competition at $p_{k,t}^m = (r_t + \delta) / \alpha_k$ and the corresponding optimal demand schedule equation then writes

$$x_{k,i,t}^m = \left[\frac{p_{k,t} \alpha_k^2 A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}}{r_t + \delta} \right]^{\frac{1}{1-\alpha_k}} q_{k,i,t}.$$

In the opposite case, the machine is supplied competitively and its price set at the marginal production cost, that is $p_{k,i,t}^c = r_t + \delta$. The corresponding optimal demand schedule equation then writes

$$x_{k,i,t}^c = \left[\frac{p_{k,t} \alpha_k A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}}{r_t + \delta} \right]^{\frac{1}{1-\alpha_k}} q_{k,i,t}.$$

As demonstrated in [Subsection 3.3.3](#), R&D activity is uniformly distributed over machines lines with a probability of success $\lambda_{k,t}$. Thus, at the beginning of each period, a fraction $\lambda_{k,t}$ of firms experience a successful innovation, gain one additional rung of size q_k in their corresponding quality ladder, and are produced under monopolistic competition. For the remaining firms, the quality remains constant and the production is perfectly competitive. Discriminating firms accordingly and introducing the average quality index $Q_{k,t}$ (defined at the beginning of the period) yields

$$\begin{aligned} Y_{k,t} &= A_{k,t} \left[\int_0^{\lambda_{k,t}} q_{k,i,t}^{1-\alpha_k} (x_{k,i,t}^m)^{\alpha_k} di + \int_{\lambda_{k,t}}^1 q_{k,i,t}^{1-\alpha_k} (x_{k,i,t}^c)^{\alpha_k} di \right] H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}, \\ &= \left[A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k} \right]^{\frac{1}{1-\alpha_k}} \left[\frac{p_{k,t} \alpha_k}{r_t + \delta} \right]^{\frac{\alpha_k}{1-\alpha_k}} \left[q_k \alpha_k^{\frac{\alpha_k}{1-\alpha_k}} \int_0^{\lambda_{k,t}} q_{k,i,t-1} di + \int_{\lambda_{k,t}}^1 q_{k,i,t-1} di \right], \\ &= \left[A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k} \right]^{\frac{1}{1-\alpha_k}} \left[\frac{p_{k,t} \alpha_k}{r_t + \delta} \right]^{\frac{\alpha_k}{1-\alpha_k}} \left[\alpha_k^{\frac{\alpha_k}{1-\alpha_k}} q_k \lambda_{k,t} + (1 - \lambda_{k,t}) \right] Q_{k,t}. \end{aligned} \quad (50)$$

Turning now to the stock of capital ultimately dedicated to the raw capital demand in each final input sector, defined by $K_{k,t} = \int_0^1 x_{k,i,t} di$, one can distinguish between machines that are produced under monopolistic and perfect competition to write

$$K_{k,t} = \left[\frac{p_{k,t} \alpha_k A_{k,t} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}}{r_t + \delta} \right]^{\frac{1}{1-\alpha_k}} \left[\alpha_k^{\frac{1}{1-\alpha_k}} q_k \lambda_{k,t} + (1 - \lambda_{k,t}) \right] Q_{k,t-1}, \quad (51)$$

Inserting Eq (51) into Eq (50) and rearranging gives a final expression for the aggregate production function in the final good sector

$$Y_{k,t} = A_{k,t} Q_{k,t}^{1-\alpha_k} \frac{\alpha_k^{\frac{\alpha_k}{1-\alpha_k}} q_k \lambda_{k,t} + (1 - \lambda_{k,t})}{\left[\alpha_k^{\frac{1}{1-\alpha_k}} q_k \lambda_{k,t} + (1 - \lambda_{k,t}) \right]^{\alpha_k}} K_{k,t}^{\alpha_k} H_{k,t}^{\beta_k} E_{k,t}^{\gamma_k}.$$

Appendix C Endogenous direction of technical change with the original specification

We maintain **Assumptions 1, 2, 3, 4, 5** and **6**. At first, using **Proposition 3**, we can compute the direction of technical change as

$$\frac{g_{Q_{r,t}}}{g_{Q_{e,t}}} = \frac{q_r - 1}{q_e - 1} \frac{\bar{\lambda}_{r,t} - \frac{\phi_k w_t}{\alpha(1-\alpha)G_{v,t} \left[\frac{\alpha^2}{r_t + \delta} \right]^{\frac{\alpha}{1-\alpha}} \left[p_{r,t} H_{r,t}^\beta E_{r,t}^\gamma \right]^{\frac{1}{1-\alpha}}}}{\bar{\lambda}_{e,t} - \frac{\phi_k w_t}{\alpha(1-\alpha)G_{v,t} \left[\frac{\alpha^2}{r_t + \delta} \right]^{\frac{\alpha}{1-\alpha}} \left[p_{e,t} H_{e,t}^\beta E_{e,t}^\gamma \right]^{\frac{1}{1-\alpha}}}}. \quad (52)$$

