Directed technical change as a response to natural-resource scarcity

John Hassler[*], Per Krusell[†], and Conny Olovsson[‡]

June 26, 2019

Abstract

How do markets economize on scarce natural resources? With an application to fossil energy, we emphasize technological change aimed at saving on the scarce resource. We develop quantitative macroeconomic theory as a tool for interpreting the past and thinking about the future. We argue, first, that aggregate U.S. data calls for a short-run substitution elasticity between energy and the capital/labor inputs that is near Leontief. Given this fact and an aggregate CES function, we note that energy-saving technical change took off right as the oil shocks hit in the 1970s. We rationalize this observation using a theory that views technical change as directed: it can be used to save on different inputs and, hence, the long-run substitutability between inputs becomes higher than Leontief. For our application, we estimate long-run dependence on fossil energy—measured by its factor share—to climb to a little below 10%; absent endogenous technical change directed toward energy-saving, it would go to 100%.

1 Introduction

What is the future of our dependence on natural resources in finite supply? How will consumption growth be affected by scarcity? We develop quantitative theory to answer these questions and apply it to the case of fossil fuel-based energy as an input into production. The

[*]Hassler: Institute for International Economic Studies (IIES), University of Gothenburg (UG), and CEPR.
Support from Mistra-Swecia and the Swedish Energy Agency is gratefully acknowledged.

[†]Krusell: IIES, CEPR, and NBER. Support from the European Research Council, Mistra-Swecia, and the Swedish Energy Agency is gratefully acknowledged.

[‡]Olovsson: Sveriges Riksbank. The present paper circulated in an earlier version with a different name ("Energy-Saving Technical Change"). The opinions expressed in this article are the sole responsibility of the authors and should not be interpreted as reflecting the views of Sveriges Riksbank.

market's first response to scarcity is a rise in the price of the scarce resource, with curbed use as a result. In this paper we focus on an implication of a higher price: endogenous resource-saving technical change, in the form of new techniques and products allowing us to save on inputs. We use the theory to understand the postwar U.S. data on fossil energy dependence but we also make projections into the future.

How potent is technology in resource-saving, however? Viewed through the lens of our structural macroeconomic model, historical data allows us to assess how effectively energy saving has responded to price movements in the past. We use this information to parameterize the structural model so that it can be used to predict how potent energy saving is likely to be in the future. Thus, it allows us to directly address the sustainability issue: what are the effects of the resource scarcity on economic growth and welfare, and how high will the payments to fossil energy be as a fraction of GDP?

Energy saving can be accomplished both by reducing energy waste and by shifting toward less fossil energy-intensive products. The structural model used in this paper is aggregate in nature and thus melds these two together by a general focus on the demand side. In particular, we formulate an aggregate production function, aimed at describing the U.S. economy, with capital, labor, and fossil energy as inputs. We focus on a function that is not necessarily Cobb-Douglas and that therefore allows us to identify separate input-augmenting technology series. We thus look at two input aggregates: a capital-labor composite and energy. The model then contains another layer where the two corresponding (input-augmenting) technology series are subject to choice, along the lines of Acemoglu (2002). This mechanism allows us to capture the natural notion that there is very low short-run substitutability between energy and other inputs, once the technology factors at a point in time have been chosen, but significantly higher substitutability over longer periods when these factors are endogenous.[1] We then show how the long-run energy share, along with consumption growth, will be determined in the model.

With parameter values in hand and an estimate of initial stocks, one can use the model to compute the future paths for technologies, output, and welfare. In the selection of parameter values, there are at least two important challenges, however. One is to determine the shape of the "ex-post" aggregate production function, i.e., the input elasticities conditional on given values of the factor-augmenting technology levels. This shape is important per se but

[1]Our approach for modeling ex-ante/ex-post distinctions between input elasticities is, we believe, novel relative to the earlier literature on the topic, which has tended to look at vintage structures. For the latter, see Atkeson and Kehoe (1999) and the interesting recent study in Abrell, Rausch, and Schwerin (2016); for an application to capital vs. labor, see Léon-Ledesma and Satchi (2019).

in our context it plays a special role. Given prices and quantities of the different inputs and of output, the shape of the ex-post production function affects our measures of how the key factor-augmenting technology levels move over time. The other challenge is to characterize the "technology technology": the production possibility frontier for the factor-augmenting technology levels or, more precisely, what future factor-augmenting technology levels are attainable given their current values. Our approach is to estimate our structural model so as to address both these questions.

