Energy Transitions in the Long Run: Theory and Evidence from English Coal*

Jevan Cherniwchan Juan Moreno-Cruz

June 19, 2026

Abstract

We study the long-run effects of England's transition from wood to coal. We develop a Malthusian model in which energy transitions arise endogenously from population growth and alter population dynamics. We derive an estimating equation from our model and take it to county population data spanning 1086 to 1750, exploiting variation in the appropriability of coal created by the Dissolution of the Monasteries to address the endogeneity of when and where transitions occur. Our estimates confirm our model's predictions: population dynamics change starkly because of the transition, raising the population density of affected counties by roughly 28% by 1750.

JEL Codes: N53, O13, Q32
Keywords: Energy Transition, Coal, England, Malthus, Growth

*Cherniwchan: Spencer Family Associate Professor, Department of Economics, McMaster University, Canada. Moreno-Cruz: Professor, School of Environment, Enterprise and Development and Canada Research Chair, University of Waterloo, Waterloo, Canada. We are indebted to Scott Taylor for planting the seeds of this project and encouraging us to pursue it further. We would also like to thank Jeff Chan, Andrew Dickens, Adam Lavecchia, Nouri Najjar and audiences at Brock University, the University of Buffalo, Wilfrid Laurier University, the annual meetings of the Canadian Economic Association, the Canadian Resource and Environmental Economics Association, and the Urban Economic Association, the Fifth Annual Scotland and Northern England Conference in Applied Microeconomics and Sustainable Resource Use and Economic Dynamics for helpful suggestions and comments that have improved this paper. We would also like to thank Imogen Walks Kowalchuk for excellent research assistance.

1 Introduction

Human history has been marked by a series of transitions between primary energy sources. Societies have moved from biomass to coal, then to oil, natural gas, and nuclear power, and are now turning towards renewable energy. A growing body of research shows that changes in access to energy resources reshape economic activity in the short run. But an energy transition is more than a change in access: it is the displacement of an economy's primary energy source by an alternative. Yet, at present, we have a limited understanding as to why some resources become dominant, or how energy transitions affect how economies grow and develop over the long run.

The purpose of this paper is to make initial progress in this area by studying what is often referred to as the first major energy transition: England's switch from wood to coal prior to the Industrial Revolution. This is a particularly attractive setting for two reasons. The first is its simplicity. While modern economies are shifting among several energy sources at once — oil, natural gas, nuclear, and a range of renewables — England moved between just two: wood and coal. The second is one of data availability. England's transition occurred around the turn of the seventeenth century: while wood accounted for nearly 75% of all non-human and non-animal energy consumption in England and Wales as of the 1560s, coal had overtaken wood as the leading source within one hundred years (Wrigley, 2010). However, England has sub-national estimates of economic activity pre-dating this period and extending back to 1086, making it possible to observe changes in how economies grow and develop as a result of an energy transition.

We begin by developing a simple theory of a Malthusian economy that can draw energy from more than one source.[1] Our model features three sectors that use land and labor: agriculture, which produces food; an organic energy sector, which produces energy from land in the form of wood; and a mineral energy sector, which produces energy from coal. In separating organic from mineral energy, we formalize a distinction long emphasized by Wrigley (2010), who contrasts an "organic economy" that draws its energy from the land with a "mineral economy" built on coal. The two energy sectors differ in factor intensity: organic energy is relatively land-intensive, since growing wood

[1]A Malthusian model is a natural choice for our setting: previous work suggests that England remained Malthusian throughout our period of study (Clark and Hamilton, 2006; Moreno-Cruz and Taylor, 2020), meaning that any income gains accrued primarily to population rather than to living standards. It is, however, worth noting that the strength of England's Malthusian checks is debated. Clark and Hamilton (2006) document the income–fertility link underlying the Malthusian mechanism, and Kelly and Ó Gráda (2012) find evidence of an operative preventive check, while Nicolini (2007) and Crafts and Mills (2009) estimate that the checks weakened over the seventeenth and early eighteenth centuries. What matters for our analysis is only that population responded to economic gains over this period, not that living standards were perfectly pinned to subsistence.

requires extensive forest, while mineral energy is relatively labor-intensive, since coal requires little surface land but a great deal of labor to extract. Wrigley draws the contrast sharply, characterizing wood production as “areal” and coal as “punctiform.” This difference in factor intensity is what drives the transition. As population grows, land becomes scarce relative to labor, and the price of coal falls relative to the price of wood. Once the population is large enough, coal becomes the cheaper source of energy and the economy transitions from wood to coal.

Two results from our model are worth emphasizing. The first is that energy transitions need not occur. Because the transition is driven by population growth, an economy whose coal is sufficiently unproductive will reach its steady state before coal ever becomes competitive, and remain reliant on organic energy indefinitely. The second is that the transition alters the dynamics of population growth, leading to permanent changes in population levels, but not living standards. To see this, compare an economy that transitions to coal with an otherwise identical economy that never gains access to it. In the economy without coal, population growth follows standard Malthusian logic: population growth reduces wages, reducing the rate of future population growth until wages reach subsistence levels. In contrast, in the economy transitioning to coal, this logic no longer binds: the reallocation of factors across all three sectors as organic energy production falls and mineral energy production rises holds up wages, maintaining a rate of population growth that would have otherwise fallen. As a result, once the economy begins to transition, it does not stop until wood is fully displaced. Once the transition is complete, the economy once again falls into the Malthusian trap where further population growth reduces wages to subsistence levels. Hence, while the living standards of the coal economy eventually match those of the one without, the increased rate of population growth during the energy transition leads to higher population levels in the long run.

As the second step in our analysis, we take our theory to the data to test if England’s transition from wood to coal changed population dynamics in the manner our model predicts. Doing so poses two key empirical challenges. The first is that we do not observe the transition. We observe population estimates for the "ancient" counties of England at six dates between the Domesday Book of 1086 and 1750, but the historical record does not tell us when, or whether, each county switched from wood to coal.[2] The second is that, even if we did observe it, the transition would not be exogenous. In our

[2] We draw our population data from Broadberry et al. (2015) and Wrigley (2009), and describe their construction in Section 3. Our main estimation sample comprises thirty-two counties observed in 1086, 1290, 1377, 1600, 1700, and 1750.

model, a county transitions once its population crosses a threshold, and that threshold depends on primitives, such as the productivity of its coal, that we do not observe and that also govern the county's population growth. The timing of the transition is therefore jointly determined with the population dynamics we are trying to explain. Identifying the effect of the transition thus requires exogenous variation that shifts this threshold for some counties but not others.

We obtain this variation from the Dissolution of the Monasteries. Between 1536 and 1540, a series of Acts of Parliament transferred property from the Catholic Church to the Crown, which sold much of it to private landowners to help finance Henry VIII's wars (Habakkuk, 1958). The transfer was among the largest reallocations of property in English history, and Heldring et al. (2021) show that it left broad and lasting marks on England's economic development through its effects on commercialization. For our purposes, what matters is one specific consequence. Much of the coal in early-sixteenth-century England lay beneath Church land, and, as Nef (1932a) documents, the Church leased that land on terms that discouraged mining, while the private owners who acquired it after the Dissolution offered far more favorable terms. We summarize this contractual shift through what we term the *appropriability* of coal: the share of mining revenue retained within the local economy rather than claimed by absentee landowners. Introduced into our model, a rise in appropriability lowers the population at which coal becomes competitive, acting much like an increase in the productivity of coal. Population growth still drives a transition, but a rise in appropriability widens the set of counties for which an energy transition is feasible and accelerates its onset for those that would have transitioned regardless. Where coal is sufficiently unproductive, or absent from Church land, the Dissolution has no effect at all. This is the basis of our research design: we compare counties that had coal on Church land, and were thus exposed to the change in appropriability, with those that did not, before and after the Dissolution.

We implement this design using a theory-derived estimating equation that links a county's population density in a given period to its population density in the previous period, whether it has begun to transition to coal and ongoing changes in agricultural productivity. We proxy for the transition using whether the county had coal on former Church land and whether the period follows the Dissolution. Our estimating equation delivers two primary coefficients of interest. The first reflects population growth before a county begins to transition, when it is governed by Malthusian dynamics. The second reflects dynamics during the transition. We use these coefficient estimates to test the central prediction of our model: that the energy transition changed the dynamics of population growth, replacing Malthusian reversion with growth at a constant rate.

Moreover, because our estimating equation is recursive, these same estimated dynamics let us recover the transition's average long-run effect on affected counties: the cumulative change in population density from the energy transition relative to what would have occurred in the absence of coal.

Our empirical results provide robust evidence that matches two key predictions from our theory. The first is that during the transition, population growth persists at a constant rate. The estimates from our preferred specification indicate that once counties begin to use coal, population dynamics feature a generational persistence of 1.0, meaning population growth does not level off, matching the value implied by our theory. The second is that the energy transition significantly changed the dynamics of population growth. Our preferred estimate indicates that in the absence of coal, English counties followed standard Malthusian dynamics, with a generational population persistence of 0.8. Moreover, this estimate is significantly different from the value during the energy transition, meaning we reject the null of no change in dynamics, as our theory predicts. Compounded over the generations after the Dissolution, this change in population dynamics implies a sizable cumulative effect: by 1750, the transition had raised the population density of affected counties by roughly 28% relative to what would have occurred in the absence of coal.[3] Our empirical estimates also indicate that this change in population dynamics persisted as late as 1801, suggesting that England did not revert to Malthusian dynamics before the Industrial Revolution began in earnest.

Altogether, our results contribute to several literatures. First, we contribute to the literature examining how changing access to energy sources affects economic activity. Whether the focus is the effects of coal after the Industrial Revolution (Matheis, 2016; Beach and Hanlon, 2018; Hanlon, 2020; Clay et al., 2024), electrification (Dinkelman, 2011; Lipscomb et al., 2013; Lewis and Severnini, 2020; Severnini, 2023), or the recent shale gas boom in the United States (Muehlenbachs et al., 2015; Feyrer et al., 2017; Allcott and Keniston, 2018), most previous research relies on data spanning years or, at most, decades, providing estimates of the short-run effects of changes in access to energy.[4] Our contribution stems from measuring something different: the long-run dynamic effects of

[3] A natural concern is that our estimates are also capturing the effects of other phenomena. We address this concern at length. We show the change in population dynamics is robust to measuring a county's exposure to the Dissolution continuously rather than with an indicator, and it is not driven by differential trends in transport costs, in migration toward London, Bristol, or Liverpool, or in the availability of timber as an alternative fuel. We also show our results are not capturing the effects of the broader commercialization that Heldring et al. (2021) attribute to the Dissolution itself, or the effects of the introduction of New World crops. Our estimates are also robust to alternative measures of agricultural productivity.

[4] A more recent body of work has begun to study the effects of transitioning *away* from fossil fuels (Hanson, 2023; Haywood et al., 2024; Rud et al., 2024), but these papers are similarly focused on short-run outcomes.

the transition from wood to coal. By examining the effects of a new energy source over such a long interval, our work is closest to that of Fernihough and O'Rourke (2021), who estimate the effects of coal on a large sample of European cities between 1300 and 1900. Our work differs from theirs in two key respects. The first is in our use of theory and empirics to study coal's effects on the dynamics of population growth; Fernihough and O'Rourke estimate the average effect of coal adoption on the population levels of affected cities. The second is in our unit of observation: we study the total population of English counties rather than the size of individual cities, and we find that the transition to coal began to raise county populations before the Industrial Revolution — somewhat earlier than the city-growth effect documented by Fernihough and O'Rourke, a difference consistent with coal reshaping rural settlement and proto-industrial population before it became visible in the growth of cities.

Second, in building a model of England's transition to coal, we contribute to a growing literature that develops dynamic models of energy transitions. Most of this work concerns the prospective transition away from fossil fuels (Acemoglu et al., 2016; Arkolakis and Walsh, 2023; Besley and Persson, 2023; Lemoine, 2024, 2025); of existing work, ours is closest to Stokey (2001) and Stern et al. (2021), who likewise study England's transition to coal but do so with quantitative models centered on technological change and the Industrial Revolution. Our model differs in treating the transition as a consequence of population growth rather than of technological progress, and in showing that institutions — here, the contractual arrangements governing coal — can shape when and where an energy transition occurs.

Third, we contribute to a broad literature studying pre-industrial growth through a Malthusian lens.[5] Our model borrows the standard machinery of Malthusian growth (Ashraf and Galor, 2011; Lagerlöf, 2019; Blanc and Wacziarg, 2025), but departs from it by allowing the steady-state population to depend on the economy's choice of energy source, and by allowing that choice to shift endogenously, as an economy's reliance on relatively land-intensive energy gives way to a relatively labor-intensive alternative. In this respect our account also differs from existing models of economic change in England and elsewhere in Europe around the Industrial Revolution, which typically attribute structural change to forces such as fertility decline, disease, warfare, technological progress, international trade, or urbanization (Galor and Weil, 2000; Hansen and Prescott, 2002; O'Rourke et al., 2013; Voigtländer and Voth, 2013b; Trew, 2014). We show instead that a transition between energy sources can itself reshape an economy's long-run population, without any change in underlying production technologies.

[5] For recent surveys, see Ashraf et al. (2021) and Galor (2024).

Finally, our work adds to a long-standing debate over the role of coal in England's economic development, one that dates at least to Jevons (1865). Much of this debate has centered on coal's contribution to the Industrial Revolution (Pomeranz, 2000; Clark and Jacks, 2007; Allen, 2009; Mokyr, 2009; Wrigley, 2010; Fernihough and O'Rourke, 2021; Kelly et al., 2023). We contribute to this literature by building on Nef (1932a) to provide evidence that coal shaped the English economy even before the Industrial Revolution. In doing so, we complement two recent accounts of this period. The first is that of Heldring et al. (2021), who, as we discussed above, show that the Dissolution of the Monasteries stimulated English growth through its effects on commercialization. We build on their work by showing that the Dissolution, by raising the appropriability of coal where it lay beneath Church land, also changed England's population dynamics.[6] The second is that of Bouscasse et al. (2025), who date the onset of sustained growth in output per person to around 1600. Our work points to a change occurring on a different margin at much the same time: where they identify the beginnings of intensive growth, in output per person, we identify a change in the dynamics of population — the extensive margin — driven by the transition to coal.

The remainder of the paper proceeds as follows. Section 2 develops our theory. Section 3 extends the theory to the English setting and derives the estimating equation we take to the data. Section 4 presents our empirical results. Section 5 concludes.

2 A Malthusian Theory of Energy Transitions

We begin by developing a simple Malthusian model featuring multiple energy sources. Our motivations for doing so are twofold. First, as we discussed above, existing theories of growth during the preindustrial era have abstracted from the explicit study of non-agricultural energy resources, making the effects of an energy transition difficult to predict ex-ante. Hence, we use our theory to formalize how such an energy transition could arise endogenously in a preindustrial society, and then show how it affects population dynamics. Second, as we show later in Section 3, we adapt this theory to our empirical setting and derive the estimating equation we use in our empirical analysis.

[6] It is worth noting that Heldring et al. (2021) show that their estimates are robust to controlling for the effects of coal; as we discuss further below, we similarly show our estimates are robust to controlling for the broader effects of the Dissolution.

2.1 Setup

We consider a Malthusian economy in discrete time in which each period τ corresponds to one generation. The economy at time τ is comprised of N_τ identical households that each live one period and supply one unit of labor inelastically, earning wage w_τ. The economy is endowed with a fixed stock of land $\bar{L} > 0$, with rental rate σ_τ. For simplicity, we assume that wages are the only source of income for households; land is owned by an infinitely-lived elite that uses the resulting income to finance external consumption.[7] All factor markets are competitive.

As is standard in Malthusian models (see, e.g. Ashraf and Galor, 2011; Voigtländer and Voth, 2013b), households in the economy derive utility from consumption, q_τ, and raising children, n_τ. Household preferences are given by:

$$U_\tau = q_\tau^{1-\eta} n_\tau^{\eta} \tag{1}$$

where $\eta \in (0, 1)$. We depart from standard formulations by letting household consumption be a Cobb-Douglas aggregate of agricultural output (F) and energy (E):

$$q_\tau = q_{F,\tau}^{\mu} q_{E,\tau}^{1-\mu} \tag{2}$$

where $\mu \in (0, 1)$. Substituting Equation (2) into Equation (1), and defining $\nu_F \equiv \mu[1 - \eta]$, $\nu_E \equiv [1 - \mu][1 - \eta]$, and $\nu_n \equiv \eta$ for convenience, the household's problem is to maximize:

$$U_\tau = q_{F,\tau}^{\nu_F} q_{E,\tau}^{\nu_E} n_\tau^{\nu_n} \tag{3}$$

subject to the budget constraint:

$$p_{F,\tau} q_{F,\tau} + p_{E,\tau} q_{E,\tau} + \phi n_\tau = w_\tau \tag{4}$$

where $p_{F,\tau}$ is the price of agricultural output, $p_{E,\tau}$ is the price of energy, and $\phi > 0$ is the cost of raising a child.