Recalling that the relative intensity in R&D is driven by: (i) an *efficiency effect*, (ii) a *price effect*, and (iii) a *market-size effect*. As **Propositions 4** and **5** still hold, price and market size effects play in opposite directions with regards to relative R&D knowledge stocks and extraction costs. We can then solve for the equilibrium wage rate, w_t , combining equilibrium input production levels and the first order conditions with regards to labor. Substituting in Eq. (52) then yields an explicit expression for directed technical change, denoted $DTC(Q_{e,t}, Q_{r,t}, \Psi_{e,t}, \Psi_{r,t})$, that only depends on R&D knowledge and extraction costs

$$DTC(Q_{e,t}, Q_{r,t}, \Psi_{e,t}, \Psi_{r,t}) = \frac{q_r - 1}{q_e - 1} \frac{\bar{\lambda}_r - \frac{\beta \phi_r}{\alpha(1-\alpha)G_{v,t}} Q_{r,t} \left[1 + \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{-(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{\gamma(\sigma-1)} \right]}{\bar{\lambda}_e - \frac{\beta \phi_e}{\alpha(1-\alpha)G_{v,t}} Q_{e,t} \left[1 + \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)} \right]}. \quad (53)$$

We can now state a result similar to **Theorem 2** regarding the direction of technical change.

Theorem 3 (Direction of technical change) *Under **Assumptions 1, 2, 3, 4, 5** and **6**, and assuming that final inputs are sufficiently substitutable (i.e., that $(1-\alpha)(\sigma-1) \geq 1$), rising relative extraction cost favors R&D in the opposite sector. Moreover, whenever sectors are technically comparable and inputs sufficiently substitutable, that is $Q_{r,t} = O(Q_{e,t})$, then R&D knowledge favors innovation in the most advanced sector.*

Proof. We consider the partial derivatives of $DTC(\cdot)$ with regards to extraction costs and R&D knowledge. First, one can show that

$$\begin{aligned} \frac{\partial DTC}{\partial \Psi_{r,t}} &= -\frac{1}{\lambda_{e,t}^2} \frac{\beta \gamma (\sigma - 1)}{\alpha(1-\alpha)G_{v,t}} \left[\phi_r (q_r - 1) \frac{\lambda_{e,t} Q_{r,t}}{\Psi_{e,t}} \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{-(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{\gamma(\sigma-1)-1} \right. \\ &\quad \left. + \phi_e (q_e - 1) \frac{\lambda_{r,t} Q_{e,t}}{\Psi_{e,t}} \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)-1} \right] < 0. \end{aligned}$$

The sign of this partial derivative is always negative, and it can be similarly shown that $\partial DTC / \partial \Psi_{e,t} > 0$. Thus, extraction cost variations unambiguously foster innovation in opposite sectors. Turning now to R&D knowledge, one can show that

$$\begin{aligned} \frac{\partial DTC}{\partial Q_{e,t}} &= -\frac{DTC(\cdot) \beta \phi_e (q_e - 1)}{\lambda_{e,t} \alpha(1-\alpha)G_{v,t}} \left[\left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{(1-\alpha)(\sigma-1)} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{-\gamma(\sigma-1)} [(1-\alpha)(\sigma-1) - 1] \right. \\ &\quad \left. + \frac{(q_r - 1) \phi_r}{(q_e - 1) \phi_e} \frac{1}{DTC(\cdot)} (1-\alpha)(\sigma-1) \left[\frac{Q_{r,t}}{Q_{e,t}} \right]^{-(1-\alpha)(\sigma-1)+1} \left[\frac{\Psi_{r,t}}{\Psi_{e,t}} \right]^{\gamma(\sigma-1)} - 1 \right]. \end{aligned}$$

A similar expression can be obtained for $\partial DTC / \partial Q_{r,t}$. The sign of this expression remains ambiguous. The first term in the sum between brackets is positive whenever inputs are sufficiently substitutable, that is $(1-\alpha)(\sigma-1) \geq 1$. The second term is positive if the condition on parameters just mentioned is met, and if sectors are quite similar for technologies in the sense that

$Q_{r,t} = O(Q_{e,t})$. In this case, there is again a threshold for substitutability that ensures that the second bracketed term is positive. Note that variations in the relative extraction cost play an intuitive role in these partial derivatives. Moreover, this condition could be numerically evaluated to assess what the direction of R&D knowledge incentives in technical change would be at a given point. \square

Three comments about the previous theorem are necessary: (i) there is no ambiguity in the impact of relative extraction cost variations on the direction of R&D (towards the cheapest sector), (ii) the possible reversal in the direction of R&D knowledge incentives might be puzzling at first, but shall be put in perspective with the specification of the R&D success probability, that admits decreasing returns (research becoming increasingly difficult whenever a threshold value is approached), and (iii) the mechanism proposed is compatible with our modelling of both the energy shift and the economic take off. On this last point, our narrative goes as follows. R&D knowledge stock is comparable at the beginning of the industrial revolution because modern R&D was only beginning. It is precisely the extraction cost evolution that directs R&D more intensively towards the exhaustible sector as the renewable resource sees a shortage while a large stock of exhaustible resource remains untapped (from this view, one might consider the exhaustible extraction cost as constant at that period). The resulting fall of the $DTC(\cdot)$ ratio then fosters R&D in the exhaustible sector (positive externality of the corresponding knowledge stock). It is only after a sustained period of biased innovation, while the R&D success probability reaches its upper limit in the exhaustible sector, that the reversal in the direction of knowledge incentives is likely to happen. At that moment, the renewable sector catches up.