We find, first, that an ex-post aggregate production function with a unitary elasticity between capital and labor and a near-zero elasticity between the capital-labor composite and fossil energy fits the data quite well. This finding is highly robust to the econometric technique we use; in particular, functions where the latter elasticity is not close to zero are extremely difficult to reconcile with the data, chiefly because the price of fossil fuel tracks the share very closely.[2]

Next, given this overall ex-post shape of the production function, we make several observations. First, the energy-saving technology trend took off very sharply after the oil-price shocks in the 1970s, after having been dormant for decades. Second, the capital/labor-saving technology series looks very much like the standard aggregate TFP series, thus mimicking the well-known productivity slowdown episode but otherwise featuring steady growth. Third, the implied energy-saving technology trend comoves negatively with the capital/labor-saving technology trend, whether we use annual observations or look across subperiods. These findings suggest, precisely, that technical change—in the form of saving on different inputs—is endogenous and, as labeled in the literature (see Hicks (1932), Kennedy (1964), Dandrakis and Phelps (1966), Acemoglu (2002), and others), "directed".

Turning to the calibration of the parameters governing the direction of technological improvements, it turns out that the theory has a rather direct link between these parameters and observables. We show, in particular, for a general constant-returns production function, that the key transformation elasticity between the growth rates of the two kinds of factor-augmenting technologies in the long run of the model must equal the relative cost shares of

[2] A substitution elasticity between energy and capital/labor much above zero would require very large short-run changes in the production technology parameters, which is challenging to rationalize and indeed is ruled out in our econometric estimates. Relatedly, electricity demand on short horizons are even viewed to be good instruments for aggregate economic activity; see Jorgenson and Griliches (1967) and the many studies following it. Applications in the literature of the type of production function we employ also use estimates consistent with what we find here; e.g., Manne et al. (1995) use an elasticity of 0.4 for a model with a ten-year time period.

the two factors.[3] I.e., a one-percent decrease in the growth rate of capital/labor-augmenting technical change allows an increase in the growth rate of energy-augmenting technical change by an amount that equals the ratio of the cost share of capital and labor to that of energy in production. Our finding that the growth rates in the two technology trends have varied historically thus allows us to identify what this transformation elasticity has been given our range of past data. Here the quantitative estimates are somewhat more sensitive to data selection and econometric technique.

Our estimated parameters imply a point estimate for the long-run energy share that is around seven percent. Thus, absent innovation into new sources of energy, the future appears to bring a significantly, but not radically, higher energy dependence than we have today. There is no collapse of economic growth, however: the implied long-run growth rate of consumption is only somewhat lower than in the past—about 1.9% per year. Thus, although the economy will show high dependence on fossil energy it will still generate high consumer welfare.

We begin our analysis with our main application: after a brief literature review in Section 2, in Section 3 we posit a nested CES function of our aggregate inputs capital/labor and an energy composite and confront it with U.S. data. This exercise uses a low value of the elasticity between inputs, given the high correlation between the energy price and the energy share; the elasticity chosen here is close to that formally estimated later in the paper. The key finding is that energy-saving technical change took off in response to the oil-price hikes in the 1970s, thus motivating our focus on endogenous directed technical change. In Section 4 we then introduce our full model of energy-saving technical change. Step one of this analysis is to extend the seminal analysis in Dasgupta and Heal (1974) in two directions. First, we look at a capital-labor composite and energy (rather than capital and energy) but second, and most importantly, we allow for exogenous input-saving technical change, all under the condition that the input substitution elasticity is less than unity. Here, unless a very specific, knife-edge combination of growth rates of the two kinds of input-saving technologies is assumed, the economy's asymptotic behavior is associated with one of the inputs losing importance. In particular, if the growth rate of natural resource-saving technical change is low enough, its cost share goes to 100% in the limit. We then introduce endogenous technology and show that these results, like in Acemoglu's (2003) work on directed capital- and labor-augmenting technology, change entirely: the economy's asymptotic behavior will now always be balanced. In particular, what was a knife-edge case when technology was

[3] We make the simplifying assumption that the total amount of R&D, appropriately defined, is exogenous.

exogenous is now selected endogenously as a result of research efforts. Moreover, we show analytically that the long-run natural resource share of income exclusively depends on how costly it is to enhance its efficiency in terms of lost capital/labor efficiency. In Section 5, we then estimate the structural parameters of our model of directed technical change on the same data as we briefly looked at in the earlier empirical section. This section also computes the long-run energy share and consumption growth rate implied by our estimates and computes a transition path into the future where fossil-fuel use peaks two decades from now. Section 6, finally, concludes.