Utility maximization implies households devote fixed fractions of their income to the

[7] Assuming wages are the only source of income for households is a standard simplification in preindustrial growth models. Our setup is closest to Voigtländer and Voth (2013a), where land is the only source of income for landlords and their expenditures "finance[s] activities that are unrelated to agricultural production (e.g., building castles or engaging in warfare)". Galor and Weil (2000, fn. 6) and Lagerlöf (2019) adopt looser variants of the convention.

consumption of agricultural output, energy and raising children:

$$q_{F,\tau} = \nu_F \frac{w_\tau}{p_{F,\tau}}, \qquad q_{E,\tau} = \nu_E \frac{w_\tau}{p_{E,\tau}}, \qquad n_\tau = \nu_n \frac{w_\tau}{\phi} \tag{5}$$

Thus, the ratio of household spending on agricultural output to energy is constant and determined by preferences:

$$\frac{p_{E,\tau} q_{E,\tau}}{p_{F,\tau} q_{F,\tau}} = \frac{\nu_E}{\nu_F} \tag{6}$$

Furthermore, Equation (5) shows that fertility depends on the income levels of households. We will return to the resulting population dynamics in Section 2.3.

The economy features three perfectly competitive sectors that use land and labor as inputs: agriculture (F), organic energy (O) and mineral energy (M). We assume that households view organic and mineral energy as perfect substitutes in consumption; as a result, households choose the cheapest energy source when making their consumption decisions, meaning the organic and mineral energy sectors are both active only when the price of both energy sources is the same. Output from sector j at time τ is given by:

$$X_{j,\tau} = A_j L_{j,\tau}^{\alpha_j} N_{j,\tau}^{1-\alpha_j} \tag{7}$$

where $\alpha_j \in (0,1)$, $L_{j,\tau}$ denote the land, and $N_{j,\tau}$ denote the labor used in the production of j at time t, and $A_j > 0$ is an exogenous sectoral productivity parameter. We assume $\alpha_O > \alpha_M$, that is, organic energy is relatively more land intensive than mineral energy. These characterizations reflect the physical realities of these production technologies in pre-industrial economies: the production of organic energy—which primarily comes from wood—requires extensive land for forests and biomass to grow. Extracting mineral energy—coal—requires little surface land.

Given Equation (7), the unit cost for sector j is:

$$c_{j,\tau}(\sigma_\tau, w_\tau) = \frac{1}{A_j} \left[\frac{\sigma_\tau}{\alpha_j}\right]^{\alpha_j} \left[\frac{w_\tau}{1-\alpha_j}\right]^{1-\alpha_j} \tag{8}$$

As such, unit costs in each sector depend on the prevailing wage, w_τ, and the rental rate on land σ_τ.

Let $p_{j,\tau}$ be the output price in sector j at time τ. Perfect competition implies $p_{j,\tau} = c_{j,\tau}(\sigma_\tau, w_\tau)$ if sector j is actively produced. Moreover, cost minimization implies that in each active sector:

$$\sigma_\tau = \alpha_j \frac{p_{j,\tau} X_{j,\tau}}{L_{j,\tau}}, \qquad w_\tau = [1-\alpha_j] \frac{p_{j,\tau} X_{j,\tau}}{N_{j,\tau}} \tag{9}$$

Finally, all land and labor is employed in production in each period:

$$\sum_j L_{j,\tau} = \bar{L}, \qquad \sum_j N_{j,\tau} = N_\tau \tag{10}$$

where each sum is defined over the set of active sectors at time t.

2.2 Equilibrium and the Regime Threshold

Given that households treat organic and mineral energy as perfect substitutes, the set of active sectors in the economy in period τ will be determined by the relative prices of the two energy sources in equilibrium. We now turn to determine the conditions under which organic and mineral energy are each economically viable. For convenience, in what follows, we let agricultural output be the numeraire.

When $p_{M,\tau} > p_{O,\tau}$, the economy is in an *organic regime*: households consume organic energy, meaning organic energy and agricultural output are actively produced. Combining the spending ratio (6) with the factor-pricing conditions (9) yields the share of labor and land allocated to agriculture in the organic regime:

$$\psi_O \equiv \frac{L_{F,\tau}}{\bar{L}} = \frac{\alpha_F \nu_F}{\alpha_F \nu_F + \alpha_O \nu_E} \qquad \lambda_O \equiv \frac{N_{F,\tau}}{N_t} = \frac{[1-\alpha_F]\nu_F}{[1-\alpha_F]\nu_F + [1-\alpha_O]\nu_E}. \tag{11}$$

Both shares are constants, independent of population. With these allocations in hand, the first-order conditions from the agricultural sector deliver wages and land rents as a function of population density:

$$w_\tau = \Omega_O^w \left[\frac{N_\tau}{\bar{L}}\right]^{-\alpha_F}, \qquad \sigma_\tau = \Omega_O^\sigma \left[\frac{N_\tau}{\bar{L}}\right]^{[1-\alpha_F]} \tag{12}$$

where:

$$\Omega_O^w \equiv [1-\alpha_F] A_F \left[\frac{\psi_O}{\lambda_O}\right]^{\alpha_F} \qquad \text{and} \qquad \Omega_O^\sigma \equiv \alpha_F A_F \left[\frac{\lambda_O}{\psi_O}\right]^{[1-\alpha_F]}$$

are constants. As Equation (12) shows, in the organic economy, when the population increases, land becomes relatively scarce; wages fall and land rents rise as a result.

In the organic regime $p_{E,\tau} = p_{O,\tau}$, so the price of energy follows from the zero profit condition for the organic sector:

$$p_{O,\tau} = \Phi_O \left[\frac{N_\tau}{\bar{L}}\right]^{[\alpha_O - \alpha_F]} \tag{13}$$

where:

$$\Phi_O \equiv \frac{1}{A_O}\left[\frac{\Omega_O^\sigma}{\alpha_O}\right]^{\alpha_O}\left[\frac{\Omega_O^w}{1-\alpha_O}\right]^{[1-\alpha_O]}$$

is a constant. As the economy's population increases, land rents increase. When organic energy is more land-intensive than agriculture ($\alpha_O > \alpha_F$), the unit cost of organic energy increases faster than the general price level, and energy becomes progressively more expensive in a crowded organic economy. When instead agriculture is the more land-intensive of the two ($\alpha_F > \alpha_O$), the unit cost of agricultural output rises faster, and energy becomes progressively cheaper as the organic economy grows. In either case, the absolute prices of both goods rise with population; what differs is only their relative prices.

When $p_{M,\tau} < p_{O,\tau}$, the economy is in a *mineral regime*. Households consume mineral energy: the economy is in a mineral regime in which mineral energy and agricultural output are actively produced. Using a similar approach as above yields the factor allocations for the mineral regime:

$$\psi_M = \frac{\alpha_F \nu_F}{\alpha_F \nu_F + \alpha_M \nu_E}, \qquad \text{and} \qquad \lambda_M = \frac{[1-\alpha_F]\nu_F}{[1-\alpha_F]\nu_F + [1-\alpha_M]\nu_E}. \tag{14}$$

Wages and land rents can then be expressed as:

$$w_\tau = \Omega_M^w \left[\frac{N_\tau}{\bar{L}}\right]^{-\alpha_F}, \qquad \sigma_\tau = \Omega_M^\sigma \left[\frac{N_\tau}{\bar{L}}\right]^{[1-\alpha_F]} \tag{15}$$

where:

$$\Omega_M^w \equiv [1-\alpha_F] A_F \left[\frac{\psi_M}{\lambda_M}\right]^{\alpha_F} \qquad \text{and} \qquad \Omega_M^\sigma \equiv \alpha_F A_F \left[\frac{\lambda_M}{\psi_M}\right]^{[1-\alpha_F]}$$

are constants. It is worth noting that the elasticities of the wage and rental rate of land with respect to population are the same in the organic and mineral regimes because they both share the same technology for agricultural production; only the constants differ across regimes.

The price of energy in the mineral regime is given by:

$$p_{M,\tau} = \Phi_M \left[\frac{N_\tau}{\bar{L}} \right]^{[\alpha_M - \alpha_F]} \tag{16}$$

where:

$$\Phi_M \equiv \frac{1}{A_M} \left[\frac{\Omega_M^\sigma}{\alpha_M} \right]^{\alpha_M} \left[\frac{\Omega_M^w}{1-\alpha_M} \right]^{[1-\alpha_M]}$$

is a constant. As the economy's population increases, land rents continue to rise and wages continue to fall. The effect on the unit cost of mineral energy depends on the ranking of α_F and α_M. When agriculture is more land-intensive than mineral energy ($\alpha_F > \alpha_M$), mineral's cost structure is dominated by labor and its relative price falls with population. When mineral energy is more land-intensive than agriculture ($\alpha_M > \alpha_F$), the opposite holds: mineral's cost is dominated by land and its relative price rises with population.

When $p_{M,\tau} = p_{O,\tau}$, the economy is in a *transition regime*: households are indifferent between the two energy sources, meaning agriculture, organic energy, and mineral energy are all actively produced. With both energy sources produced, perfect competition imposes two zero profit conditions that involve factor prices alone: agricultural output is the numeraire, so $c_{F,\tau}(\sigma_\tau, w_\tau) = 1$, and the two energy sources must sell at a common price, so $c_{O,\tau}(\sigma_\tau, w_\tau) = c_{M,\tau}(\sigma_\tau, w_\tau)$. Population appears in neither condition.

The equality of energy costs fixes the ratio of land rents to wages:

$$\frac{\sigma_\tau}{w_\tau} = \Xi \equiv \left[\frac{A_O\, \alpha_O^{\alpha_O} [1-\alpha_O]^{[1-\alpha_O]}}{A_M\, \alpha_M^{\alpha_M} [1-\alpha_M]^{[1-\alpha_M]}} \right]^{\frac{1}{\alpha_O - \alpha_M}} \tag{17}$$

The zero profit condition for agriculture then delivers the levels of wages and land rents in the transition regime:

$$w_\tau = \Omega_T^w, \qquad \sigma_\tau = \Omega_T^\sigma \tag{18}$$

where:

$$\Omega_T^w \equiv A_F\, \alpha_F^{\alpha_F} [1-\alpha_F]^{[1-\alpha_F]}\, \Xi^{-\alpha_F} \qquad \text{and} \qquad \Omega_T^\sigma \equiv A_F\, \alpha_F^{\alpha_F} [1-\alpha_F]^{[1-\alpha_F]}\, \Xi^{[1-\alpha_F]}$$

are constants. In contrast with Equations (12) and (15), population does not appear in

Equation (18): as long as the economy remains in the transition regime, wages do not fall and land rents do not rise as the population grows.

In the transition regime $p_{E,\tau} = p_{O,\tau} = p_{M,\tau}$, so the price of energy follows from the zero profit condition for either energy sector:

$$p_{E,\tau} = \Phi_T \tag{19}$$

where:

$$\Phi_T \equiv \frac{1}{A_O} \left[\frac{\Omega_T^{\sigma}}{\alpha_O}\right]^{\alpha_O} \left[\frac{\Omega_T^{w}}{1-\alpha_O}\right]^{[1-\alpha_O]} = \frac{1}{A_M} \left[\frac{\Omega_T^{\sigma}}{\alpha_M}\right]^{\alpha_M} \left[\frac{\Omega_T^{w}}{1-\alpha_M}\right]^{[1-\alpha_M]}$$

is a constant; the two expressions coincide by the definition of Ξ. Like wages and land rents, the price of energy does not vary with population in the transition regime.

With factor prices pinned by costs, it is the allocation of land and labor that absorbs population growth. Cost minimization at the factor prices in Equation (18) fixes the land-labor ratio of every sector at $\frac{L_{j,\tau}}{N_{j,\tau}} = \frac{\alpha_j}{1-\alpha_j}\,\Xi^{-1}$. Combining the spending ratio (6) with the factor-pricing conditions (9) and the full employment of land and labor yields the share of labor and land allocated to agriculture in the transition regime:

$$\begin{aligned} \psi_{T,\tau} &\equiv \frac{L_{F,\tau}}{\bar{L}} = \frac{\alpha_F \nu_F}{\nu_F + \nu_E} \left[1 + \Xi^{-1} \frac{N_\tau}{\bar{L}}\right], \\ \lambda_{T,\tau} &\equiv \frac{N_{F,\tau}}{N_\tau} = \frac{[1-\alpha_F]\nu_F}{\nu_F + \nu_E} \left[1 + \Xi \left[\frac{N_\tau}{\bar{L}}\right]^{-1}\right]. \end{aligned} \tag{20}$$

In contrast with Equations (11) and (14), both shares now vary with population, and they do so between the allocations of the two other regimes. Because the energy sectors must employ non-negative quantities of land and labor, the transition regime prevails over a range of populations: at the lower end of this range, mineral employment is zero and the shares in Equation (20) coincide with the organic allocations in Equation (11); at the upper end, organic employment is zero and they coincide with the mineral allocations in Equation (14).

As the economy's population increases in the transition regime, agriculture's land share rises while its labor share falls, and the two energy sectors trade places: land and labor reallocate from the land-intensive organic sector to the labor-intensive mineral sector, so organic energy production contracts as mineral energy production expands. This compositional margin is what keeps factor prices constant. In the organic and mineral regimes, additional workers must be absorbed by the two active sectors, and with

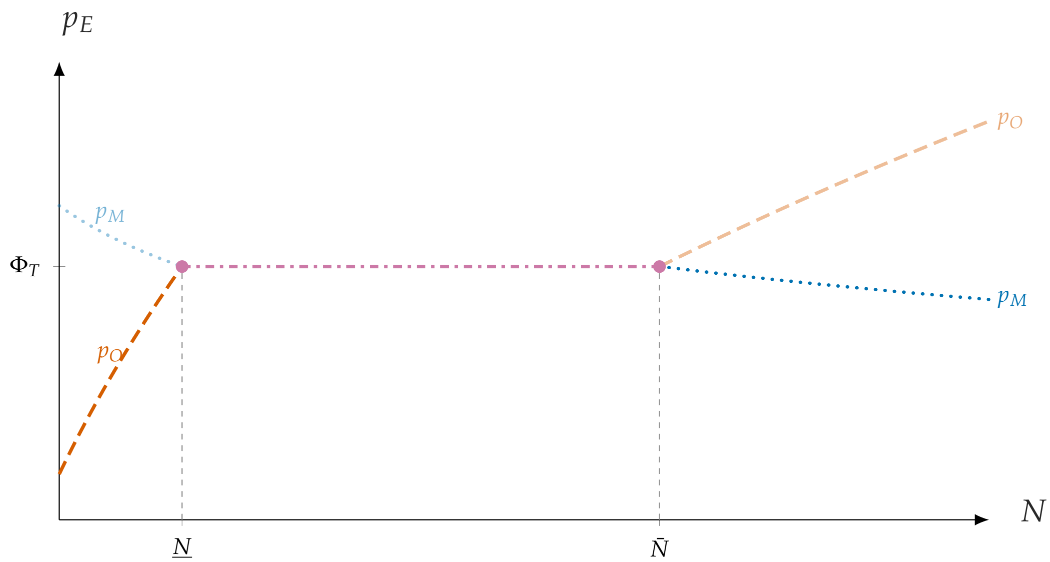


Figure 1: Energy Prices and the Transition Regime.

Notes: The figure plots the price of energy as a function of population. For populations below $\underline{N}$, the economy is in the organic regime and energy is priced at the organic price p_O from Equation (13) (orange, dashed). For populations above $\bar{N}$, the economy is in the mineral regime and energy is priced at the mineral price p_M from Equation (16) (blue, dotted). For populations between $\underline{N}$ and $\bar{N}$, the economy is in the transition regime: both energy sources are produced and the price of energy is constant at Φ_T from Equation (19) (purple, dash-dotted). The faded curves depict the break-even price of the energy source that is not produced in each range, evaluated at the prevailing factor prices. The figure depicts the case $\alpha_O > \alpha_F > \alpha_M$, in which the organic price rises and the mineral price falls with population; the boundaries of the transition regime arise identically under the other admissible orderings of α_F.

land fixed this depresses wages and raises land rents. When both energy sources are produced, the economy can instead accommodate a larger workforce by producing relatively more of the labor-intensive energy source and relatively less of the land-intensive one, which employs the additional workers and releases exactly the land they would otherwise have competed for. Population growth is thus absorbed by the composition of energy production rather than by factor prices.

The behavior of the price of energy is shown in Figure 1, which plots the prices given by Equations (13), (16) and (19) over the corresponding ranges of population. How the level of each energy price moves with population depends on the ordering of α_F relative to α_O and α_M; three cases are consistent with the maintained assumption $\alpha_O > \alpha_M$. In Case 1 ($\alpha_O > \alpha_F > \alpha_M$), depicted in Figure 1, the price of organic energy rises with population while the price of mineral energy falls. In Case 2 ($\alpha_F > \alpha_O > \alpha_M$), both energy prices fall with population, but the labor-intensive mineral price falls more steeply than the land-intensive organic price. In Case 3 ($\alpha_O > \alpha_M > \alpha_F$), both energy prices rise with population, but the land-intensive organic price rises more steeply than the labor-intensive mineral price. The three cases differ in how the individual prices move, but they share the feature that drives the energy transition: because organic energy is more land-intensive than mineral energy, the cost of mineral energy relative to organic energy is strictly decreasing in the ratio of land rents to wages, and this ratio rises with population in any single-source regime. As a growing population bids up land rents and depresses wages, mineral energy becomes progressively more competitive and it eventually becomes viable to produce, and organic energy eventually stops being viable. These two events define the boundaries of the transition regime:

Proposition 1. *There exist unique population thresholds $\underline{N}$, at which mineral energy enters, and $\bar{N}$, at which organic energy exits, given by:*

$$\underline{N} = \bar{L}\left[\frac{1-\alpha_F}{\alpha_F}\right]\frac{\psi_O}{\lambda_O}\,\Xi \qquad \textit{and} \qquad \bar{N} = \bar{L}\left[\frac{1-\alpha_F}{\alpha_F}\right]\frac{\psi_M}{\lambda_M}\,\Xi \tag{21}$$

with $\underline{N} < \bar{N}$. The economy is in the organic regime when $N_t < \underline{N}$, in the transition regime when $\underline{N} \leq N_t < \bar{N}$, and in the mineral regime when $N_t \geq \bar{N}$.