Appendix D Best-fit calibration process

In this section, we provide some technical information about our numerical simulation process and its calibration on historical data.

Appendix D.1 Simulation process

We use the formal calculation software *Maple* (2017 version) to perform the numerical simulation of our model. As we cannot provide a closed form solution for the general version of our model, we use optimization commands (*Optimization* package) to solve for prices and quantities, accounting for market clearing and physical constraint conditions, as well as optimal (and closed-form) responses of the representative household. Our Monte-Carlo simulations are also performed on this software using the *Statistics* and *RandomTools* packages. For our numerical results, we performed a total of 10,000 runs on each scenario analyzed in [Section 5](#). All the presented figures are realized with the *RStudio* software.

Provided that fertility, savings, and educational choices are endogenous in our model according to [Problem HH](#) (and thus depend on the clearing wage rate), we allow for two shimming periods (1660 and 1680) before starting the numerical simulations. These shimming periods ensure that household's decisions, population's allocation (between young and retired adults) and capital accumulation (resulting from previous savings) consistently adjust prior to the formal start of the simulation process. We preclude any knowledge accumulation during this shimming time, so knowledge stocks remain consistent with their normalization to unity in the first simulation period. It is worth mentioning that the current stock of capital only depends on savings from the previous period according to Eq. (7), therefore, the initial value chosen for the stock of capital does not matter for capital accumulation after these shimming periods.

Appendix D.2 Calibration process

Due to the high dimensionality of our calibration problem and the complexity of the general equilibrium interactions in the full version of our model, we adopt a three-step calibration process:

- *Preferences calibration problem*: for given series of wage rate and total applied knowledge stock, the closed form solution of [Problem HH](#) allows us to obtain the partial equilibrium

solution of the full model for population and human capital per capita. In a first calibration step, we thus optimize for household preferences and the initial population stock to fit these two time series, that is, we set $\{\chi, \rho, \sigma, \tau, A_E, N_0\}$ to minimize the squared error between the simulated model and historical data.⁴⁶

- *Deterministic calibration problem*: for a given series of GPT arrivals (determined according to historical data), solving numerically for **Problem GE** allows us to obtain a deterministic solution of our model. In a second calibration step, we thus optimize for all parameters except GPT related parameters to fit our historical times series of interest. That is, we set $\{Y_0, \Omega, \omega_H, \omega_G, \phi, q_c, q_d, \bar{\lambda}, \bar{Y}_r, \bar{Y}_e, \psi_{\mathcal{R},r}, \psi_{\mathcal{R},e}, \psi_Q\}$ to minimize various squared error metrics between the simulated model and historical data on GDP, population, energy consumption flows and human capital per capita.⁴⁷ For each assessed design, we also optimize for A_E to minimize the squared error in regards to population and human capital per capita (with a 10 times lower weight for the second) to ease the coupling between the *preferences* and *deterministic calibration problems*.
- *Stochastic GPT calibration problem*: for a given calibration design (chosen as the best candidate from the *deterministic calibration problem*), repeatedly solving for **Problem GE** allows us to obtain a Monte-Carlo solution for our full model. In practice, due to computational constraints, we performed 100 runs for each calibration design assessed in this step. This number is gauged to be a good approximation for Monte-Carlo convergence. We tested the designs that score best on a higher number of steps (1,000 for the calibration process, 10,000 for the numerical results). In this last calibration step, we thus optimize for GTP related parameters only to fit the historical sequence of GPT arrival, that is, we set $\{\mu_0, \bar{G}, \bar{\zeta}, \bar{\xi}\}$ to minimize the squared errors between the simulated and historical GDP frequency.⁴⁸ In this calibration step, given that we assess the relationship between total applied knowledge accumulation, Q_t , and GPT's frequency and diffusion, we only performed this step for the British calibration and maintain the obtained parameters for the western European and eastern Asian cases. As illustrated in additional **Figure 9**, we observe that the frequency of GPTs in the *baseline* calibration of Western Europe closely follows the one obtained for Great Britain, while it is lower for Eastern Asia, as one would logically expect.

If necessary (i.e., whenever a significant difference is observed), these three steps are cyclically repeated until the calibration converged within an increasing series of compact sets for LHS procedures (at each repeated step, we removed ranges of parameters that yielded the poorest results based on the best 2 to 3 runs on each dimension of interest).