2 Connections to the literature

Our finding that there is active, directed technical change is perhaps not surprising given a variety of studies using disaggregated analysis; see, e.g., Popp (2002) for energy-saving and Aghion et al. (2016) for the application to "clean" and "dirty" technologies in the case of autos. In terms of magnitudes, the microeconomic estimates in these specific studies suggest a somewhat lower ability to substitute across technologies than what we find using aggregate data.

Our aggregate focus resembles that in the literature on directed technical change toward high- vs. low-skilled labor (or products intensive in these respective inputs). Beginning with Katz and Murphy's (1992) paper, an argument was put forth—also using an aggregate CES technology—that there has been skill-biased technical change since the late 1970s. Acemoglu (1998) then looked at how changes in the sizes of college-graduating cohorts could have explained this fact. We conduct these exercises jointly, including an estimation of both the CES elasticity and the technology available for choosing factor-augmenting technologies.

The large literature following Acemoglu (1998), which in turn builds conceptually on Hicks (1932), Kennedy (1964), and Dandrakis and Phelps (1966), exploits modern endogenous-growth techniques to formulate models of endogenous directed technical change. We stand on the shoulders of all this work and our purpose is very applied: we look at the natural resource case, with a focus on fossil energy, where a stock is depleted over time. We also restrict attention to cases where the natural resource is "hard" to replace in the short run: we assume an input substitution elasticity less than unity. We also focus mostly on a planning problem, and where we consider market allocations, we look at pure technology spillovers rather than R&D, mostly for simplicity. We argue, again for our application, that the higher positive spillovers generated by focusing more on one kind of input-saving are always counteracted by the lower spillovers associated with focusing less on the other input. In fact, in a simple version the model, we can show that these factors cancel exactly and that the direction of technical change in the market equilibrium is efficient. More general technologies and market settings are studied, e.g., in Acemoglu (2002, 2003, and 2007).

We estimate the elasticity between energy and the capital/labor composite to be close to zero based on annual aggregate U.S. data. This estimate is broadly consistent with other parameterizations in the applied literature, e.g., in the integrated assessment literature where the same functional form is used and five-year elasticities are of the order of 0.5; see Manne, Mendelsohn, and Richels (1995) for the MERGE model, Bosetti et al. (2006) for the WITCH model, and Werf (2008), as well as with a broad range of econometric estimates (see, e.g., Dahl and Sterner, 1991).[4]

Our model can produce a time path for fossil use that is initially rising for a rather long time, only to eventually fall. This pattern is clear in the data but is difficult to produce in more standard models of finite natural resources; see, for example, the literature following the oil-price shocks of the 1970s, e.g., Dasgupta and Heal (1974), Solow (1974), and Stiglitz (1974, 1979).[5,6] Relatedly, Jones's (2002) textbook on economic growth has a chapter on non-renewable resources with quantitative observations related to those we make here. Recently, a growing concern for the climate consequences of the emission of fossil CO_2 into the atmosphere has stimulated research into the supply of, but also demand for, fossil fuels as well as alternatives; see, e.g., Acemoglu et al. (2012) or Hémous (2016). The recent literature, including the present paper, differs from the earlier contributions to a large extent because of the focus on endogenous technical change, making use of the theoretical advances from the endogenous-growth literature.[7] Our paper is focused on the role of technology for resource saving, but the development of alternative sources of energy is also possible through technical change. This appears important and valuable to study but is not our main focus here.[8]

[4]A related elasticity is the percentage response of fossil energy use to a one-percent increase in its price, which some studies argue is large; see, e.g., Kilian (2008). Such an estimate can still be consistent with our production function parametrization if all inputs can be varied in the short run due to movements in capacity utilization in response to fossil-energy prices.

[5]New discoveries and technological improvements on the supply side can also be invoked.

[6]Given the consistently upward-sloping path for fossil-fuel use and the large deposits of coal and non-conventional oil and gas, the finiteness of fossil fuel could even be be questioned from a quantitative perspective. However, because of climate concerns, and because these alternative fossil resources all have their challenges—in being transported or feasibly extracted—we find the supposition of a limited stock of fossil fuel a realistic one.