Proof. Consider an economy in the organic regime. Mineral energy is produced only if a producer can cover its unit cost at the prevailing factor prices, that is, only if $c_{M,\tau}(\sigma_\tau, w_\tau) \leq c_{O,\tau}(\sigma_\tau, w_\tau)$. The ratio of the two unit costs is:

$$\frac{c_{M,\tau}(\sigma_\tau, w_\tau)}{c_{O,\tau}(\sigma_\tau, w_\tau)} = \left[\Xi\,\frac{w_\tau}{\sigma_\tau}\right]^{\alpha_O-\alpha_M},$$

which is strictly decreasing in σ_τ/w_τ and equals one exactly when $\sigma_\tau/w_\tau = \Xi$, as in Equation (17). From Equation (12), the factor-price ratio in the organic regime,

$$\frac{\sigma_\tau}{w_\tau} = \frac{\Omega_O^\sigma}{\Omega_O^w}\frac{N_\tau}{\bar{L}} = \frac{\alpha_F}{1-\alpha_F}\frac{\lambda_O}{\psi_O}\frac{N_\tau}{\bar{L}},$$

is strictly increasing in population, rising from zero to infinity, so there is a unique population at which it equals Ξ; solving yields $\underline{N}$ in Equation (21). For $N_t < \underline{N}$, mineral energy cannot cover its unit cost and the organic regime prevails; for $N_t \geq \underline{N}$, an inactive mineral producer could profitably enter, so the organic regime cannot be an equilibrium. The argument for the mineral regime is symmetric: from Equation (15), the factor-price ratio is $[\alpha_F/[1-\alpha_F]][\lambda_M/\psi_M][N_\tau/\bar{L}]$, and organic energy covers its unit cost only when this ratio is at most Ξ. The unique solution of $\sigma_\tau/w_\tau = \Xi$ is $\bar{N}$ in Equation (21): for $N_t > \bar{N}$, organic energy cannot cover its cost and the mineral regime prevails, while for $N_t < \bar{N}$, an inactive organic producer could profitably enter, so the mineral regime cannot be an equilibrium. Dividing the two thresholds gives:

$$\frac{\bar{N}}{\underline{N}} = \frac{\psi_M/\lambda_M}{\psi_O/\lambda_O} > 1$$

since $\alpha_O > \alpha_M$ implies $\psi_M > \psi_O$ and $\lambda_O > \lambda_M$; thus $\underline{N} < \bar{N}$ and the transition regime prevails over a non-empty range of populations. Finally, for $\underline{N} \leq N_t < \bar{N}$, neither single-source regime is an equilibrium, while the transition-regime allocations in Equation (20) assign non-negative quantities of land and labor to all three sectors exactly on this interval, with mineral employment equal to zero at $\underline{N}$ and organic employment equal to zero at $\bar{N}$. □

As Figure 1 shows, at any population $N_t < \underline{N}$, mineral energy cannot cover its cost of production: the economy is in the organic regime, and agriculture and organic energy are the active sectors. When $N_t \geq \bar{N}$, the opposite holds: organic energy can no longer cover its cost, and agriculture and mineral energy are the active sectors. For populations between the two thresholds, both energy sources break even at the prevailing factor prices, all three sectors are active, and the economy is in the transition regime, where factor prices and the price of energy are constant. It is also worth noting that both thresholds are decreasing in the productivity of the mineral energy sector:

$$\frac{\partial \underline{N}}{\partial A_M} = -\frac{\underline{N}}{A_M[\alpha_O - \alpha_M]} < 0 \qquad \text{and} \qquad \frac{\partial \bar{N}}{\partial A_M} = -\frac{\bar{N}}{A_M[\alpha_O - \alpha_M]} < 0. \tag{22}$$

Hence, an economy in which A_M is high, as would be the case when mineral energy is easy to produce, both enters and completes its energy transition at lower population levels than an economy for which A_M is low. Because the sectoral productivities enter the two thresholds only through Ξ, their ratio $\bar{N}/\underline{N}$ is independent of A_M. The productivity of mineral energy determines where the transition regime is located, not its proportional width.

2.3 Population Dynamics and the Energy Transition

The evolution of the economy's population depends on the initial number of households $N_0 > 0$ and their fertility choices n_t. Thus, the size of the population at $\tau + 1$, $N_{\tau+1}$, is:

$$N_{\tau+1} = n_\tau N_\tau. \tag{23}$$

Recall from Equation (5) that each household raises $n_\tau = \nu_n w_\tau / \phi$ children. As w_τ varies across regimes, substituting for the wages in Equation (12), Equation (18), and Equation (15) gives a law of motion with one branch per regime:

$$N_{\tau+1} = \begin{cases} \dfrac{\nu_n \Omega_O^w}{\phi} \bar{L}^{\alpha_F} N_\tau^{1-\alpha_F} & N_\tau < \underline{N} \\ \dfrac{\nu_n \Omega_T^w}{\phi} N_\tau, & \underline{N} \leq N_\tau \leq \bar{N} \\ \dfrac{\nu_n \Omega_M^w}{\phi} \bar{L}^{\alpha_F} N_\tau^{1-\alpha_F} & N_\tau > \bar{N} \end{cases} \tag{24}$$

The three branches join continuously at the thresholds: substituting $\underline{N}$ and $\bar{N}$ from Equation (21) into Equations (12) and (15) yields $w_O(\underline{N}) = \Omega_T^w = w_M(\bar{N})$, so the wage—and with it fertility—passes between regimes without jumps.

Since $\alpha_F \in (0, 1)$, $N_{\tau+1}$ is strictly concave in N_τ in the organic and mineral regimes. In the transition regime, by contrast, the constant wage fixes fertility, so the law of motion is linear with constant growth factor $n_T = (\nu_n \Omega_T^w)/\phi$. We then have the following:

Proposition 2. *Starting from $N_0 > 0$, each single-source regime has a unique steady state given by:*

$$N_e^* = \left[\frac{\nu_n \Omega_e^w}{\phi} \right]^{\frac{1}{\alpha_F}} \bar{L}, \qquad e \in \{O, M\} \tag{25}$$

Moreover, the steady state population in the mineral regime exceeds that of the organic regime; that is, $N_O^* < N_M^*$. Despite this, the steady state wage in both regimes is $w^* = \phi/\nu_n$. The transition regime contains no steady state unless $\Omega_T^w = \phi/\nu_n$.

Proof. In a steady state $N_{\tau+1} = N_\tau$. Making this substitution in Equation (24) and rearranging yields Equation (25). To show that $N_O^* < N_M^*$, note that the ratio of steady state population levels across the two regimes can be written as:

$$\frac{N_M^*}{N_O^*} = \left[\frac{\Omega_M^w}{\Omega_O^w} \right]^{\frac{1}{\alpha_F}} = \frac{\psi_M}{\psi_O} \cdot \frac{\lambda_O}{\lambda_M} \tag{26}$$

where $\psi_M/\psi_O > 1$ and $\lambda_O/\lambda_M > 1$ since $\alpha_O > \alpha_M$. Thus, $N_M^* > N_O^*$. Next, note that from Equation (23), $N_{\tau+1} = N_\tau$ in steady state requires $n_\tau = 1$. From Equation (5), this implies $w^* = \phi/\nu_n$, regardless of regime. Finally, on the transition branch of Equation (24), $N_{\tau+1} = N_\tau$ would require $n_T = \nu_n \Omega_T^w / \phi = 1$; as Ω_T^w is fixed by technology, this holds only in the knife-edge case $\Omega_T^w = w_O(\underline{N}) = \phi/\nu_n$, so generically the transition regime contains no steady state. □

Proposition 2 shows that the mineral regime is capable of supporting a larger steady state population than the organic regime. Yet, as the Proposition shows, both regimes have the same steady state level of income per household, reflecting the Malthusian trap. Furthermore, it is straightforward to verify $\partial N_i^* / \partial A_F > 0$, meaning increases in agricultural productivity increase steady state population levels, as is typical in previous Malthusian models of the preindustrial era.

Whether and when the economy transitions between energy sources depends on how its initial population $N_0 > 0$ compares to the entry and exit thresholds, $\underline{N}$ and $\bar{N}$, and the organic and mineral steady states, N_O^* and N_M^*:

Proposition 3. *Suppose the economy starts with an initial population $N_0 \in (0, N_M^*)$. Then population growth leads to one of three outcomes:*

1. ***An Organic Economy** ($N_0 \leq N_O^* < \underline{N}$): The economy converges to the organic steady state N_O^*; mineral energy is never economically viable.*
2. ***An Energy Transition** ($N_0 < \underline{N} < N_O^*$): The economy starts in the organic regime and grows until the population reaches $\underline{N}$ in finite time. It then enters the transition regime, growing at the constant rate $n_T > 1$ until the population exceeds $\bar{N}$, and converges to the mineral steady state $N_M^* > \bar{N}$.*
3. ***A Mineral Economy** ($\bar{N} < N_0 < N_M^*$): The economy converges to the mineral steady state; organic energy is never viable.*

Proof. To start, note that the organic and mineral branches of Equation (24) are strictly concave, each with a unique positive fixed point given by Equation (25), so an economy that remains in a single-source regime converges monotonically to that regime's steady state. Second, note that from Equation (12), $w_O(N)$ is strictly decreasing with $w_O(N_O^*) = w^* = \phi/\nu_n$; since $w_O(\underline{N}) = \Omega_T^w$, it follows that $n_T = \nu_n \Omega_T^w / \phi > 1$ if and only if $\underline{N} < N_O^*$.

1. If $N_0 \leq N_O^* < \underline{N}$, then the economy grows monotonically to N_O^* and reaches the organic steady state without ever crossing $\underline{N}$; mineral energy never covers its cost.
2. If $N_0 < \underline{N} < N_O^*$, then the economy begins growing monotonically towards N_O^*. Since $\underline{N} < N_O^*$, there exists some finite T with $N_T \geq \underline{N}$, at which point mineral energy becomes economically viable and the economy enters the transition regime. Because $\underline{N} < N_O^*$, we have $n_T > 1$: the transition branch of Equation (24) lies strictly above the 45° line, so it contains no rest point, and the population grows

by the constant factor n_T each generation, exiting the transition regime in finite time. Upon exit, $w_M(\bar{N}) = \Omega_T^w > \phi/\nu_n = w_M(N_M^*)$; as the mineral wage is strictly decreasing in population, $\bar{N} < N_M^*$, and the economy grows monotonically to N_M^*.

3. If $\bar{N} < N_0 < N_M^*$, then organic energy cannot cover its cost at any t and the economy grows monotonically to N_M^*.

□

Proposition 3 highlights how an energy transition can arise endogenously as a product of population growth. Given the difference in factor intensities between the two energy sources, population growth makes land progressively scarcer, raising the cost of land-intensive organic energy relative to labor-intensive mineral energy. Once the population reaches $\underline{N}$, mineral energy becomes economically viable and begins to displace organic energy. Proposition 3 also shows that such an energy transition need not occur: if the entry threshold $\underline{N}$ is large, as would be the case if productivity in the mineral energy sector (A_M) is low, the economy reaches the organic steady state before mineral energy ever becomes viable. Moreover, an energy transition, once begun, always runs to completion.

Proposition 4. *If the population grows into the transition regime at some date τ_0, it exits into the mineral regime at a finite date $\tau_1 > \tau_0$, with:*

$$\tau_1 - \tau_0 \leq 1 + \frac{\ln[\bar{N}/\underline{N}]}{\ln n_T} \tag{27}$$

and converges to the mineral steady state N_M^*. The duration of the transition is therefore bounded by a constant that depends only on technologies and preferences, and not on the date or the population at which the transition begins.

Proof. From the proof of Proposition 3, an economy in the organic regime grows past $\underline{N}$ only if $\underline{N} < N_O^*$, which implies $n_T > 1$. Let τ_0 be the first date with $N_{\tau_0} \geq \underline{N}$. While $N_\tau \in [\underline{N}, \bar{N}]$, the population grows by the constant factor n_T, so after k generations in the transition regime $N_{\tau_0+k} = n_T^k N_{\tau_0} \geq n_T^k \underline{N}$. The smallest k for which $n_T^k \underline{N} > \bar{N}$ satisfies $k \leq 1 + \ln[\bar{N}/\underline{N}]/\ln n_T$, which establishes Equation (27). By Proposition 2, the transition regime generically contains no steady state, so the economy cannot come to rest mid-transition; and as established in the proof of Proposition 3, $\bar{N} < N_M^*$ whenever the transition begins, so upon exit the economy grows monotonically to N_M^*. Finally, Equation (21) gives $\bar{N}/\underline{N} = [\psi_M/\lambda_M]/[\psi_O/\lambda_O]$, so the bound in Equation (27) depends only on the factor intensities, preferences, and n_T. □

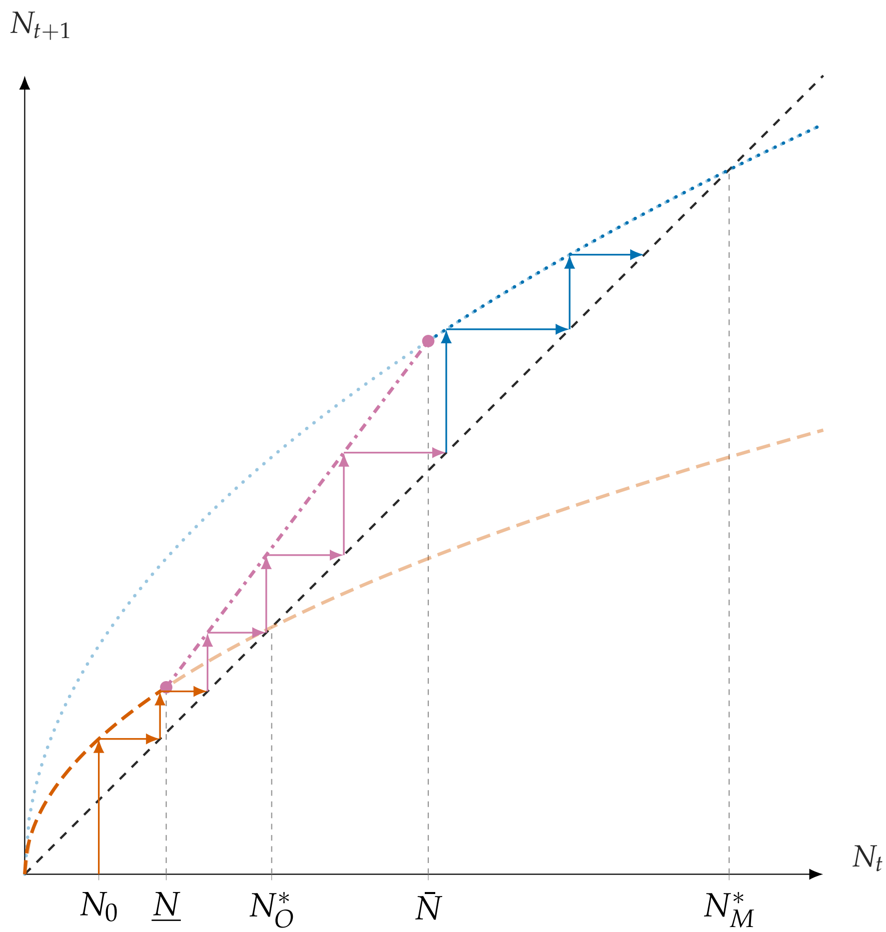


Figure 2: An Energy Transition

Notes: Figure depicts the cobweb dynamics of Equation (24) for an economy that completes an energy transition. Starting from N_0, the economy follows the concave organic branch (orange, dashed). When the population crosses $\underline{N}$, mineral energy enters and the economy moves onto the linear transition branch (purple, dash-dotted), along which the population grows at the constant rate n_T. When the population crosses $\bar{N}$, organic energy exits and the economy follows the concave mineral branch (blue, dotted), converging to the mineral steady state N_M^*. The faded curves depict the organic and mineral branches outside the ranges of population where they apply; the gray dashed line is the 45° line.

Proposition 4 reflects the distinctive economics of the transition regime. Because wages are constant while both energy sources are produced, fertility is constant as well; and because an economy can only grow into the transition regime with a wage above subsistence, population growth continues at a constant rate until organic energy is fully displaced—there is no level of population within the transition regime at which the economy can come to rest. An energy transition is a process that unfolds over a bounded number of generations, and an economy observed using both energy sources is necessarily mid-transition. Once started, energy transitions run to completion.

The dynamics of the energy transition are illustrated in Figure 2. Starting from the initial population N_0, the economy follows the concave dynamics of the organic regime (depicted by the dashed orange curve). If the entry threshold had exceeded the organic steady state population, the economy would have converged to N_O^* along this branch and remained organic. As depicted, however, $\underline{N} < N_O^*$: once the population crosses $\underline{N}$, mineral energy enters and the economy moves onto the linear dynamics of the transition regime (depicted by the dash-dotted purple line), along which the population grows by the constant factor n_T each generation. Once the population crosses $\bar{N}$, organic energy exits, and the economy follows the concave dynamics of the mineral regime (depicted by the dotted blue curve) until it converges to the mineral steady state N_M^*.