⁴⁶We use a direct method algorithm from the *DirectSearch* package to cover the whole compact calibration set. We give 10 times greater weight to population than to human capital per capita (squared-error) scores, and use a Hodrick-Prescott (HP) filter to smooth historical time series so as to downplay the extreme values of the Great Depression and World War (I and II). N_0 denotes the population stock that will be applied to the shimming periods. To account for this slight approximation (the population however being initially low) we perform the *preference calibration problem* in the 1660-2000 period to account for the two shimming periods.

⁴⁷We use a Latin Hypercube Sampling (LHS) with a *maximin* criteria to test different sets with up to 3200 samples from a compact calibration set. The sampling is performed using the *lhs* package used on the *RStudio* software. We computed various squared-errors metrics between the simulated model and historical data, that are smoothed with a HP-filter to downplay the extreme values of the Great Depression and World War (I and II). More precisely, each variable in $\{E_r, E_e, E_r/E_e, \mathcal{R}_r, \mathcal{R}_e, Y, h, P, Y/P\}$ is mapped to a constructed and normalized historical time series, from which we compute a squared-error and weighted squared-error (allowing for a 10 times greater weight for periods prior 1900). For the British case, we then pick designs that perform well favoring (i) the weighted Y score (fit on GDP during the economic take-off), and (ii) the sum of squared-error differences for the score of variables in $\{E_r/E_e, \mathcal{R}_r, \mathcal{R}_e, Y, h, P, Y/P\}$ (fit on the main time-series of interest on the whole simulation horizon). For both Western Europe and Eastern Asia calibrations, we pick designs that perform well according to the weighed sum of squared-error differences for the score of variables in $\{E_e, E_r, Y, h, P, Y/P\}$. In the Western European case, we emphasize the pre-WWII period, provided the industrial take-off and transition towards fossil fuels occurred in the XIXth century. For this reason we ascribe 10 times more weight to the squared-error differences prior 1940. In the Eastern Asian case, we emphasize the post-1900 period provided the industrial take-off and transition towards fossil energy occurred in the XXth century. This is why we ascribe 10 times more weight to the squared-error differences after 1900.

⁴⁸We use a LHS procedure with a *maximin* criteria to test different sets with up to 1600 samples from a compact calibration set. The sampling is performed using the *lhs* package used on the *RStudio* software.

Appendix E Calibration in the baseline run for Great Britain

Table 5 and 6 respectively present the initial variables and parameters (rounded to the nearest 10^{-3}) obtained for the baseline calibration of the model to the British data. If no index $k \in \{r, e\}$ is provided (e.g., A), the initial condition or parameter concerns both sectors.

Table 5 Initial values of variables for the *baseline* calibration of the model on British data.

Variable	A_0	G_0	h_0	K_0	N_0	Q_0	$\mathcal{R}_{r,0}$	$\mathcal{R}_{e,0}$
Initial value	1	1	1	0.002	0.094	1	1.073	463.097

Table 6 Parameters for the *baseline* calibration of the model on British data.

Parameter	α	β	γ	ϑ_k	Y_0	Ω	ω_H	ω_G	ϕ
Value	1/3	1/2	1/6	1	3.3	0.012	0.027	0.445	0.183
Parameter	q	$\bar{\lambda}$	$\bar{\Psi}_r$	$\bar{\Psi}_e$	$\psi_{\mathcal{R},r}$	$\psi_{\mathcal{R},e}$	ψ_Q	η	χ
Value	1.26	0.377	0.016	0.106	-4.506	-7.296	-7.993	0.202	0.3
Parameter	ρ	τ	A_E	\bar{h}	μ_0	\bar{G}	ζ	ξ	
Value	0.048	0.102	51.74	1	0.1	1.69	0.179	0.244	

Appendix F Calibration in the baseline run for Western Europe

Table 7 and 8 respectively present the initial variables and parameters (rounded to the nearest 10^{-3}) obtained for the baseline calibration of the model for Western Europe. If no index $k \in \{r, e\}$ is provided (e.g., A), the initial condition or parameter concerns both sectors.

Table 7 Initial values of variables for the *baseline WE* calibration

Variable	A_0	G_0	h_0	K_0	N_0	Q_0	$\mathcal{R}_{r,0}$	$\mathcal{R}_{e,0}$
Initial value	1	1	1	0.04	0.018	1	5.76	227.19

Table 8 Parameters for the *baseline WE* calibration

Parameter	α	β	γ	ϑ_k	Y_0	Ω	ω_H	ω_G	ϕ
Value	1/3	1/2	1/6	1	3.145	0.014	0.059	0.492	0.179
Parameter	q_r	q_e	$\bar{\lambda}$	$\bar{\Psi}_r$	$\bar{\Psi}_e$	$\psi_{\mathcal{R},r}$	$\psi_{\mathcal{R},e}$	ψ_Q	η
Value	1.361	1.371	0.686	0.03	0.192	-8	-4.209	-4.14	0.479
Parameter	χ	ρ	τ	A_E	μ_0	\bar{G}	ζ	ξ	
Value	0.37	0.037	0.224	73.696	0.1	1.69	0.179	0.244	

Appendix G Calibration in the baseline run for Eastern Asia

Table 9 and 10 respectively present the initial variables and parameters (rounded to the nearest 10^{-3}) obtained for the baseline calibration of the model to the Eastern Asia. If no index $k \in \{r, e\}$ is provided (e.g., A), the initial condition or parameter concerns both sectors.