[7]See, e.g., Aghion and Howitt (1992).

[8]In follow-up work (Hassler, Krusell, Olovsson, and Reiter, 2019), we apply the insights in the present

3 Empirical motivation

Our focus is the U.S. economy. Figure 1 below shows the evolution of the fossil energy's share of output, the fossil energy price, as well as its use. The notion we use for "energy" is an index of the three main fossil fuels: oil, coal, and natural gas, and its price is an index defined accordingly (see the Appendix for details). We thus look at a broad measure of fossil-fuel energy, even though the sub-components are somewhat heterogeneous; natural gas and oil are rather similar in terms of their production technologies but coal is different, with a higher marginal cost as a fraction of the market price and with higher estimated proved reserves. We also abstract from non-fossil sources of energy (such as hydro and nuclear power).[9] Thus, we take energy's share of output to be ep/q in Figure 1, where e is fossil-energy use, p is the fossil-fuel price in chained (2005) dollars, and q is our measure of output, to be discussed below: GDP minus net export of fossil fuel in chained (2005) dollars. The specific data on e and p are both taken from the U.S. Energy Information Administration.

As can be seen from top part of the figure, (i) energy's share is highly correlated with its price and (ii) it is quite volatile. Specifically, the share starts out around three percent in 1949 and then decreases somewhat up to the first oil price shock when it increases dramatically. The share then falls drastically between 1981 and the second half of the nineties and then finally increases again. The share does not seem to have an obvious long-run trend. The bottom graph shows that fossil energy use has been increasing throughout the post-war period in the United States, with a slowdown beginning in the early 1970s.

Our approach is to interpret these data from the perspective of an aggregate production function delivering "output" using three inputs: aggregate capital, aggregate labor, and aggregate fossil energy. We use quotation marks here because—abstracting from the use of energy outside productive domestic activities—we take the aggregate production function to be equal to GDP, even though energy has the appearance of an intermediate output, and thus our function is a gross production function in this sense. In the data, of course, GDP is total value added produced from non-energy-producing sectors (using energy) plus that from the energy-producing sector itself. We could thus define two production functions, one for each of these sectors, with arguments being the inputs used in the respective sectors. To simplify, we instead assume that the energy-producing sector delivers a pure rent, i.e., that energy is

paper to this case in the context of a climate-economy model.

[9] One can straightforwardly consider broader aggregates, but the fossil share is and has been very high. As for our fossil-fuel aggregate, we adopted a standard formulation here but also considered an oil-only case, with very similar results to those reported in the text.

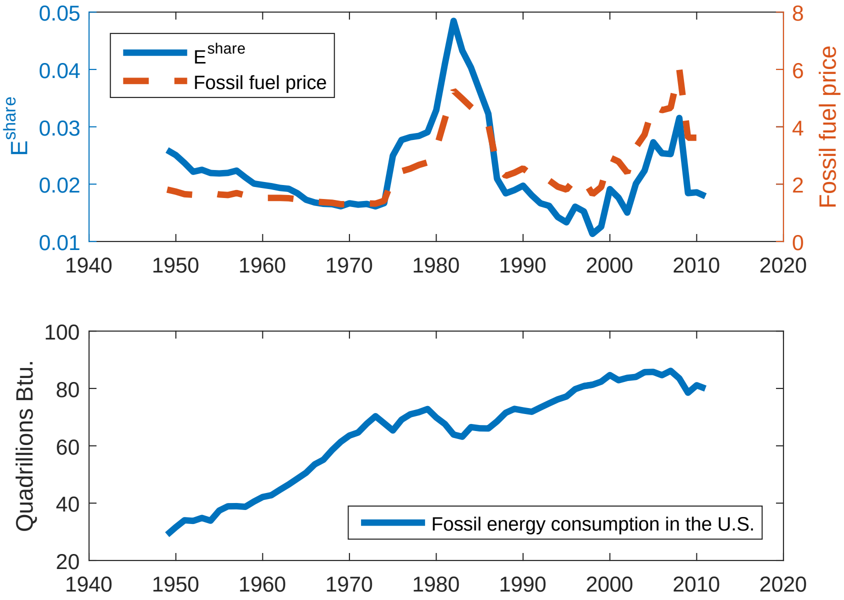

Figure 1: Top graph: fossil energy share (scale to the left) and its price (scale to the right). Bottom graph: fossil energy consumption in the United States.