An energy transition also leaves a distinct imprint on factor prices. We focus on wages because fertility—and hence population—responds only to the wage; land rents move in the opposite direction but play no direct role in the population dynamics. To trace this imprint, we compare an economy that undergoes an energy transition—entering the transition regime at date τ_0 and exiting into the mineral regime at date τ_1—with a counterfactual economy that is identical in every respect, including its initial population, except that mineral energy is unavailable to it. The counterfactual economy remains organic and converges to N_O^*; we denote its path by $[w_\tau^O, N_\tau^O]$.

Proposition 5. *Consider an economy that undergoes an energy transition and its organic counterfactual. Then:*

1. *The wage path of the transitioning economy is continuous. Wages fall along the organic locus, are constant at Ω_T^w while $N_\tau \in [\underline{N}, \bar{N}]$, and fall along the mineral locus thereafter, with $w_O(\underline{N}) = \Omega_T^w = w_M(\bar{N})$.*

2. *The two populations coincide for all $\tau \leq \tau_0$ and the two wages for all $\tau < \tau_0$. For $\tau \in [\tau_0, \tau_1]$, the wage gap $\ln w_\tau - \ln w_\tau^O$ is nonnegative and strictly increasing and, as a consequence, the population gap $\ln N_\tau - \ln N_\tau^O$ is strictly increasing as well.*

3. *For $\tau \geq \tau_1$, the wage gap decays geometrically:*

$$\ln w_{\tau+1} - \ln w_{\tau+1}^O = [1 - \alpha_F] \left[\ln w_\tau - \ln w_\tau^O\right] \tag{28}$$

   so both economies return to the subsistence wage w^*. The population gap instead converges to its permanent level:

$$\lim_{\tau\to\infty} \left[\ln N_\tau - \ln N_\tau^O\right] = \ln\left[\frac{N_M^*}{N_O^*}\right] = \ln\left[\frac{\bar{N}}{\underline{N}}\right] > 0 \tag{29}$$

Proof. (1) The wage in the transition regime is Ω_T^w by Equation (18). Substituting $\underline{N}$ from

Equation (21) into Equation (12) yields $w_O(\underline{N}) = A_F \alpha_F^{\alpha_F} [1 - \alpha_F]^{[1-\alpha_F]} \Xi^{-\alpha_F} = \Omega_T^w$, and substituting $\bar{N}$ into Equation (15) yields $w_M(\bar{N}) = \Omega_T^w$ by the same algebra.

(2) The two economies share the organic branch of Equation (24) until the transitioning economy first crosses $\underline{N}$ at τ_0, so their populations coincide for $\tau \leq \tau_0$ and their wages for $\tau < \tau_0$. At τ_0, $N_{\tau_0}^O = N_{\tau_0} \geq \underline{N}$ implies $w_{\tau_0}^O = w_O(N_{\tau_0}) \leq \Omega_T^w = w_{\tau_0}$, so the wage gap is nonnegative. For $\tau \in [\tau_0, \tau_1]$, the transitioning economy's wage is constant at Ω_T^w, while the counterfactual economy remains organic: its population continues to grow towards N_O^*, so its wage strictly falls, and the wage gap strictly increases. For the population gap, Equation (23) and Equation (5) give:

$$\ln N_{\tau+1} - \ln N_{\tau+1}^O = \left[\ln N_\tau - \ln N_\tau^O\right] + \left[\ln w_\tau - \ln w_\tau^O\right]$$

so the population gap accumulates past wage gaps and is strictly increasing whenever the wage gap is positive.

(3) In any single-source regime, Equation (24) implies $w_{\tau+1}/w_\tau = [N_{\tau+1}/N_\tau]^{-\alpha_F} = n_\tau^{-\alpha_F}$, and Equation (5) gives $n_\tau = w_\tau / w^*$. Combining the two:

$$\ln w_{\tau+1} - \ln w^* = [1 - \alpha_F] \, [\ln w_\tau - \ln w^*]$$

For $\tau \geq \tau_1$, the transitioning economy is in the mineral regime and the counterfactual economy is in the organic regime, and both obey this law because the two regimes share the same agricultural technology. Differencing across the two economies eliminates w^* and yields Equation (28); since $\alpha_F \in (0, 1)$, the wage gap converges monotonically to zero and both wages converge to w^*. The population gap accumulates these geometrically vanishing wage gaps and therefore converges. Its limit follows from Proposition 3: the transitioning economy converges to N_M^* and the counterfactual economy to N_O^*. Finally, from Equation (25) and Equation (21):

$$\frac{N_M^*}{N_O^*} = \left[\frac{\Omega_M^w}{\Omega_O^w}\right]^{\frac{1}{\alpha_F}} = \frac{\psi_M}{\psi_O} \cdot \frac{\lambda_O}{\lambda_M} = \frac{\psi_M / \lambda_M}{\psi_O / \lambda_O} = \frac{\bar{N}}{\underline{N}}$$

which exceeds one because $\alpha_O > \alpha_M$ implies $\psi_M > \psi_O$ and $\lambda_O > \lambda_M$. □

Proposition 5 traces the full imprint of an energy transition on wages, and thus population in our model. The plateau arises for the same reason factor prices are pinned in the transition regime: with both energy sources produced, the economy absorbs additional workers by shifting the composition of energy production toward the labor-intensive mineral source rather than by bidding down wages. While the transitioning economy's

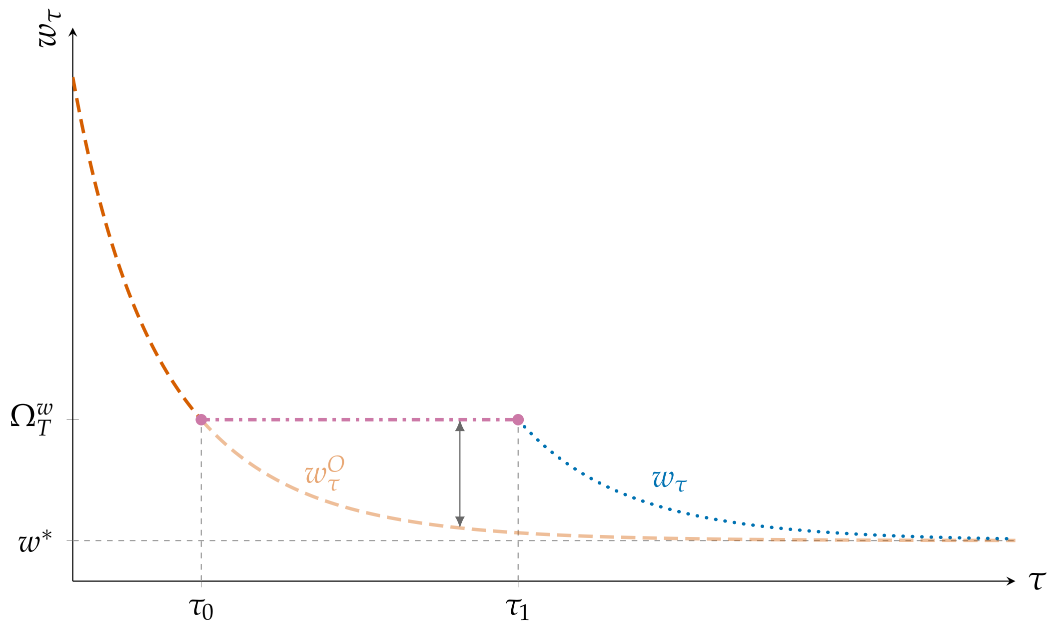


Figure 3: Wages During the Energy Transition: Time Paths.

Notes: Figure plots the wage of an economy that undergoes an energy transition and the wage of its organic counterfactual, w_τ^O (faded orange, dashed), as functions of time. The two paths coincide before τ_0, falling together as the population grows. At τ_0 the transitioning economy enters the transition regime and its wage is held constant at Ω_T^w (purple, dash-dotted) while the counterfactual wage continues to fall: the wage gap (gray arrow) widens. At τ_1 the transition is complete and the transitioning economy's wage falls along the mineral path (blue, dotted); the gap closes geometrically at rate $1 - \alpha_F$ per generation as both economies converge to the subsistence wage w^*.

wage is held at Ω_T^w, its organic counterfactual continues the Malthusian descent, so the wage gap between the two widens generation by generation and, because fertility moves with the wage, the population gap widens with it. Once the transition is complete, the Malthusian logic reasserts itself: population growth in the mineral regime erodes labor scarcity, and the wage gap closes geometrically at rate $1 - \alpha_F$, the same persistence that governs population adjustment in Equation (24). What remains is the population gap: the wage gains of the transition are transitory, but they are converted, one generation at a time, into a population difference that is permanent. The permanent gap in Equation (29) equals the proportional width of the transition regime itself, and the same primitive ratio governs the wage comparison across regimes, $\Omega_M^w / \Omega_O^w = [\bar{N} / \underline{N}]^{\alpha_F}$. This pattern of a transitory wage gap that builds a permanent population gap is the signature of an energy transition in a Malthusian economy.

Figure 3 plots the transitioning economy alongside the organic counterfactual of Proposition 5. The two economies are indistinguishable before the transition begins. Between τ_0 and τ_1, the transitioning economy's wage is held at Ω_T^w while the counterfactual wage continues its descent, so the wage gap between the two widens generation by generation. After τ_1, the gap closes geometrically at rate $1 - \alpha_F$ as the mineral regime carries the transitioning economy back towards subsistence. The wage gap is transitory; what it leaves behind—the population difference accumulated while it was open—is permanent.

3 From Theory to Empirics

While the theory we developed in Section 2 illustrates how an energy transition can arise endogenously as a result of population growth in a pre-industrial economy, this very endogeneity presents a challenge in taking our theory to the data. This is because the entry threshold $\underline{N}$ is unobserved; it depends on primitives that vary across counties in ways that we cannot readily measure given the available data and it is crossed by population growth that is endogenous to those same primitives. Identifying the effects of an energy transition requires an exogenous shock that plausibly shifts $\underline{N}$ for some economies but not others. Here, we address this issue using a change in land ownership that occurred during the Dissolution of the Monasteries as part of the English Reformation. In what follows, we first extend our theory to show how changing land ownership created plausibly exogenous variation in $\underline{N}$ across regions within England and then derive the county-level estimating equation that we take to the data.

3.1 The Dissolution and Appropriability

Our starting point is the classic work of Nef (1932a), who attributes the rapid expansion of English coal mining in the 1600s to the effects of changes in property ownership during the preceding century. Between 1536 and 1540, a series of Parliamentary Acts known collectively as the Dissolution of the Monasteries transferred property from the Catholic Church to the Crown. Although the Dissolution was initially motivated by fiscal considerations, much of the seized monastic land was quickly sold to private landowners to finance Henry VIII's wars (Habakkuk, 1958). The transfer was transformational for the English economy more broadly: parishes affected by the Dissolution were more likely to industrialize by the 1830s, as the transfer stimulated commercialization and facilitated the reallocation of factors of production through markets (Heldring et al., 2021).

Of greater interest for our purposes is Nef's documentation of changes in the contractual terms for mineral extraction that accompanied the transfer of land ownership. Prior to the Dissolution coal deposits were disproportionately located on land owned by the Catholic Church, which had little incentive to invest in mining and offered prospective lessees terms that discouraged them from doing so:

> "If the ecclesiastics had continued unmolested the enjoyment of their original holdings, they would have perhaps hindered the growth of coal mining under Elizabeth and the Stuarts, by retarding the investment of the large capital sums necessary to this growth. The arguments in favour of such a view are three. Religious bodies were not themselves prepared to invest heavily in their own mines. The conditions which they offered lessees were not so favourable as to induce others to invest heavily in them. The Church was not entirely in harmony with the municipal trading class, which controlled the sale of coal, and this opposition sometimes led to a conflict over privileges, which made trade difficult." (Nef, 1932a, pg. 135)

Pre-Dissolution leases were typically short, often of one year, with restrictions on output and high royalty rents (Nef, 1932a, pg. 139). Post-Dissolution leases offered by the Crown and private landowners differed sharply:

> "...the Crown lease had certain well-defined characteristics. Usually drawn for a term of twenty-one years, and never apparently for a shorter period, these leases offered sufficient length of tenure to invite deeper mining, and the investment of capital in new drainage devices. Seldom was a limit put on output. In return for a fine and a fixed annual rent, the tenant commonly acquired an unlimited right to take coal within the territory leased him." (Nef, 1932a, pg. 144)

As Nef documents, the Dissolution of the Monasteries effectively functioned as a change in the appropriation of mineral rents by elites. We summarize this shift via an "appropriability" parameter $\theta \in (0,1]$ equal to the share of revenue from mineral energy that is retained within the local economy. The residual $[1-\theta]$ is appropriated by the elite. Letting the case with Church ownership be denoted by θ_C and the case with ownership by the Crown and private landowners be denoted by θ_P the Dissolution of the Monasteries represents a one-time shift from θ_C to θ_P for the parts of England for which coal lay beneath former Church lands, where, following Nef, $\theta_C < \theta_P$.[8]

[8]Other authors emphasize related but distinct mechanisms. Hatcher (1993), for example, attributes the post-Dissolution expansion to the transfer of land to the gentry rather than to the change in lease terms per se. Our framework accommodates either interpretation: what matters for our argument is that θ rose discretely at the Dissolution.

3.2 The Effects of Changing Appropriability

To formalize the effects of the changes in appropriability that accompanied the Dissolution, we introduce the parameter θ into our model and trace its effect on the energy transition.

Suppose that of every unit of mineral revenue $p_{M,t}X_{M,t}$ generated in the economy, only the fraction θ is retained locally; the remaining $[1-\theta]p_{M,t}X_{M,t}$ is appropriated by the absentee elite and, following our previous assumptions, finances external consumption. The zero-profit condition for the mineral sector then becomes $\theta p_{M,t} = c_{M,t}(\sigma_t, w_t)$, and combining the spending ratio with this modified condition, the labor and land allocations to agriculture in the mineral regime become:

$$\tilde{\psi}_M = \frac{\alpha_F \nu_F}{\alpha_F \nu_F + \alpha_M \theta \nu_E}, \quad \text{and} \quad \tilde{\lambda}_M = \frac{[1-\alpha_F]\nu_F}{[1-\alpha_F]\nu_F + [1-\alpha_M]\theta \nu_E}. \tag{30}$$

Both reduce to their no-appropriation values from Equation (14) when $\theta = 1$.

Appropriability reshapes the transition regime in the same way. With both energy sources produced, households' indifference requires $p_{O,t} = p_{M,t}$; since the mineral sector retains only θ of its revenue, this indifference becomes $c_{M,\tau}(\sigma_\tau, w_\tau) = \theta\, c_{O,\tau}(\sigma_\tau, w_\tau)$. Proceeding exactly as in the proof of Proposition 1, this cost equality pins the rent–wage ratio at a θ-modified value:

$$\frac{\sigma_\tau}{w_\tau} = \tilde{\Xi}(\theta) \equiv \Xi\, \theta^{-\frac{1}{\alpha_O - \alpha_M}}, \tag{31}$$

where Ξ is the no-appropriation ratio from Equation (17). Because $\alpha_O > \alpha_M$, $\tilde{\Xi}$ exceeds Ξ whenever $\theta < 1$ and is strictly decreasing in θ.

The entry and exit thresholds then follow as in Proposition 1, with Ξ replaced by $\tilde{\Xi}(\theta)$ and, at the exit threshold, the mineral allocations replaced by their θ-modified counterparts:

$$\underline{\tilde{N}}(\theta) = \bar{L}\, \frac{1-\alpha_F}{\alpha_F}\, \frac{\psi_O}{\lambda_O}\, \tilde{\Xi}(\theta) \qquad \text{and} \qquad \bar{\tilde{N}}(\theta) = \bar{L}\, \frac{1-\alpha_F}{\alpha_F}\, \frac{\tilde{\psi}_M}{\tilde{\lambda}_M}\, \tilde{\Xi}(\theta). \tag{32}$$

Because the organic allocations do not depend on θ, the entry threshold inherits the monotonicity of $\tilde{\Xi}$: $\partial \underline{\tilde{N}}/\partial\theta < 0$. Higher appropriability lowers the population at which mineral energy first becomes viable, shifting the onset of the energy transition toward smaller populations. As in Equation (22), we also have $\partial \underline{\tilde{N}}/\partial A_M < 0$: an increase in the productivity of mineral energy reduces the entry threshold as well. In this sense, reductions in appropriation by elites mimic the effects of productivity gains, raising the return to economic activity in the mineral energy sector.

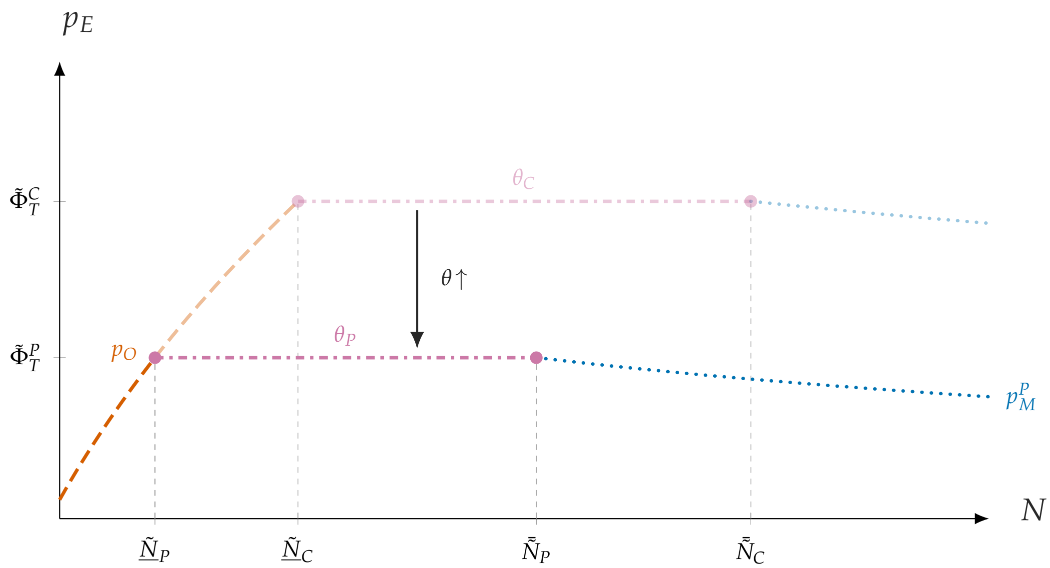


Figure 4: The Effects of Changing Appropriability.