Table 9 Initial values of variables for the *baseline EA* calibration

Variable	A_0	G_0	h_0	K_0	N_0	Q_0	$\mathcal{R}_{r,0}$	$\mathcal{R}_{e,0}$
Initial value	1	1	1	0.005	0.022	1	20.43	72.57

Table 10 Parameters for the *baseline EA* calibration

Parameter Value	α 1/3	β 1/2	γ 1/6	ϑ_k 1	Y_0 2.294	Ω 0.001	ω_H 0.044	ω_G 0.335	ϕ 0.184
Parameter Value	q_r 1.179	q_e 1.55	$\bar{\lambda}$ 0.789	$\bar{\Psi}_r$ 0.01	$\bar{\Psi}_e$ 0.222	$\psi_{\mathcal{R},r}$ -5.221	$\psi_{\mathcal{R},e}$ -3.182	ψ_Q -5.278	η 0.431
Parameter Value	χ 0.358	ρ 0.052	τ 0.179	A_E 58.626	μ_0 0.1	\bar{G} 1.69	ζ 0.179	ξ 0.244	

Appendix H Additional figure: GPT frequency in baseline calibrations

GPTs are calibrated on the British case. The values obtained for the parameters $\{\mu_0, \bar{G}, \zeta, \xi\}$ are assumed to reflect the relationship between total applied knowledge, Q_t , and the arrival and diffusion time of GPTs. These values are consequently maintained for both Western Europe and Eastern Asia. In [Figure 9](#), one can observe that GPT frequency closely followed the British case for Western Europe, but are consistently lower for Eastern Asia, as one would logically expect.

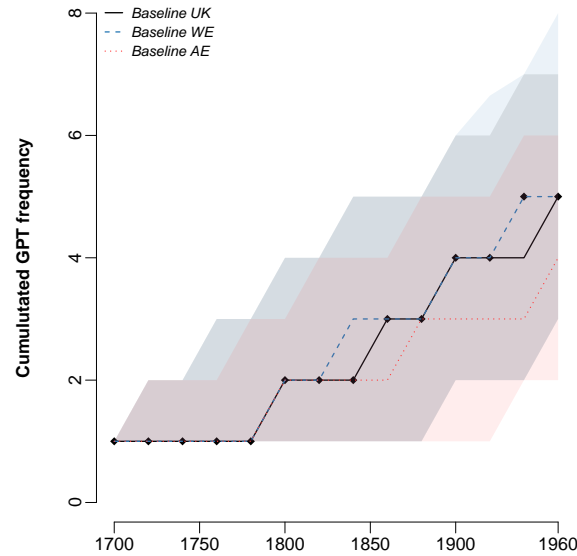


Fig. 9 GPT frequency in the *baseline* calibration (10,000 simulations) for Great Britain (*black*), Western Europe (*blue*) and Eastern Asia (*red*), 1700-1960 (*solid, dashed and dotted lines are medians, color share are 90% probability intervals, diamonds are historical data*). Data source: [Lipsey et al. \(2005, p. 132\)](#)

References

- Acemoglu, Daron. 2002. Directed Technical Change. *Review of Economic Studies*, **69**(4), 781–809.
- Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth*. Princeton, NJ: Princeton University Press.