Notes: The figure plots the energy price as a function of population under two appropriability regimes, $\theta_C < \theta_P$, for the case $\alpha_O > \alpha_F > \alpha_M$. Both economies share the rising organic price p_O (orange, dashed) until mineral energy becomes viable. Under private ownership (θ_P, solid), the transition regime runs from $\underline{\tilde{N}}_P$ to $\bar{\tilde{N}}_P$, over which the energy price is constant at $\tilde{\Phi}_T^P$ (purple, dash-dotted), and the economy then follows the declining mineral price p_M^P (blue, dotted). Under Church ownership (θ_C, faded), the higher entry threshold places the transition regime later, from $\underline{\tilde{N}}_C$ to $\bar{\tilde{N}}_C$, and at the higher price $\tilde{\Phi}_T^C$. A rise in appropriability ($\theta \uparrow$) slides the entry point down and to the left along the organic price curve.

Reducing appropriation also raises the economy's long-run population. The mineral-regime steady state can be written as:

$$\tilde{N}_M^* = \left[\frac{\nu_n[1-\alpha_F]A_F}{\phi}\right]^{\frac{1}{\alpha_F}} \left[\frac{\tilde{\psi}_M}{\tilde{\lambda}_M}\right] \bar{L}. \tag{33}$$

By the identity established in Proposition 5, a completed transition raises long-run population by $\ln[\tilde{N}_M^*/N_O^*] = \ln[\bar{\tilde{N}}(\theta)/\underline{\tilde{N}}(\theta)]$ — the proportional width of the transition regime. When agriculture is more land-intensive than mineral energy ($\alpha_F > \alpha_M$), as in our empirical setting, this gain is increasing in θ, so $\partial\tilde{N}_M^*/\partial\theta > 0$. Yet $w^* = \phi/\nu_n$ and the Malthusian trap still hold: the higher return to the mineral sector is eventually eroded by population growth, leaving a larger population at the same subsistence wage.

Figure 4 shows how a change in appropriability reshapes the energy transition in energy-price space. The figure plots the price of energy against population, in the manner of Figure 1, for two appropriability regimes. Because the organic sector is unaffected by θ, both economies share the same rising organic price p_O until mineral energy enters; the energy price then flattens at the constant transition level $\tilde{\Phi}_T$ — the appropriability analog of Φ_T from Equation (19) — and, once organic energy exits, follows the declining mineral price p_M^P. A rise in appropriability from θ_C to θ_P lowers the entry threshold, sliding the entire transition regime down and to the left along the organic price curve: the transition begins earlier ($\underline{\tilde{N}}_P < \underline{\tilde{N}}_C$), ends earlier ($\bar{\tilde{N}}_P < \bar{\tilde{N}}_C$), and clears at a lower energy price ($\tilde{\Phi}_T^P < \tilde{\Phi}_T^C$).

Whether this leftward shift actually triggers a transition depends on the economy's organic steady state. By Propositions 3 and 4, an economy that begins in the organic regime undergoes a complete transition if and only if its entry threshold lies below its organic steady state N_O^*, which is itself independent of θ. A rise in appropriability that pushes the entry threshold below N_O^* thus converts an economy that would have remained organic into one that transitions and converges to a higher mineral steady state. More generally, the effect of the Dissolution depends on where N_O^* lies relative to the two entry thresholds, which we summarize in the following proposition.

Proposition 6. *Consider two appropriability regimes $\theta_C < \theta_P$ with corresponding entry thresholds $\underline{\tilde{N}}_P < \underline{\tilde{N}}_C$, and suppose the economy starts in the organic regime with initial population $N_0 \in (0, N_O^*)$. A change from θ_C to θ_P:*

1. *has* ***no effect*** *if $N_O^* \leq \underline{\tilde{N}}_P$: the entry threshold exceeds the organic steady state under both regimes, so the economy remains trapped in the organic regime.*

2. ***triggers an energy transition*** *if $N_0 < \underline{\tilde{N}}_P < N_O^* < \underline{\tilde{N}}_C$: the economy is trapped in the organic regime under θ_C, but under θ_P it crosses $\underline{\tilde{N}}_P$ and undergoes a complete transition to the mineral steady state. Long-run population rises.*

3. ***accelerates an energy transition*** *if $N_0 < \underline{\tilde{N}}_C \leq N_O^*$: the economy transitions under either regime, but the change to θ_P moves its onset to a smaller population and raises the mineral steady state. Long-run population rises further.*

4. *leaves the economy in a* ***completed transition*** *if $\underline{\tilde{N}}_C < N_0$: the economy has already entered the transition under both regimes and converges to the mineral steady state, which is higher under θ_P.*

Proof. Each case applies Proposition 3 and 4 to the entry threshold $\underline{\tilde{N}}(\theta)$, using $\underline{\tilde{N}}_P < \underline{\tilde{N}}_C$ (from $\partial\underline{\tilde{N}}/\partial\theta < 0$) and the fact that the organic steady state N_O^* is independent of θ. The long-run population comparisons follow from $\partial\tilde{N}_M^*/\partial\theta > 0$. □

Proposition 6 highlights how the Dissolution of the Monasteries created plausibly exogenous variation in the onset of the energy transition across regions that can be exploited to identify the long run effects of an energy transition within England. Taken together, Parts 2–4 of the Proposition show that regions where coal was sufficiently productive were *treated* by the Dissolution. Part 1, in contrast, identifies a natural control group: regions where coal was either of sufficiently low productivity, or not physically available.[9] As such, Proposition 6 suggests the Dissolution of the Monasteries can be used as the basis of a difference-in-difference type research design that compares regions where coal was available on Church land, with regions where it was not, before and after the Dissolution occurred.

3.3 The Estimating Equation

To implement this research design we make further use of our theory. To start, note that Equation (24) implies that when an economy is in the organic regime, its population density evolves according to:

$$\ln\left[\frac{N_{\tau+1}}{\bar{L}}\right] = \ln\left[[1-\alpha_F]\left[\frac{\psi_O}{\lambda_O}\right]^{\alpha_F}\right] + [1-\alpha_F]\ln\left[\frac{N_\tau}{\bar{L}}\right] + \ln A_F + \ln\left[\frac{\nu_n}{\phi}\right] \tag{34}$$

where we have made use of the fact that $\Omega_O^w = [1-\alpha_F]A_F[\psi_O/\lambda_O]^{\alpha_F}$. Similarly, when an economy is transitioning, its population density evolves according to:

$$\ln\left[\frac{N_{\tau+1}}{\bar{L}}\right] = \ln\left[\alpha_F^{\alpha_F}[1-\alpha_F]^{[1-\alpha_F]}\Xi^{-\alpha_F}\right] + \ln\left[\frac{N_\tau}{\bar{L}}\right] + \ln A_F + \ln\left[\frac{\nu_n}{\phi}\right] \tag{35}$$

where we have substituted for $\Omega_T^w = A_F\,\alpha_F^{\alpha_F}[1-\alpha_F]^{[1-\alpha_F]}\,\Xi^{-\alpha_F}$.

Let $R_\tau = \mathbf{1}[N_\tau \geq \underline{N}]$ be an indicator equal to one if the economy has started its energy transition as of generation τ. Then the dynamics of the economy can be written compactly as:

$$\begin{aligned}\ln\left[\frac{N_{\tau+1}}{\bar{L}}\right] &= \ln\left[[1-\alpha_F]\left[\frac{\psi_O}{\lambda_O}\right]^{\alpha_F}\right] + [1-\alpha_F]\ln\left[\frac{N_\tau}{\bar{L}}\right] + \ln A_F + \ln\left[\frac{\nu_n}{\phi}\right] \\ &\quad + R_\tau\left[\ln\left[\alpha_F^{\alpha_F}[1-\alpha_F]^{[1-\alpha_F]}\Xi^{-\alpha_F}\right] - \ln\left[[1-\alpha_F]\left[\frac{\psi_O}{\lambda_O}\right]^{\alpha_F}\right] + \alpha_F\ln\left[\frac{N_\tau}{\bar{L}}\right]\right]\end{aligned} \tag{36}$$

Letting i index counties, as this is the finest level of geographic detail at which we

[9] The physical absence of coal can be thought of as the case where $A_M \to 0$, so that $\underline{\tilde{N}} \to \infty$.

have population estimates, this suggests the following estimating equation:

$$\ln\left[\frac{N_{i,\tau}}{\bar{L}_i}\right] = \kappa_0 + \kappa_1 \ln\left[\frac{N_{i,\tau-1}}{\bar{L}_i}\right] + \left[[\kappa_2 - \kappa_0] + [\rho - \kappa_1]\ln\left[\frac{N_{i,\tau-1}}{\bar{L}_i}\right]\right] R_{i,\tau-1} + \gamma X_i \eta_\tau + \mu_\tau + \varepsilon_{i,\tau} \quad (37)$$

where $\kappa_0 = \ln[[1-\alpha_F][\psi_O/\lambda_O]^{\alpha_F}]$ and $\kappa_2 = \ln\left[\alpha_F^{\alpha_F}[1-\alpha_F]^{[1-\alpha_F]}\Xi^{-\alpha_F}\right]$ are constants, $\kappa_1 = [1-\alpha_F]$, $\mu_\tau = \ln[\nu_n/\phi]_\tau$, $\varepsilon_{i,\tau}$ is a mean-zero error term, and we have introduced the parameter ρ to capture population persistence during the energy transition. While our model implies $\rho = 1$, adopting this parameterization allows us to test this assumption empirically. As such, κ_1 and ρ are the primary coefficients of interest in Equation (37); they allow us to empirically test whether the dynamics of population growth change as a result of the transition from wood to coal, specifically $\rho = 1$ and $\kappa_1 < \rho$. Three additional aspects of Equation (37) also bear mention.

First, $R_{\tau-1}$ is also indexed by i to capture the fact that, as Proposition 6 shows, each county's transition to mineral energy depends on its underlying change in appropriability. Second, we have replaced the constant agricultural productivity parameter $\ln A_F$ with the interaction $X_i\eta_\tau$ that is indexed by both county and generation. We include the variable X_i to capture the fact that, in reality, counties varied in their underlying agricultural productivities due to factors such as geography. We interact X_i with generation fixed effects (η_τ) to allow agricultural productivity to evolve flexibly across generations, reflecting factors such as changes in farming techniques or other related forms of technological progress in the agricultural sector.[10] Third, while our model takes the parameters governing demographics ν_n and ϕ as constant, we have indexed μ by τ in Equation (37) to allow for systematic changes in preference for children or the costs of raising children across generations.

Two complications arise when Equation (37) is taken to the data. The first is that we do not observe $R_{i,\tau-1}$; that is, we do not know when each county first began to transition from wood to coal as, to the best of our knowledge, such information is not available from the historical record. The second is that estimating Equation (37) requires population data by county and generation both before and after the Dissolution of the Monasteries. Such data are not available for our period of study. Instead, as we discuss further below, we have population estimates by county for the years 1086, 1290, 1377, 1600, 1700, 1750, and 1801. Given that the demographers put the average generation

[10]Existing research, such as that of Allen (2000); Clark (2002); Allen (2005), suggests there were significant changes in agricultural productivity during our period of study.

length during our period of study at around 30 years, this means each of the intervals in our data spans more than one generation.

To address the fact that $R_{i,\tau-1}$ is unobserved, we exploit the intuition from Proposition 6 and proxy for $R_{i,\tau-1}$ using the interaction $[\text{Coal}_i \times \text{Post}_{\tau-1}]$, where Coal_i is an indicator of Church ownership of coal-bearing land in county i that captures each county's exposure to the change in appropriability that occurred during the Dissolution of the Monasteries, and $\text{Post}_{\tau-1}$ is an indicator equal to one if the preceding generation was after the Dissolution. This proxy is consistent with the model under two assumptions.

Assumption 1: Counties with coal on Church land are initially in the Organic Regime. With Church land ownership, θ_C, all coal-bearing counties must be in the organic regime: $R_{i,\tau-1} = 0$ for all i with coal, and all generations prior to 1535. This requires $N_{i,\tau} < \tilde{N}_{i,C}$ for all $\tau \leq 1535$.

Assumption 2: All coal counties eventually transition: With private land ownership θ_P, treated coal-bearing counties eventually transition. This requires $\tilde{N}_{i,P} < N^*_O$.

Under these assumptions, substituting $R_{i,\tau-1}$ with $[\text{Coal}_i \times \text{Post}_{\tau-1}]$ gives us a revised estimating equation:

$$\ln\left[\frac{N_{i,\tau}}{\bar{L}_i}\right] = \kappa_0 + \kappa_1 \ln\left[\frac{N_{i,\tau-1}}{\bar{L}_i}\right] + \left[[\kappa_2 - \kappa_0] + [\rho - \kappa_1]\ln\left[\frac{N_{i,\tau-1}}{\bar{L}_i}\right]\right] \times [\text{Coal}_i \times \text{Post}_{\tau-1}] + \gamma X_i \eta_\tau + \mu_\tau + \varepsilon_{i,\tau} \quad (38)$$

We address the fact that each period in our data spans more than one generation via the recursive structure of our estimating equation. Let t denote calendar time, g denote the length of a generation and Δ_t denote each interval in our data. Then iterating over any interval when a county is in the organic regime yields:

$$\ln\left[\frac{N_{i,t}}{\bar{L}}\right] = \kappa_1^{h_t} \ln\left[\frac{N_{i,t-1}}{\bar{L}}\right] + \Gamma_t X_i + \chi_t + \epsilon_{i,t} \quad (39)$$

where $h_t = \lfloor \Delta_t / g \rfloor$[11] and:

$$\chi_t \equiv \sum_{j=0}^{h_t-1} \kappa_1^j [\kappa_0 + \mu_j], \qquad \Gamma_t \equiv \gamma \sum_{j=0}^{h_t-1} \kappa_1^j \eta_j, \qquad \epsilon_{i,t} \equiv \sum_{j=0}^{h_t-1} \kappa_1^j \varepsilon_{i,j}.$$

[11] Here $\lfloor\cdot\rfloor$ is the floor operator so h_t is the number of complete generations contained in interval Δ_t given generation length g.

Similarly, iterating over any interval when a county is undergoing an energy transition yields:

$$\ln\left[\frac{N_{i,t}}{\bar{L}}\right] = \rho^{h_t}\ln\left[\frac{N_{i,t-1}}{\bar{L}}\right] + \tilde{\Gamma}_t X_i + \tilde{\chi}_t + \tilde{\epsilon}_{i,t} \tag{40}$$

where:

$$\tilde{\chi}_t \equiv \sum_{j=0}^{h_t-1}\rho^j[\kappa_2+\mu_j], \qquad \tilde{\Gamma}_t \equiv \gamma\sum_{j=0}^{h_t-1}\rho^j\eta_j, \qquad \tilde{\epsilon}_{i,t} \equiv \sum_{j=0}^{h_t-1}\rho^j\varepsilon_{i,j}.$$

Thus, combining Equations (39) and (40) gives an estimating equation expressed in calendar time:

$$\begin{aligned}\ln\left[\frac{N_{i,t}}{\bar{L}}\right] &= \kappa_1^{h_t}\ln\left[\frac{N_{i,t-1}}{\bar{L}}\right] \\ &\quad + \left[\Delta\chi_t + [\rho^{h_t}-\kappa_1^{h_t}]\ln\left[\frac{N_{i,t-1}}{\bar{L}}\right] + \Delta\Gamma_t X_i\right] \times [\text{Coal}_i \times \text{Post}_t] \\ &\quad + \Gamma_t X_i + \chi_t + \epsilon_{i,t}\end{aligned} \tag{41}$$

where $\Delta\chi_t \equiv \tilde{\chi}_t - \chi_t$, $\Delta\Gamma_t \equiv \tilde{\Gamma}_t - \Gamma_t$ and Post_t is an indicator equal to one in the post 1600 period.

While it is possible to estimate Equation (41) directly, it is convenient to redefine terms such that population persistence is expressed in terms of annual, rather than generational changes as h_t depends on our choice of g. Let $r \equiv \kappa_1^{1/g}$, and $\varrho \equiv \rho^{1/g}$. Then Equation (41) can be rewritten as:

$$\begin{aligned}\ln\left[\frac{N_{i,t}}{\bar{L}}\right] &= r^{\Delta_t}\ln\left[\frac{N_{i,t-1}}{\bar{L}}\right] \\ &\quad + \left[\Delta\chi_t + [\varrho^{\Delta_t}-r^{\Delta_t}]\ln\left[\frac{N_{i,t-1}}{\bar{L}}\right] + \Delta\Gamma_t X_i\right] \times [\text{Coal}_i \times \text{Post}_t] \\ &\quad + \Gamma_t X_i + \chi_t + \epsilon_{i,t}\end{aligned} \tag{42}$$

Equation (42) can be estimated directly using nonlinear least squares. Clearly, our main parameters of interest can be recovered given a choice of g. Given this, we can test the restrictions implied by our model directly using the estimates $\hat{r}$ and $\hat{\varrho}$. If England's transition to coal changed the dynamics of population growth, as our model predicts, then $\hat{\varrho} = 1$, and $\hat{r} < \hat{\varrho}$. In what follows, we cluster standard errors by county to address potential serial correlation (Bertrand et al., 2004).