- Acemoglu, Daron, Aghion, Philippe, Bursztyn, Leonardo, & Hemous, David. 2012. The Environment and Directed Technical Change. *American Economic Review*, **102**(1), 131–166.
- Aghion, Philippe, & Howitt, Peter. 1992. A Model of Growth Through Creative Destruction. *Econometrica*, **60**(2), 323–351.
- Aghion, Philippe, & Howitt, Peter. 1998. *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- Aghion, Philippe, & Howitt, Peter. 2009. *The Economics of Growth*. Cambridge, MA: MIT Press.
- Allen, R. C. 2011. Why the industrial revolution was British: commerce, induced invention, and the scientific revolution. *Economic History Review*, **64**(2), 357–384.
- Allen, Robert C. 2009. *The British Industrial Revolution in Global Perspective*. Cambridge, UK: Cambridge University Press.
- Allen, Robert C. 2015. The high wage economy and the industrial revolution: a restatement. *Economic History Review*, **68**(1), 1–22.
- Allen, Robert C. 2019. Real wages once more: a response to Judy Stephenson. *The Economic History Review*, **in press**.
- Allen, Robert C., Bassino, Jean-Pascal, Ma, Debin, Moll-Murata, Christine, & van Zanden, Jan Luiten. 2011. Wages, prices, and living standards in China, 1738-1925: in comparison with Europe, Japan, and India. *Economic History Review*, **64**(2), 8–38.
- Angrist, Joshua, Lavy, Victor, & Schlosser, Analia. 2010. Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, **28**(4), 773–824.
- Ayres, Robert U., & Warr, Benjamin. 2009. *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*. Cheltenham, UK: Edward Elgar Publishing.
- Ayres, Robert U., van den Bergh, Jeroen C.J.M., Lindenberger, Dietmar, & Warr, Benjamin. 2013. The underestimated contribution of energy to economic growth. *Structural Change and Economic Dynamics*, **27**, 79–88.
- Becker, Gary Stanley. 1960. An Economic Analysis of Fertility. In: Becker, Gary Stanley (ed), *Demographic and Economic Change in Developed Countries*. Princeton, NJ: Princeton University Press.
- Becker, Gary Stanley. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- Becker, Sascha O., Cinnirella, Francesco, & Woessmann, Ludger. 2010. The trade-off between fertility and education: evidence from before the demographic transition. *Journal of Economic Growth*, **15**, 177–204.
- Berg, Maxine. 1994. *The Age of Manufactures, 1700–1820: Industry, Innovation, and Work in Britain*. London, UK: Routledge.
- Black, Sandra E., Devereux, Paul J., & Salvanes, Kjell. G. 2005. The More the Merrier ? The Effect of Family Size and Birth Order on Children’s Education. *Quarterly Journal of Economics*, **120**(2), 669–700.
- Bolt, Jutta, Inklaar, Robert, de Jong, Herman, & van Zanden, Jan Luiten. 2018. *Maddison Project Database 2018*.
- Boucekkine, Raouf, de la Croix, David, & Peeters, Dominique. 2007. Early literacy achievements, population density, and the transition to modern growth. *Journal of the European Economic Association*, **5**(1), 183–226.
- British Petroleum. 2015. *Statistical Review of World Energy 2015*.
- Broadberry, Stephen, & Wallis, John Joseph. 2017. Growing, shrinking, and long run economic performance: historical perspectives on economic development. *NBER Working Paper*, **No. 23343**.
- Broadberry, Stephen, Campbell, Bruce M. S., Klein, Alexander, Overton, Mark, & Leeuwen, Bas van. 2015. *British Economic Growth, 1270–1870*. Cambridge, UK: Cambridge University Press.
- Broadberry, Stephen, Guan, Hanhui, & Li, David Daokui. 2018. China, Europe, and the Great Divergence: a study in historical national accounting, 980–1850. *Journal of Economic History*, **78**(4), 1–46.
- Cáceres-Delphiano, Julio. 2006. The Impacts of Family Size on Investment in Child Quality. *Journal of Human Resources*, **41**(4), 738–754.
- Cervellati, Matteo, & Sunde, Uwe. 2005. Human Capital Formation, Life Expectancy, and the Process of Development. *American Economic Review*, **95**(5), 1653–1672.
- Clark, G., & Jacks, D. 2007. Coal and the Industrial Revolution, 1700-1869. *European Review of Economic History*, **11**(1), 39–72.
- Clark, Gregory, & Cummins, Neil. 2016. The Child Quality-Quantity Tradeoff, England, 1780-1880: A

Fundamental Component of the Economic Theory of Growth is Missing. *CEPR Discussion Papers*, **No. DP1123**.

Court, Victor, Jouvet, Pierre-André, & Lantz, Frédéric. 2018. Long-term endogenous economic growth and energy transitions. *Energy Journal*, **39**(1), 29–57.

Crosby, Alfred W. 2007. *Children of the Sun: A History of Humanity's Unappeasable Appetite for Energy*. New York, NY: W.W. Norton & Company.

de La Croix, David, & Michel, Philippe. 2002. *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*. Cambridge, UK: Cambridge University Press.

de Pleijt, Alexandra, Nuvolari, Alessandro, & Weisdorf, Jacob. 2019. Human Capital Formation During the First Industrial Revolution: Evidence from the use of Steam Engines. *Journal of the European Economic Association*, 03.

de Vries, Jan. 1994. The Industrial Revolution and the Industrious Revolution. *Journal of Economic History*, **54**(2), 249–270.

Debeir, Jean-Claude, Deleage, Jean-Paul, & Hemery, Daniel. 1991. *In the Servitude of Power: Energy and Civilisation Through the Ages*. London, UK: Zed Books.

Denison, Edward Fulton. 1979. *Accounting for slower economic growth : the United States in the 1970s*. Washington, DC: The Brookings Institution.

Desmet, Klaus, & Parente, Stephen L. 2012. The evolution of markets and the revolution of industry: a unified theory of growth. *Journal of Economic Growth*, **17**(3), 205–234.

Diamond, Jared M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. New York, NY: W.W. Norton & Company.

Doepke, Matthias. 2004. Accounting for Fertility Decline During the Transition to Growth. *Journal of Economic Growth*, **9**(3), 347–383.

Fizaïne, Florian, & Court, Victor. 2016. Energy expenditure, economic growth, and the minimum EROI of society. *Energy Policy*, **95**.