It is worth emphasizing that Equation (42) embeds the logic of the difference-in-difference-type research design suggested by Proposition 6: it compares changes in average population density across "treated" counties that have coal deposits on Monastic land with "control" counties that do not before and after the Dissolution of the Monasteries. However, unlike a traditional difference in difference formulation in which identification relies on a parallel trends assumption, here identification requires conditional independence, namely that treatment ($\text{Coal}_i \times \text{Post}_t$) is as good as randomly assigned, conditional on lagged population density, period fixed effects and additional controls.

It is also worth noting that we can also use the estimates from Equation (42) to measure the cumulative effect of the transition to coal on the population density of affected counties. To see this, consider a treated county i and compare its realized population density to the counterfactual density that would have occurred had it not been exposed to the change in appropriability brought about by the Dissolution. That is, consider the same county with Coal_i set to zero, holding all else fixed. After differencing the two paths implied by Equation (42), the effect of the energy transition at time t, the estimated gap $\hat{G}_{i,t} \equiv \ln[N^T_{i,t}/\bar{L}_i] - \ln[N^O_{i,t}/\bar{L}_i]$, is given by:

$$\hat{G}_{i,t} = \hat{r}^{\Delta_t}\hat{G}_{i,t-1} + \left[\widehat{\Delta\chi}_t + \left[\hat{\varrho}^{\Delta_t} - \hat{r}^{\Delta_t}\right]\ln\left[\frac{N_{i,t-1}}{\bar{L}_i}\right] + \widehat{\Delta\Gamma}_t X_i\right], \tag{43}$$

where the bracketed term is the change in the gap between t and $t-1$, evaluated at the county's observed density at $t-1$. Two features of Equation (43) are worth mentioning. First, the previous period's gap is carried forward at the *untreated* persistence $\hat{r}^{\Delta_t}$: because the factual and counterfactual lags enter Equation (42) with different coefficients, the residual slope contrast is absorbed into the increment rather than the carry-forward term. Second, the change between t and $t-1$ is evaluated using observed lagged density, which from the second post-Dissolution period onward already embeds the effect of prior exposure. Thus, measuring the cumulative effect of the energy transition requires the recursion in Equation (43), not a running sum of per-interval changes.

Iterating Equation (43) forward in the post-Dissolution period yields the cumulative effect at any time t:

$$\widehat{CG}_{i,t} = \sum_{s\,:\,1600<\text{yr}_s\leq\text{yr}_t} \hat{r}^{\,\text{yr}_t-\text{yr}_s}\left[\widehat{\Delta\chi}_s + \left[\hat{\varrho}^{\Delta_s} - \hat{r}^{\Delta_s}\right]\ln\left[\frac{N_{i,s-1}}{\bar{L}_i}\right] + \widehat{\Delta\Gamma}_s X_i\right]. \tag{44}$$

where we have used $\prod_m r^{\Delta_{t_m}} = r^{\text{yr}_t-\text{yr}_s}$. To calculate the cumulative effect for treated counties, we evaluate Equation (44) at the treated-county means of lagged density and caloric suitability in each period, using our estimates of $\hat{r}$ and $\hat{\varrho}$. In what follows, we compute standard errors for $\widehat{CG}_{i,t}$ via the delta method.

3.4 Data and Measurement

Our empirical analysis requires county level data on population density, access to coal, and agricultural productivity. We obtain these data from a variety of sources.

Constructing measures of population density is complicated by the fact that most historical estimates of English population levels are at the national level; there are relatively few historical estimates of population by county. As such, we construct our population data from the two available sources that report county-population estimates spanning a number of years: Broadberry et al. (2015) and Wrigley (2009). Broadberry et al. report estimates for 1086, 1290, 1377 and 1600, while Wrigley reports estimates for 1600, 1700, and 1750.[12] Given that two series overlap in 1600, we use the average from the two datasets as the estimate for that year in our analysis.[13]

Constructing a measure of Coal_i, that is, a measure of the importance coal deposits on Monastic land owned by the Catholic Church is more straightforward. We first use a shapefile on coal deposits from Fernihough and O'Rourke (2021) and a shapefile on English parish boundaries from Satchell et al. (2023) to determine which parishes in England contained coal deposits. We then use the data from Heldring et al. (2021) to determine which parishes were Monastic. Aggregating to the county level then gives us the amount of land containing coal in each county, as well as the share of this land that was controlled by the Church.[14] In our baseline regressions we define Coal_i to be equal to one if the county contained coal deposits that were located on Monastic land, and zero otherwise.[15] Doing so gives us fifteen "coal counties" that have coal on Monastic land and could have been affected by the Dissolution of the Monasteries.

Measuring agricultural productivity is complicated by the fact that we do not observe agricultural output or input use by county consistently throughout our period of

[12] Wrigley (2009) also includes county-level estimates for 1801. For most of our analysis we do not make use of these data to ensure that we are not capturing the effects the technological change created by Watt's steam-engine.

[13] Doing so has little impact on the baseline results we present below. For brevity, we do not include these results here, but they available on request.

[14] It is worth noting that the county of Kent contains positive measures of coal using this approach. However, we reassign Kent to have zero values because the historical record shows that the relevant subterranean coal deposit in Kent that yields positive measures was not discovered via exploratory boring until the late 1800s.

[15] We examine the effects of adopting a continuous measure that reflects the share of county i's coal that is located on monastic land in our empirical analysis.

study. This means we are unable to estimate agricultural productivity using standard methods. As such, we follow the approach taken elsewhere in the literature (e.g. Nunn and Qian (2011); Cherniwchan and Moreno-Cruz (2019)) and use a time-invariant measure of physical agricultural productivity interacted with time fixed effects as a proxy. We construct this agricultural productivity measure using raster data on caloric suitability from Galor and Özak (2016), who construct measures of the average and maximum potential caloric yield for $5' \times 5'$ grid cells across the globe given the sets of crops that were available both before and after the Columbian Exchange.[16] Here, we rely on the average potential yields for pre-exchange crops and measure agricultural productivity as the median value across cells in each county.[17] However, since agricultural productivity also potentially changed during our period of study due to the introduction of new crops during the Columbian Exchange, we also include measures of the change in caloric suitability in our analysis below.

A final issue for analysis concerns the fact that the fifteen counties that contained coal deposits on Church land were not the only ones potentially affected by the change in appropriability that accompanied the Dissolution. While coal was very difficult to transport over land – the upper estimates by historians indicate that it was prohibitively costly to transport coal more than 20 miles (32 kilometers) from a mine – some coal reserves were located in close proximity to the ocean, greatly lowering the transport costs associated with moving coal over long distances. For example, shipping coal three hundred miles by sea from Newcastle in Northumberland to London in Middlesex cost the same as transporting coal three to four miles over land (Freese, 2003). The historical record suggest that this trade included a number of counties; Nef (1932b) documents coal imports shortly after the Dissolution at ports in the counties of Devon, Essex, Hampshire, Lincolnshire, Middlesex, Norfolk and Suffolk.

Such imports are problematic because our research design requires that access to coal be as good as randomly assigned, conditional on lagged population and the other controls. This identification assumption is unlikely to hold for counties that are importers, as the choice to import is endogenous. The historical record suggests this endogeneity is material. For example, consider the case of Middlesex, which imported coal via coastal trade from Newcastle. This trade, was itself a result of metropolitan demand: Wrigley (1967) identifies "the large and steadily growing demand for coal afforded by the London coal market" as one of the "chief supports" of the expanding coal industry,

[16]We use the caloric suitability data from Galor and Özak (2016) rather than the agricultural suitability data from IIASA/FAO (2012), as the former better captures the potential energy produced by agriculture on a given plot of land. See Galor and Özak (2015) for further discussion.

[17]In the online appendix, we show that our baseline results do not depend on this choice.

Table 1: Summary Statistics

	Non-Coal Counties			Coal Counties		
	Mean	Std. Dev	Median	Mean	Std. Dev	Median
Population (1,000s)	70.33	41.19	66.16	102.57	92.36	78.31
County Area (sq. miles)	869.50	377.34	780.42	1,537.52	1,263.87	1,133.94
Population Density	82.81	36.23	84.45	71.62	39.72	72.23
Coal Area (sq. miles)	0.00	0.00	0.00	420.81	301.76	311.73
Church Coal Share (%)	0.00	0.00	0.00	31.39	17.82	32.36
Median CSI (1,000s)	4.08	0.41	4.25	3.78	0.28	3.80
Obs.	102			90		
No. Counties	17			15		

Notes: Table reports key summary statistics for the two groups of counties considered in our analysis. The unit of observation is county×year. See text for further details.

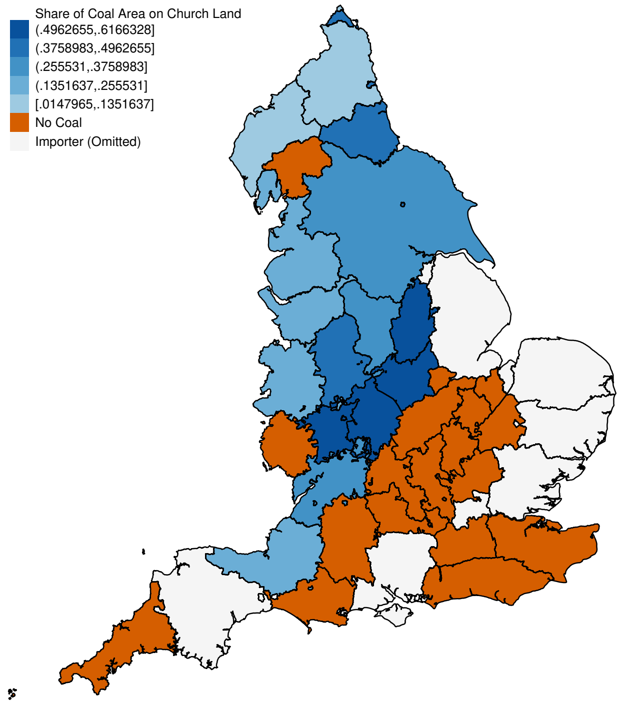
and the tonnage shipped down the east coast to London roughly doubled in the century before 1750. This creates a problem of reverse causality: imports of coal are as much of a product of population growth as they are a cause of it. To address this issue, we exclude the "early" importers of coal identified by Nef from our analysis.

With our exclusion of importing counties, after combining our data sources, we are left with data for 32 counties for the years 1086, 1290, 1377, 1600, 1700 and 1750. We report summary statistics in Table 1, which divides our sample into coal counties that have coal deposits and non-coal counties that do not. As the Table shows, coal counties have larger populations, but are also larger in area, than non-coal counties, leading to slightly lower population densities on average. Coal counties also have lower agricultural productivity levels than non-coal counties. Finally, the table shows that the average coal county has coal deposits that cover close to 421 square miles of land; of this over 31% lies on land controlled by the Catholic Church, suggesting there is potential scope for the Dissolution of the Monasteries to impact coal extraction.[18]

The geographic variation in coal located on land controlled by the Catholic church for the counties in our analysis is depicted in Figure 5. The fifteen counties with coal on Monastic land are depicted in blue. The seventeen non-coal counties that do not contain coal deposits are depicted in orange. Importing counties, which are omitted from our analysis, are depicted in white. As the figure shows, counties endowed with coal lie in the north or close to Wales, while counties without coal are primarily located in the south. The figure also highlights what share of each county's coalfield area is located on

[18]While it is not immediately clear from Table 1, for all of the counties in our analysis that contain coal, some of their coalfield lies on Church land.

Figure 5: The Fraction of Coal Fields on Church Land by County

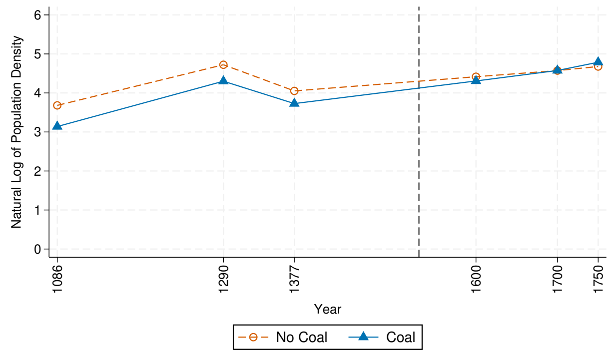


Notes: Figure depicts share of each county's coalfield area that is located on land controlled by the Catholic Church prior to the Dissolution of the Monasteries. The counties in blue are endowed with coal on Monastic land, while counties in orange lack coal deposits. Counties that were early importers of coal (excluded from our sample) are depicted in white.

land controlled by the Catholic Church; as the figure shows, counties with coal in the Midlands have a much higher share of coalfields located on Monastic land than those in the North East or North West regions of England.

While the summary statistics presented in Table 1 suggests that coal counties were less dense than non-coal counties on average, this difference masks significant changes in population density over time. This can be seen in Figure 6, which plots the natural log of average population density for counties with (depicted by the solid blue line) and without (depicted by the dashed orange line) coal over our period of study. As this Figure shows, the average population density of coal and non-coal counties tracked each other closely prior to the Dissolution of the Monasteries (depicted by the vertical dashed grey line), suggesting the two groups of counties exhibited similar population dynamics for the first three periods in our sample. Following the Dissolution, the population density of coal counties appears to have grown substantially faster than non-coal counties; as of 1700, the population density of the average coal county exceeds that of the average non-coal county. This evidence is strongly suggestive that coal counties grew more dense relative to the non-coal counties following the Dissolution of the Monasteries, as

Figure 6: Average Population Density With and Without Coal: 1086-1750



Notes: Figure reports the natural log of average population density (in persons per square mile) in counties with and without coal for each year in our data.

our theory would suggest. We now turn to investigate this energy transition formally.

4 Empirical Results

We begin our empirical analysis by estimating six specifications based on Equation (42). The results of this exercise are presented in Table 2. Column (1) reports the simplest possible specification: it assumes that agricultural productivity is constant, and is thus embedded in the intercept, κ_0. Column (2) is our preferred specification and allows for differential trends in agricultural productivity by interacting average caloric suitability index from Galor and Özak (2016) with time fixed effects.[19] Column (3) replicates our preferred specification, but replaces the indicator Coal_i with the share of county i's coal deposits that are located on Monastic land. Column (4) imposes the restriction $\rho = 1$. Columns (5) and (6) allow ρ to vary across periods. Column (6) also expands our sample to include observations from 1801, the next period for which we have population estimates by county. Throughout, standard errors clustered by county are reported in round parentheses. For each specification, we also report p-values from a number of hypothesis tests, including tests of the restrictions $\hat{\varrho} = 1$ and $\hat{\varrho} = r$ implied by our model. We also report estimates of $\hat{\kappa}_1$ and $\hat{\rho}$, calculated under the assumption that $g = 30$, as well as the estimated cumulative effect of the energy transition at the end of the sample period. For each of these estimates, we also report standard errors calculated

[19]Throughout, we refer to this set of interactions as "CSI $\times$ Time FE" for caloric suitability index $\times$ time fixed effects.

by the delta method in square parentheses.[20]

The estimates reported in Table 2 suggest that the transition to coal significantly altered the population growth of counties that had access to it relative to those that did not. Consider our preferred estimates in column (2). We find that $\hat{r}$ is significantly below one, meaning that in counties without access to coal, populations grow toward their prevailing organic steady state. Population dynamics are starkly different in counties with coal: for these counties, we cannot reject the hypothesis that $\varrho = 1$ and we reject the hypothesis that $\varrho = r$, the null of no change in dynamics. These results suggest that the energy transition led to a sustained increase in population growth, as predicted by our theory. The remaining coefficients are also informative. Our estimates of the $\widehat{\Delta\chi}_t$ suggest that the transition to coal did not lead to a level shift in population growth, but our estimates of the $\widehat{\Delta\Gamma}_t$ suggest that ongoing changes to agricultural productivity also impacted the effects of coal. All together the estimates reported in Column (2) suggest that the transition to coal significantly increased the population density of affected counties. The cumulative effect reported in the last row of the table indicates that, by 1750, the transition to coal had raised the population density of affected counties by 0.248 log points relative to their organic level, an increase of roughly 28%.

As the remaining columns of in Table 2 show, the results presented in Column (2) are quite robust. Two additional findings from Table 2 are worth highlighting. The first is that the conclusions from our main specification are robust to the adoption of a continuous measure of treatment. As the estimates reported in Column (3) show, when we adopt a continuous measure of treatment that better captures cross-county differences in exposure to the Dissolution of the Monasteries, our estimates of r and ϱ are essentially unchanged. Adopting a continuous measure does reduce our estimate of the cumulative effect of the energy transition by 1750, but as these estimates are relatively imprecise, it is not statistically different than the estimate from our preferred specification. The second finding worth highlighting is that the change in population dynamics from the energy transition appear to have continued to persist in the long run. As the estimates in Columns (5) and (6) show, our estimates of ϱ are essentially unchanged by 1801, suggesting that affected counties had not completed transitioning to the mineral regime before the start of the Industrial Revolution. However, our estimates of $\widehat{\Delta\chi}$ and $\widehat{\Delta\Gamma}$ for 1801 from Column (6) also suggest that the underlying dynamic process governing population growth also began to change after the invention of the commercial

[20]It is also worth highlighting that population density for 1086 enters as the initial lag in our estimating equation, so each specification using our 1086-1750 sample has 160 observations rather than the 192 reported in Table 1.

steam engine.

4.1 Alternative Explanations

As the second step in our empirical analysis, we investigate several alternative explanations for our findings. The results from this exercise are reported in the eight columns of Table 3. For the sake of brevity, here we only report estimates of r and ϱ. For convenience, Column (1) reproduces our preferred estimates from Column (2) of Table 2. As in this specification, the specifications reported in the remaining columns of the table include year fixed effects and allow for caloric suitability by time fixed effects. Throughout, standard errors clustered by county are reported in parentheses. Each specification also reports estimates of $\hat{\kappa}_1$ and $\hat{\rho}$ calculated assuming a generation length $g = 30$ and the cumulative effect of coal, as well as their associated standard errors calculated via the Delta method (in square parentheses).

We begin by considering whether our estimates are capturing differential trends in transport costs across ancient counties. As we noted above, coal was notoriously difficult to transport over long distances. This led to the development of a large transportation network to reduce transport costs and move coal to final consumers (e.g. Allen (2023)). As such, our estimates could be capturing the effects of changes in transport costs as opposed to the effects of the energy transition. Since such changes are endogenous and comprehensive data on transportation infrastructure is not available throughout our period of study, we instead rely on the Human Mobility Index of Özak (2010, 2018). Specifically, we investigate whether our main estimates are capturing the effects of differential changes in transport costs by supplementing our preferred specification with a measure of county isolation, calculated as the average travel time to major county towns across all parishes in a county, interacted with time fixed effects to allow for differential trends in the relative ease of moving coal and other goods over space. As shown in Column (2) of Table 3, doing so has little effect on estimates, suggesting our preferred specification is not capturing differential trends in transport costs across counties.

We next turn to examine whether our results are potentially capturing the effects of migration to major urban centers. While the available evidence from the historical record suggests that long-distance migration was limited during our period of study – less than 5% of people moved more than 100 miles (Clark, 1979) – London's growth was sustained through in-migration from the rest of the country, particularly many of the non-coal counties that serve in our control group. As major ports, Bristol and Liverpool also experienced similar in-migration, although on a smaller scale. Such migration would

Table 2: The Effects of Coal on Population Density

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{r}$	0.9959	0.9928	0.9930	0.9928	0.9928	0.9933
	(0.0005)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0016)
$\hat{\varrho}$	0.9997	1.0002	1.0053			
	(0.0014)	(0.0017)	(0.0043)			
$\hat{\varrho}_{1700}$					1.0002	1.0002
					(0.0016)	(0.0016)
$\hat{\varrho}_{1750}$					1.0004	1.0004
					(0.0029)	(0.0029)
$\hat{\varrho}_{1801}$						1.0008
						(0.0040)
$\widehat{\Delta\chi}_{1700}$	-1.239	-0.958	-1.044	-0.921	-0.950	-0.863
	(0.571)	(0.723)	(3.781)	(0.674)	(0.732)	(0.733)
$\widehat{\Delta\chi}_{1750}$	-0.664	-1.059	-3.972	-1.032	-1.077	-1.017
	(0.306)	(0.787)	(2.466)	(0.761)	(0.801)	(0.804)
$\widehat{\Delta\chi}_{1801}$						-1.511
						(0.921)
$\widehat{\Delta\Gamma}_{1700}$		-0.323	-0.995	-0.309	-0.320	-0.315
		(0.167)	(0.468)	(0.139)	(0.166)	(0.165)
$\widehat{\Delta\Gamma}_{1750}$		-0.052	0.397	-0.046	-0.056	-0.051
		(0.173)	(0.489)	(0.171)	(0.183)	(0.182)
$\widehat{\Delta\Gamma}_{1801}$						0.026
						(0.216)
Time FE	X	X	X	X	X	X
CSI$\times$ Time FE		X	X	X	X	X
R^2	0.997	0.998	0.997	0.998	0.998	0.998
Obs.	160	160	160	160	160	192
p-value: $\varrho = 1$	0.824	0.903	0.214		0.992	0.995
p-value: $\varrho = r$	0.004	0.000	0.011		0.000	0.001
p-value: ϱ_t equal					0.934	0.990
p-value: $\Delta\chi$ joint	0.094	0.396	0.100	0.381	0.392	0.386
p-value: $\Delta\Gamma$ joint		0.016	0.005	0.000	0.045	0.106
$\hat{\kappa}_1 = \hat{r}^g$	0.883	0.805	0.811	0.805	0.805	0.817
	[0.014]	[0.045]	[0.044]	[0.045]	[0.045]	[0.040]
$\hat{\rho} = \hat{\varrho}^g$	0.990	1.006	1.173			
	[0.043]	[0.052]	[0.150]			
Cumulative effect	0.180	0.248	0.157	0.248	0.248	0.343
	[0.076]	[0.114]	[0.082]	[0.114]	[0.114]	[0.151]

Notes: The dependent variable in all columns is the natural log of county population density. Column (3) replaces $Coal_i$ with the share of county i's coal that is located on Monastic land. Column (4) imposes $\varrho = 1$. Column (5) and (6) allows ϱ to vary across years; Column (6) expands the sample to include 1801. Standard errors clustered by county are shown in parentheses. $\hat{\kappa}_1$ and $\hat{\rho}$ calculated assuming a generation length $g = 30$. Standard errors for $\hat{\kappa}_1$, $\hat{\rho}$ and the cumulative effect of coal calculated via the Delta method are shown in square brackets. Coal importing counties are omitted from the sample.

lead us to overestimate the effects of the energy transition. As we do not observe cross-county migration flows directly, we investigate this possibility by calculating the average distance across each parish to London, Bristol and Liverpool. We interact these measures with time fixed effects to allow for differential trends in the ease of migrating. As the estimates in Columns (3) and (4) show, allowing for such trends has little effect on our estimates, suggesting our baseline results are not capturing the effects of migration to major cities.

We then investigate whether our results are capturing the effects of differential changes in the relative abundance of wood. While the estimating equation we derived from our theory does not explicitly depend on organic energy, the transition between energy sources could also be driven by changes in its price. Indeed, Nef (1932a) also suggests that the rise of coal was in part due to a timber shortage, although this "timber crisis" theory is disputed (e.g. Allen (2003)). Since we do not observe changes in the price of wood directly, we proxy for it by interacting measures of initial relative abundance of wood, constructed from estimates of wooded area in 1086 reported by Rackham (1980), with year fixed effects. The estimates reported in Column (5) show that further controlling for differential trends based on relative wood abundance has little impact on our estimates of r and ϱ, suggesting our results are not capturing the effects of changes in the availability of wood.

We then turn to two other phenomena often linked to the long-run economic development of England in this period: the Dissolution of the Monasteries itself, and the introduction of New World crops. As Heldring et al. (2021) show, the Dissolution of the Monasteries had long-run effects on commercialization and industrialization in affected parishes that are distinct from any effects operating through coal. To ensure we are not capturing these effects, in Column (6) we interact the share of monastic land in each county that was subject to the Dissolution with time fixed effects. As these estimates show, accounting for this channel has little effect on our results. Column (7) then allows for possible changes in agricultural productivity due to the introduction of New World crops as part of the Columbian Exchange, as previous studies (e.g. Nunn and Qian (2011), Chen and Kung (2016), or Cherniwchan and Moreno-Cruz (2019)) have shown their introduction significantly impacted the growth of affected populations. We do so by interacting change in county-level caloric suitability, again drawn from Galor and Özak (2016), with time fixed effects. As the estimates in Column (7) show, accounting for New World crops also has little effect on our results.

Finally, in Column (8), we allow for differential trends in all of these factors simultaneously. This results in a modest decrease in our estimates of r and ϱ and the null

Table 3: Alternative Explanations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{r}$	0.9928	0.9926	0.9919	0.9938	0.9928	0.9924	0.9926	0.9919
	(0.0018)	(0.0019)	(0.0025)	(0.0014)	(0.0016)	(0.0020)	(0.0020)	(0.0033)
$\hat{\varrho}$	1.0002	1.0000	0.9995	0.9992	1.0004	1.0000	1.0002	0.9964
	(0.0017)	(0.0019)	(0.0016)	(0.0020)	(0.0016)	(0.0017)	(0.0018)	(0.0034)
Isolation $\times$ Time FE		X						X
Distance to London $\times$ Time FE			X					X
Distance to Bristol $\times$ Time FE				X				X
Distance to Liverpool $\times$ Time FE				X				X
% Woodland $\times$ Time FE					X			X
Monastic Share $\times$ Time FE						X		X
Post-1500 CSI $\times$ Time FE							X	X
R^2	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999
Obs.	160	160	160	160	160	160	160	160
p-value: $\varrho = 1$	0.903	0.990	0.766	0.702	0.784	0.984	0.904	0.290
p-value: $\varrho = r$	0.000	0.000	0.010	0.009	0.001	0.000	0.000	0.303
$\hat{\kappa}_1 = \hat{r}^g$	0.805	0.801	0.784	0.829	0.804	0.794	0.801	0.783
	[0.045]	[0.046]	[0.059]	[0.035]	[0.039]	[0.049]	[0.048]	[0.079]
$\hat{\rho} = \hat{\varrho}^g$	1.006	1.001	0.986	0.977	1.013	1.001	1.006	0.899
	[0.052]	[0.058]	[0.048]	[0.060]	[0.049]	[0.050]	[0.053]	[0.091]
Cumulative effect	0.248	0.330	0.344	0.329	0.253	0.244	0.273	0.353
	[0.114]	[0.117]	[0.099]	[0.107]	[0.106]	[0.114]	[0.113]	[0.112]

Notes: The dependent variable in all columns is the natural log of county population density. All specifications include time fixed effects and allow for differential trends in caloric suitability. The additional controls indicated in the rows below the regression coefficients are each interacted with time fixed effects. “Isolation” is average remoteness of a county’s parishes, calculated using the Human Mobility Index of Özak (2018), “Distance to London”, “Distance to Bristol” and “Distance to Liverpool” is the average distance to London, Bristol and Liverpool, respectively, from each of a county’s parishes; “Woodland” is initial relative timber abundance constructed from Rackham (1980); “Monastic Share” is the share of monastic land in each county subject to the Dissolution from Heldring et al. (2021); “New Crops” is the change in caloric suitability from the introduction of new crops following the Columbian Exchange from Galor and Özak (2016). Standard errors clustered by county are shown in parentheses. $\hat{\kappa}_1$ and $\hat{\rho}$ calculated assuming a generation length $g = 30$. Standard errors for $\hat{\kappa}_1$, $\hat{\rho}$ and the cumulative effect of coal calculated via the Delta method are shown in square brackets. Coal importing counties are omitted from the sample.

hypothesis $\hat{r} = \hat{\varrho}$ can no longer be rejected at conventional levels, but the estimates from this specification are much less precise, owing to the fact that this specification is estimating fifty-six parameters. Taken together, the estimates reported in Table 3 suggest that our main estimates are not being driven by the effects of differential trends in transport costs, migration, timber availability, the Dissolution of the Monasteries itself, or the introduction of New World crops, but rather the effect of the transition from wood to coal.

4.2 Other Robustness Tests

As the final step in our empirical analysis, we probe the robustness of our results along two additional dimensions. For brevity, we briefly describe these tests here, and relegate a full discussion to the Appendix.

First, we examine whether our main estimates are robust to our choice of agricultural productivity measure. Recall that for the results presented in Table 2, we rely on data from Galor and Özak (2016) and measure agricultural productivity using the median value of average potential caloric yields for crops available prior to 1500 in each county. To ensure our results are not driven by this choice, we re-estimate our preferred specification, but use three alternative measures based on the data reported by Galor and Özak. First, we measure agricultural productivity as the mean value of average potential caloric yields for crops available prior to 1500 in each county. Given Galor and Özak also report data on maximum potential caloric yields, we then measure agricultural productivity as the median and mean values of maximum potential caloric yields for crops available prior to 1500 by county. As the results presented in Section A.1 show, adopting these alternative measures has little effect on our estimates, suggesting that our findings are not driven by how we measure agricultural productivity.

Second, we examine the effects of completely forgoing our empirical approach where we adopt an estimation strategy that is tied tightly to theory and instead adopting a standard difference-in-difference research design that compares the effects of the Dissolution of the Monasteries across counties with and without coal on Monastic land, before and after the dissolution occurred. While doing so no longer allows us to examine the dynamics effects of the transition to coal, it serves as a useful sanity check. As differences-in-differences measures the change in treated counties relative to untreated counties, if our theory is correct we should observe that the population density of coal counties should increase relative to counties without coal after the Dissolution of the Monasteries. As the results presented in Section A.1 show, this is exactly what we find:

the difference-in-difference estimate corresponding to our preferred estimate from Column (2) of Table 2 suggests that, on average, the change in appropriability following the Dissolution of the Monasteries increased the population density of the average county by 25%, and event study estimates show that this difference is not simply due to pre-existing differences in trends across coal and non-coal counties.

5 Conclusion

This paper examines the long-run effects of the first major energy transition in England, the replacement of wood by coal. We develop a simple Malthusian model of an economy with three sectors that draw on land and labor: agriculture, a land-intensive organic energy sector based on wood, and a labor-intensive mineral energy sector based on coal. We show that a transition from one energy source to the other can arise endogenously from population growth alone. As population grows, land becomes scarce relative to labor, raising the cost of the land-intensive organic sector relative to the labor-intensive mineral one, and the economy reallocates land and labor across the three sectors, contracting wood production and expanding coal until coal displaces wood entirely. Such a transition need not occur; but when it does, its lasting effect is on population rather than living standards, as a coal-using economy comes to support a permanently larger population than an otherwise identical economy that does not. Guided by this theory, we take an estimating equation derived directly from the model to population data for the ancient counties of England over the period 1086–1750, exploiting the Dissolution of the Monasteries — which raised the share of mining revenue retained locally in the counties whose coal lay beneath Church land — as a source of identifying variation in the transition from wood to coal.

We find that the transition from wood to coal significantly altered the dynamics of population growth in affected ancient counties. Our preferred estimates indicate that as of 1750, the transition increased the population density of affected counties by close to 28% relative to their densities had the transition not occurred. The change in population dynamics persisted until the start of the Industrial Revolution, and is unlikely to reflect some other coincident change as our estimates are robust to a range of alternative explanations.

Our findings have three primary implications. First, they suggest that an energy transition may alter an economy's dynamics for an extended period of time. As such, it is not clear if the existing evidence on the effects of changes in access to energy resources, drawn almost entirely from horizons of at most a few decades, will be a reliable guide to what a future energy transition ultimately does. Second, our findings suggest that an energy transition may do more than change how much energy an economy uses or how rich it is on average. Because coal was distributed unevenly across England, its transition raised the populations of some counties relative to others, suggesting an energy transition may reshape where economic activity is located. Third, they suggest that institutional change can alter where and when an energy transition occurs. In our model the transition is driven by population growth, but the terms on which a new energy source can be exploited affect when, and whether, it occurs. The Dissolution did not create England's coal; by changing who captured the returns to mining it, it hastened the transition in the places it touched.

Our analysis is subject to at least two limitations. The first is that we study the effects of the transition on population levels rather than on welfare. This is largely due to data limitations; we do not observe wages or incomes at a detailed enough geographic level for a long enough period to measure how England's energy transition affected living standards. The second is that our evidence comes from a single transition in a single country. As such it is unclear whether the type of change in economic dynamics that we have identified arises elsewhere. We leave addressing these limitations to future work.

References

Acemoglu, D., U. Akcigit, D. Hanley, and W. Kerr (2016). Transition to Clean Technology. *Journal of Political Economy 124*(1), 52–104.

Allcott, H. and D. Keniston (2018). Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America. *Review of Economic Studies 85*(2), 695–731.

Allen, R. C. (2000). Economic Structure and Agricultural Productivity in Europe, 1300–1800. *European Review of Economic History 4*(1), 1–25.

Allen, R. C. (2003). Was There A Timber Crisis in Early Modern Europe? In *Economia e energia secc. Xiii-xviii*, Number 34 in Serie 'II–Atti delle "Settimane di Studi" e altri Convegni, Prato, pp. 469–482. Instituto Internazionale di Storia Economica "F. Datini".

Allen, R. C. (2005). English and Welsh Agriculture, 1300-1850: Output, Inputs, and Income. Mimeo.

Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge, UK: Cambridge University Press.

Allen, R. C. (2023). The Transportation Revolution and the English Coal Industry, 1695–1842: A Geographical Approach. *Journal of Economic History 83*(4), 1175–1220.

Arkolakis, C. and C. Walsh (2023). Clean Growth. Technical report, National Bureau of Economic Research.

Ashraf, Q. and O. Galor (2011, August). Dynamics and Stagnation in the Malthusian Epoch. *American Economic Review 101*(5), 2003–41.

Ashraf, Q. H., O. Galor, and M. Klemp (2021). The Ancient Origins of the Wealth of Nations. In *The Handbook of Historical Economics*, pp. 675–717. Elsevier.

Beach, B. and W. W. Hanlon (2018). Coal Smoke and Mortality in an Early Industrial Economy. *Economic Journal 128*(615), 2652–2675.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Besley, T. and T. Persson (2023). The Political Economics of Green Transitions. *Quarterly Journal of Economics 138*(3), 1863–1906.

Blanc, G. and R. Wacziarg (2025). Malthusian Migrations. Technical report, National Bureau of Economic Research.