Fouquet, Roger. 2008. *Heat, Power and Light : Revolutions in Energy Services*. Cheltenham, UK: Edward Elgar Publishing.

Fouquet, Roger. 2011. Divergences in Long Run Trends in the Prices of Energy and Energy Services. *Review of Environmental Economics and Policy*, **5**(2), 196–218.

Fouquet, Roger. 2014. Long run demand for energy services: income and price elasticities over 200 years. *Review of Environmental Economics and Policy*, **8**(2), 186–207.

Fouquet, Roger, & Broadberry, Stephen. 2015. Seven Centuries of European Economic Growth and Decline. *Journal of Economic Perspectives*, **29**(4), 227–244.

Fröling, Maria. 2011. Energy use, population and growth, 1800–1970. *Journal of Population Economics*, **24**(3), 1133–1163.

Galindev, Ragchaasuren. 2011. Leisure goods, education attainment and fertility choice. *Journal of Economic Growth*, **16**(2), 157–181.

Galor, Oded. 2005. From Stagnation to Growth: Unified Growth Theory. *Chap. 4 of: Aghion, P., & Durlauf, S.N. (eds), Handbook of Economic Growth, Vol. 1A*, vol. 1A. Amsterdam, NL: Elsevier B.V.

Galor, Oded, & Michalopoulos, Stelios. 2012. Evolution and the growth process: Natural selection of entrepreneurial traits. *Journal of Economic Theory*, **147**, 759–780.

Galor, Oded, & Moav, Omer. 2002. Natural Selection and the Origin of Economic Growth. *Quarterly Journal of Economics*, **117**(4), 1133–1191.

Galor, Oded, & Weil, David N. 2000. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review*, **90**(4), 806–828.

Gentvilaite, Ruta, Kander, Astrid, & Warde, Paul. 2015. The Role of Energy Quality in Shaping Long-Term Energy Intensity in Europe. *Energies*, **8**(1), 133–153.

Goldstone, Jack A. 2002. Efflorescences and Economic Growth in World History: Rethinking the "Rise of the West" and the Industrial Revolution. *Journal of World History*, **13**(2), 323–389.

Goldstone, Jack A. 2009. *Why Europe? The Rise of the West in World History, 1500-1850*. Boston, MA: McGraw-Hill Higher Education.

- Hansen, Gary D., & Prescott, Edward C. 2002. Malthus to Solow. *American Economic Review*, **92**, 1205–1217.
- Hayami, Akira. 2015. *Japan's Industrious Revolution: Economic and Social Transformations in the Early Modern Period*. Berlin, DE: Springer.
- Hazan, Moshe, & Zoabi, Hosny. 2006. Does longevity cause growth? A theoretical critique. *Journal of Economic Growth*, **11**, 363–376.
- Humphries, Jane. 2013. The lure of aggregates and the pitfalls of the patriarchal perspective: a critique of the high wage economy interpretation of the British industrial revolution. *Economic History Review*, **66**(3), 693–714.
- IIASA. 2012. *Global Energy Assessment - Toward a Sustainable Future*. Tech. rept. International Institute for Applied Systems Analysis, Schloss Laxenburg, AT.
- IPCC. 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- Jacob, Margaret C. 2014. *The First Knowledge Economy: Human Capital and the European Economy, 1750-1850*. Cambridge, UK: Cambridge University Press.
- Jones, Charles I. 1995. R&D-Based Models of Economic Growth. *Journal of Political Economy*, **103**(4), 759–784.
- Jones, Charles I. 2001. Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run. *Advances in Macroeconomics*, **1**(2), 1–43.
- Kander, Astrid, & Stern, David I. 2014. Economic growth and the transition from traditional to modern energy in Sweden. *Energy Economics*, **46**(11), 56–65.
- Kander, Astrid, Malanima, Paolo, & Warde, Paul. 2013. *Power to the People: Energy in Europe Over the Last Five Centuries*. Princeton, NJ: Princeton University Press.
- Kilian, Lutz. 2008. The Economic Effects of Energy Price Shocks. *Journal of Economic Literature*, **46**(4), 871–909.
- Klemp, Marc, & Weisdorf, Jacob. 2018. Fecundity, Fertility and the Formation of Human Capital. *Economic Journal*, **6**.
- Kortum, Samuel. 1993. Equilibrium R and D and the Patent–R and D Ratio: U.S. Evidence. *American Economic Review*, **83**(2), 450–457.
- Kümmel, Reiner. 2011. *The Second Law of Economics: Energy, Entropy, and the Origins of Wealth*. New York, NY: Springer.
- Kümmel, Reiner, & Lindenberger, Dietmar. 2014. How energy conversion drives economic growth far from the equilibrium of neoclassical economics. *New Journal of Physics*, **16**(12), 125008.
- Lagerlöf, Nils-Petter. 2003. Gender Equality and Long-Run Growth. *Journal of Economic Growth*, **8**, 403–426.
- Lee, Jong-Wha, & Lee, Hanol. 2016. Human capital in the long run. *Journal of Development Economics*, **122**(9), 147–169.
- Li, Hongbin, Zhang, Junsen, & Zhu, Yi. 2008. The Quantity-Quality Trade-off of Children in a Developing Country: Identification Using Chinese Twins. *Demography*, **45**(1), 223–243.
- Lipsey, Richard G., Carlaw, Kenneth., & Bekar, Clifford. 2005. *Economic Transformations: General Purpose Technologies and Long-Term Economic Growth*. Oxford, UK: Oxford University Press.
- Maddison, Angus. 2007. *Chinese Economic Performance in the Long Run – Second edition, revised, and updated: 960-2030 AD*. Paris, FR: OECD Development Centre Studies.
- Malanima, P. 2011. The long decline of a leading economy: GDP in central and northern Italy, 1300-1913. *European Review of Economic History*, **15**(2), 169–219.
- Malanima, Paolo. 2009. *Pre-modern European Economy : One Thousand Years (10th-19th Centuries)*. Leiden, NL: Brill.
- Malanima, Paolo. 2016. Energy consumption in England and Italy, 1560-1913. Two pathways toward energy transition. *Economic History Review*, **69**(1), 78–103.
- Mokyr, Joel. 1990. *The Lever of Riches: Technological Creativity and Economic Progress*. New York, NY: Oxford University Press.

- Mokyr, Joel. 2011. *The Enlightened Economy: Britain and the Industrial Revolution, 1700-1850*. London, UK: Penguin Books.
- Mokyr, Joel. 2017. *A Culture of Growth: The Origins of the Modern Economy*. Princeton: Princeton University Press.
- Morris, Ian. 2010. *Why the West Rules For Now: The Patterns of History, and What They Reveal About the Future*. New York, NY: Farrar, Straus and Giroux.
- Morris, Max D., & Mitchell, Toby J. 1995. Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference*, **43**(3), 381–402.
- Mourmouras, Alex, & Rangazas, Peter. 2009. Reconciling Kuznets and Habbakuk in a unified growth theory. *Journal of Economic Growth*, **14**(2), 149–181.
- Nguyen Dao, Thang, & Dávila, Julio. 2013. Can Geography Lock a Society in Stagnation? *Economics Letters*, **120**, 442–446.
- O'Rourke, Kevin Hjortshøj, Rahman, Ahmed S., & Taylor, Alan M. 2013. Luddites, the industrial revolution, and the demographic transition. *Journal of Economic Growth*, **18**(4), 373–409.
- Pomeranz, Kenneth. 2000. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton, NJ: Princeton University Press.
- Schaefer, A, Schiess, D, & Wehrli, R. 2014. Long-term growth driven by a sequence of general purpose technologies. *Economic Modelling*, **37**(2), 23–31.
- Smil, Vaclav. 2017. *Energy and Civilization: A History*. Cambridge, MA: MIT Press.
- Squicciarini, Mara P., & Voigtländer, Nico. 2015. Human Capital and Industrialization: Evidence from the Age of Enlightenment. *Quarterly Journal of Economics*, **130**(4), 1825–1883.
- Stephenson, Judy Z. 2018. 'Real' wages? Contractors, workers, and pay in London building trades, 1650-1800. *Economic History Review*, **71**(1), 106–132.
- Stokey, Nancy L. 1995. R and D, and Economic Growth. *Review of Economic Studies*, **62**(3), 469–489.
- Strulik, Holger, & Weisdorf, Jacob. 2008. Population, food, and knowledge: a simple unified growth theory. *Journal of Economic Growth*, **13**(3), 195–216.
- Strulik, Holger, Prettnner, Klaus, & Prskawetz, Alexia. 2013. The past and future of knowledge-based growth. *Journal of Economic Growth*, **18**(4), 411–437.
- Warde, Paul. 2007. *Energy Consumption in England and Wales, 1560–2000*. Consiglio Nazionale delle Ricerche, Istituto di Studi sulle Società del Mediterraneo.
- World Bank. 2018. *World Bank Indicators, population, total*.
- Wrigley, E. A. 2013. Energy and the English Industrial Revolution. *Philosophical Transactions of the Royal Society. Series A: Mathematical, physical, and engineering sciences*, **371**(3), 1–10.
- Wrigley, E. A, Davies, R. S., Oeppen, J. E., & Schofield, R. S. 1997. *English Population History from Family Reconstitution, 1580-1837*. Cambridge, UK: Cambridge University Press.
- Wrigley, E.A. 1962. The Supply of Raw Materials in the Industrial Revolution. *Economic History Review*, **15**(1), 1–16.
- Wrigley, Edward Anthony. 2016. *The Path to Sustained Growth: England's Transition from an Organic Economy to an Industrial Revolution*. Cambridge, UK: Cambridge University Press.
- Yakita, Akira. 2010. Human capital accumulation, fertility and economic development. *Journal of Economics*, **99**(2), 97–116.