Bouscasse, P., E. Nakamura, and J. Steinsson (2025). When Did Growth Begin? New Estimates of Productivity Growth in England from 1250 to 1870. *Quarterly Journal of Economics 140*(2), 835–888.

Broadberry, S., B. M. S. Campbell, A. Klein, M. Overton, and B. van Leeuwen (2015). *British Economic Growth, 1270–1870*. Cambridge University Press.

Chen, S. and J. K.-s. Kung (2016). Of Maize and Men: The Effect of a New World Crop on Population and Economic Growth in China. *Journal of Economic Growth 21*(1), 71–99.

Cherniwchan, J. and J. Moreno-Cruz (2019). Maize and Precolonial Africa. *Journal of Development Economics 136*, 137–150.

Clark, G. (2002). The Agricultural Revolution and the Industrial Revolution: England, 1500-1912. Mimeo.

Clark, G. and G. Hamilton (2006). Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England. *Journal of Economic History 66*(3), 707–736.

Clark, G. and D. Jacks (2007). Coal and the Industrial Revolution, 1700–1869. *European Review of Economic History 11*(1), 39–72.

Clark, P. (1979). Migration in England During the Late Seventeenth and Early Eighteenth Centuries. *Past & Present 83*(1), 57–90.

Clay, K., J. Lewis, and E. Severnini (2024). Canary in a Coal Mine: Infant Mortality and Tradeoffs Associated with Mid-20th Century Air Pollution. *Review of Economics and Statistics 106*(3), 698–711.

Crafts, N. and T. C. Mills (2009). From Malthus to Solow: How Did the Malthusian Economy Really Evolve? *Journal of Macroeconomics 31*(1), 68–93.

Dinkelman, T. (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review 101*(7), 3078–3108.

Fernihough, A. and K. H. O'Rourke (2021). Coal and the European Industrial Revolution. *Economic Journal 131*(635), 1135–1149.

Feyrer, J., E. T. Mansur, and B. Sacerdote (2017). Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution. *American Economic Review 107*(4), 1313–1334.

Freese, B. (2003). *Coal: A Human History*. New York, USA: Penguin Books.

Galor, O. (2024). Unified Growth Theory: Roots of Growth and Inequality in the Wealth of Nations. *Annual Review of Economics 17*.

Galor, O. and Ö. Özak (2015). Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability. "Department of Economics Working Paper" 5, Brown University.

Galor, O. and Ö. Özak (2016). The Agricultural Origins of Time Preference. *American Economic Review 106*(10), 3064–3103.

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

Habakkuk, H. J. (1958). The Market for Monastic Property, 1539-1603. *Economic History Review 10*(3), 362–380.

Hanlon, W. W. (2020). Coal Smoke, City Growth, and the Costs of the Industrial Revolution. *Economic Journal 130*(626), 462–488.

Hansen, G. D. and E. C. Prescott (2002). Malthus to Solow. *American Economic Review 92*(4), 1205–1217.

Hanson, G. H. (2023). Local Labor Market Impacts of the Energy Transition: Prospects and Policies. Technical report, National Bureau of Economic Research.

Hatcher, J. (1993). *The History of the British Coal Industry. Volume 1, Before 1700: Towards the Age of Coal*. Oxford, UK: Clarendon Press.

Haywood, L., M. Janser, and N. Koch (2024). The Welfare Costs of Job Loss and Decarbonization: Evidence from Germany's Coal Phaseout. *Journal of the Association of Environmental and Resource Economists 11*(3), 577–611.

Heldring, L., J. A. Robinson, and S. Vollmer (2021). The Long-Run Impact of the Dissolution of the English Monasteries. *Quarterly Journal of Economics 136*(4), 2093–2145.

IIASA/FAO (2012). Global agro-ecological zones (gaez v3.0). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Jevons, W. S. (1865). *The Coal Question; An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of our Coal-Mines*. London and Cambridge, UK: MacMillan and Co.

Kelly, M., J. Mokyr, and C. Ó Gráda (2023). The Mechanics of the Industrial Revolution. *Journal of Political Economy 131*(1), 59–94.

Kelly, M. and C. Ó Gráda (2012). The Preventive Check in Medieval and Preindustrial England. *Journal of Economic History 72*(4), 1015–1035.

Lagerlöf, N.-P. (2019). Understanding Per-Capita Income Growth in Preindustrial Europe. *International Economic Review 60*(1), 219–240.

Lemoine, D. (2024). Innovation-led Transitions in Energy Supply. *American Economic Journal: Macroeconomics 16*(1), 29–65.

Lemoine, D. (2025). Climbing the Energy Ladder: How Energy Resources Hinder, Facilitate, and Fuel Economic Growth. Technical report, National Bureau of Economic Research.

Lewis, J. and E. Severnini (2020). Short-and Long-Run Impacts of Rural Electrification: Evidence from the Historical Rollout of the US Power Grid. *Journal of Development Economics 143*, 102412.

Lipscomb, M., A. M. Mobarak, and T. Barham (2013). Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil. *American Economic Journal: Applied Economics 5*(2), 200–231.

Matheis, M. (2016). Local Economic Impacts of Coal Mining in the United States 1870 to 1970. *Journal of Economic History 76*(4), 1152–1181.

Mokyr, J. (2009). *The Enlightened Economy: An Economic History of Britain, 1700-1850*. New Haven, CT: Yale University Press.

Moreno-Cruz, J. and M. S. Taylor (2020). Food, Fuel and the Domesday Economy. *European Economic Review 128*, 103501.

Muehlenbachs, L., E. Spiller, and C. Timmins (2015). The Housing Market Impacts of Shale Gas Development. *American Economic Review 105*(12), 3633–3659.

Nef, J. U. (1932a). *The Rise of the British Coal Industry*, Volume 1. London, UK: Routledge & Sons, Ltd.

Nef, J. U. (1932b). *The Rise of the British Coal Industry*, Volume 2. London, UK: Routledge & Sons, Ltd.

Nicolini, E. A. (2007). Was Malthus Right? A VAR Analysis of Economic and Demographic Interactions in Pre-Industrial England. *European Review of Economic History 11*(1), 99–121.

Nunn, N. and N. Qian (2011). The Potato's Contribution to Population and Urbanization: Evidence From a Historical Experiment. *Quarterly Journal of Economics 126*(2), 593–650.

Özak, Ö. (2010). The Voyage of Homo-Economicus: Some Economic Measures of Distance. Mimeo.

Özak, Ö. (2018). Distance to the Pre-Industrial Technological Frontier and Economic Development. *Journal of Economic Growth 23*(2), 175–221.

O'Rourke, K. H., A. S. Rahman, and A. M. Taylor (2013). Luddites, the Industrial Revolution, and the Demographic Transition. *Journal of Economic Growth 18*(4), 373–409.

Pomeranz, K. (2000). *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton, NJ: Princeton University Press.

Rackham, O. (1980). *Ancient Woodland: Its History, Vegetation and Uses in England*. Edward Arnold (Publishers), Ltd.

Rud, J.-P., M. Simmons, G. Toews, and F. Aragon (2024). Job Displacement Costs of Phasing Out Coal. *Journal of Public Economics 236*, 105167.

Satchell, M., L. M. W. S. Shaw-Taylor, E. A. Wrigley, P. M. Kitson, G. H. Newton, and G. Stanning (2023). 1831 England and Wales ancient counties. [Data Collection].

Severnini, E. (2023). The Power of Hydroelectric Dams: Historical Evidence from the United States over the Twentieth Century. *Economic Journal 133*(649), 420–459.

Stern, D. I., J. C. Pezzey, and Y. Lu (2021). Directed Technical Change and the British Industrial Revolution. *Journal of the Association of Environmental and Resource Economists 8*(6), 1079–1114.

Stokey, N. L. (2001). A Quantitative Model of the British Industrial Revolution, 1780–1850. *Carnegie-Rochester Conference Series On Public Policy 55*(1), 55–109.

Trew, A. (2014). Spatial Takeoff in the First Industrial Revolution. *Review of Economic Dynamics 17*(4), 707–725.

Voigtländer, N. and H.-J. Voth (2013a). How the West "Invented" Fertility Restriction. *American Economic Review 103*(6), 2227–2264.

Voigtländer, N. and H.-J. Voth (2013b). The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe. *Review of Economic Studies 80*(2), 774–811.

Wrigley, E. A. (1967). A Simple Model of London's Importance in Changing English Society and Economy 1650–1750. *Past & Present* (37), 44–70.

Wrigley, E. A. (2009). Rickman Revisited: The Population Growth Rates of English counties in the Early Modern Period. *Economic History Review 62*(3), 711–735.

Wrigley, E. A. (2010). *Energy and the English Industrial Revolution*. Cambridge, UK: Cambridge University Press.

A Appendix

A.1 Additional Empirical Results

In this section, we discuss the two robustness exercises referenced in Section 4.2 of the main text. First, we examine whether our main estimates are sensitive to our choice of agricultural productivity measure. Second, we examine the effects of forgoing our theory-tied empirical approach and instead estimating a standard difference-in-difference specification. As the results presented below show, our findings are robust along each of these dimensions.

A.1.1 Alternative Measures of Agricultural Productivity

For our first robustness exercise, we examine whether the estimates reported inTable 2 of the main text are sensitive to our choice of agricultural productivity measure. Recall from Section 3.4 of the main text that we measure agricultural productivity using the median value of average potential caloric yields for crops available prior to 1500 in each county, using data from Galor and Özak (2016). Galor and Özak also report mean values of average potential caloric yields, as well as the median and mean values of the *maximum* potential caloric yields for the same set of crops. Each of these alternative measures captures a slightly different aspect of agricultural potential. To examine the sensitivity of our results to our choice of measure, we re-estimate our preferred specification using each of these three alternative measures of agricultural productivity in turn.

The results from this exercise are presented in Table A1. For convenience, Column (1) reproduces the estimates from our preferred specification reported in Column (2) of Table 2. Column (2) reports estimates from a specification in which agricultural productivity is measured as the mean value of average potential caloric yields for crops available prior to 1500. Columns (3) and (4) report estimates from specifications in which agricultural productivity is measured as the median and mean values, respectively, of maximum potential caloric yields for the same set of crops.

As the estimates reported in Table A1 show, adopting these alternative measures has little effect on our results. The estimates of r and ϱ are essentially unchanged across columns. Similarly, the conclusions from the tests of $\varrho = 1$ and $\varrho = r$ remain unchanged. Adopting alternative measures of agricultural productivity does lead to modest increases in our estimate of the cumulative effect of the energy transition, suggesting if anything, our preferred specification understates the effects of the transition from wood to coal.

Table A1: Alternative Measures of Agricultural Productivity

	(1)	(2)	(3)	(4)
$\hat{r}$	0.9928	0.9926	0.9921	0.9921
	(0.0018)	(0.0029)	(0.0025)	(0.0034)
$\hat{\varrho}$	1.0002	0.9996	0.9995	0.9993
	(0.0017)	(0.0014)	(0.0016)	(0.0014)
R^2	0.998	0.998	0.998	0.998
Obs.	160	160	160	160
p-value: $\varrho = 1$	0.903	0.753	0.748	0.618
p-value: $\varrho = r$	0.000	0.028	0.006	0.048
$\hat{\kappa}_1 = \hat{r}^g$	0.805	0.799	0.789	0.788
	[0.045]	[0.070]	[0.059]	[0.082]
$\hat{\rho} = \hat{\varrho}^g$	1.006	0.987	0.985	0.980
	[0.052]	[0.041]	[0.046]	[0.040]
Cumulative effect	0.248	0.301	0.263	0.310
	[0.114]	[0.101]	[0.114]	[0.097]

Notes: Table reports estimates of the effects of coal on the population density of ancient counties in England. The dependent variable in all columns is the natural log of population density. All specifications include time fixed effects and allow for differential trends in caloric suitability. In column (1), caloric suitability is measured as the median value of average potential caloric yields for crops available prior to 1500 in each county. In column (2), caloric suitability is measured as the mean value of average potential caloric yields for crops available prior to 1500 in each county. In column (3), caloric suitability is measured as the median value of maximum potential caloric yields for crops available prior to 1500 in each county. In column (4), caloric suitability is measured as the mean value of maximum potential caloric yields for crops available prior to 1500 in each county. Standard errors clustered by county are shown in parentheses. $\hat{\kappa}_1$ and $\hat{\rho}$ calculated assuming a generation length $g = 30$. Standard errors for $\hat{\kappa}_1$, $\hat{\rho}$ and the cumulative effect of coal calculated via the Delta method are shown in square brackets. Coal importing counties are omitted from the sample.

A.1.2 Difference-in-Difference Estimates

For our second robustness exercise, we examine the effects of forgoing our theory-tied empirical approach and instead estimating a standard difference-in-difference specification. As we discussed in Section 3.3 of the main text, Equation (42) embeds a difference-in-difference research design, but its recursive structure allows us to separately identify the dynamic changes induced by the transition to coal. Setting this structure aside, the corresponding two-way fixed effects specification is given by:

$$\ln[N_{i,t}/\bar{L}_i] = \iota_0 + \iota_1[\text{Coal}_i \times \text{Post}_t] + \gamma X_i \eta_t + \mu_t + \delta_i + \varepsilon_{i,t}. \quad (45)$$

Equation (45) replaces the conditional independence assumption that supports our preferred specification with the parallel trends assumption typical of a difference-in-difference research design. It also no longer allows us to separately identify the dynamic growth effects of the transition to coal. Instead, ι_1 in Equation (45) captures the *average* effect of the transition to coal on population density for affected counties across all post-Dissolution

Table A2: Difference in Difference Estimates

	(1)	(2)	(3)	(4)	(5)
$Coal_i \times Post_t$	0.212	-0.008	0.595	0.384	0.250
	(0.086)	(0.091)	(0.108)	(0.142)	(0.121)
County FE		X		X	X
Time FE			X	X	X
CSI× Time FE					X
R^2	0.04	0.40	0.53	0.79	0.81
Obs.	160	160	160	160	160

Notes: Table reports estimates of the effects of coal on the population density of ancient counties in England. Standard errors clustered by county are shown in parentheses. Coal importing counties are omitted from the sample.

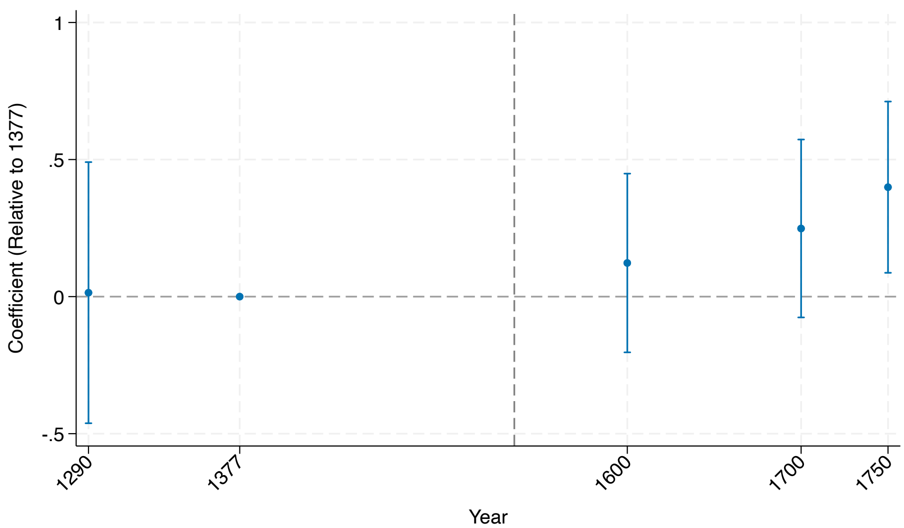
periods in our sample.

The results from estimating Equation (45) are reported in Table A2. Column (1) reports estimates from the simplest possible specification, which omits county and time fixed effects. Column (2) adds county fixed effects, while Column (3) instead adds time fixed effects. Column (4) includes both county and time fixed effects, and Column (5) corresponds most closely to our preferred specification in the main text by additionally allowing for differential trends in caloric suitability. The estimated coefficient on $Coal_i \times Post_t$ is positive and statistically significant across most specifications. In Column (5), which is closest in spirit to our preferred specification in the main text, the estimate implies access to coal after the Dissolution of the Monasteries increased the population density of affected counties by close to 28%.

Forgoing our theory based empirical approach in favor of a difference-in-difference design also allows us to estimate an event study version of Equation (45) to examine the possibility of pre-existing differences in trends across coal and non-coal counties. The results of this exercise are presented in Figure A1. As the figure shows, the coefficients on the interaction for 1290 is statistically indistinguishable from zero, suggesting there are no differential trends prior to the Dissolution. Although imprecise, the coefficients then climb monotonically after 1600, from roughly 0.25 in 1700 to 0.40 in 1750, consistent with the cumulative effect of the change in population dynamics predicted by our theory.

Taken together, the estimates reported in Table A2 and Figure A1 suggest that even when we abandon our theory-tied empirical approach in favor of a standard difference-in-difference research design, we recover an estimate of the effect of the transition to coal that is consistent with the findings from our main empirical approach.

Figure A1: Event Study: Log Population Density and Coal Access



Notes: Figure reports the estimated event-study coefficients from the event-study version of Equation (45). The specification includes county fixed effects, time fixed effects, and caloric suitability interacted with time fixed effects. The vertical dashed line marks 1535, the start of the Dissolution of the Monasteries; 1377 is the reference period. Whiskers report 95% confidence intervals constructed from standard errors clustered by county. Importing counties are omitted from the sample.