produced at zero cost. Hence our aggregate production function is a function of aggregate capital and aggregate labor, as well as of aggregate energy used in domestic production. Our abstraction from costs in the energy-producing sector sharpens the ensuing theoretical analysis and also avoids the empirical challenge of separately tracking input use (capital, labor, and energy) in the two sectors over time. For robustness, we have also elaborated with alternatives and found that, because the energy sector is small relative to the total, they deliver only very marginal changes to our quantitative results (both in terms of the basic plots, such as Figure 1, and the estimations).[10] In sum, we assume that our aggregate production delivers output, q, defined as GDP minus the value of energy use outside of domestic production; this outside energy use equals net export of (fossil) energy plus the household use of fossil energy as a final good (which largely consists of auto fuel). We set the latter, which is a very small amount compared to the total, to zero for lack of a consistent time series on it.[11] Our appendix briefly describes the data sources and construction.

[10]Our robustness checks included assuming (i) that the energy-producing sector has the same isoquant shapes as in the non-energy-producing sector and (ii) that the energy-sector production function is Cobb-Douglas.

[11]Energy used for heating homes is different: it is considered an intermediary good in the production of

In most quantitative-theory applications in macroeconomics, a Cobb-Douglas function is used when the inputs are capital and labor, though recently substitution elasticities slightly different than one have been considered.[12] Moreover, our aim is to use a production function that allows us, conditional on having data on prices and quantities, to back out a measure of energy-saving (technological change). Furthermore, since the energy share tracks the energy price so closely in the short run, having a Leontief function as a special case appears desirable. Several structures are consistent with these features but the most parsimonious case is a nested CES: it allows for two factor-specific technology trends, in addition to a Hicks-neutral factor. Moreover, since the capital and labor shares are remarkably stable over the period we consider, we opt for an even more parsimonious structure by adopting a nested CES where one of the nestings has a Cobb-Douglas structure, thus also eliminating one of the factor-specific trends.[13] Three possible nestings remain and we opted for one with a Cobb-Douglas composite of capital and labor, in turn forming a CES with energy. It turns out that the key features of our technology trends—to be displayed below—do not appreciably depend on this choice. A structure where either capital or labor forms a composite with energy would imply very sharp changes in the capital or labor shares in response to the oil shocks in the 1970s; we do not seem to observe this and hence favored a structure where capital and labor form a Cobb-Douglas composite (and, hence, their respective incomes will decline evenly in response to a short-run price hike for fossil energy). Thus, we consider the following production technology:

$$q_t \equiv F\left(A_t k_t^{\alpha} l_t^{1-\alpha}, A_{et} e_t\right) = \left[(1-\gamma)\left(A_t k_t^{\alpha} l_t^{1-\alpha}\right)^{\frac{\varepsilon-1}{\varepsilon}} + \gamma\left(A_{et} e_t\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \qquad (1)$$

where $\alpha \in (0,1)$ and where ε is the elasticity of substitution between capital/labor and fossil energy. γ is a share parameter.[14] Note that when $\varepsilon = \infty$, the Cobb-Douglas composite and fossil energy are perfect substitutes, when $\varepsilon = 1$, the production function collapses to being Cobb-Douglas in all input arguments; and when $\varepsilon = 0$ the Cobb-Douglas composite and energy are perfect complements, implying a Leontief function in the capital-labor composite and energy. The two variables A_t and A_{et} are the *input-saving* technology levels for capital/labor and energy, respectively; these are well-defined so long as ε is not equal to 1.

housing services, which is a(n imputed) part of GDP.

[12]See, e.g., Karabarbounis and Neiman (2013), who use an elasticity a little above one along with a declining relative price of capital.

[13]Throughout, we define shares in terms of our notion of aggregate output, which differs slightly from GDP, as just discussed above.

[14]A similar production function is considered by Stern and Kander (2012).

A key observation now is that it is possible, conditional on a value for the substitution elasticity, to use this production function, along with data on inputs and outputs, to back out the two energy-saving technology series. This is interesting for two reasons. One is that one expects these series to behave "like technology", i.e., be rather smooth and increasing. Thus, a preliminary test of the production function-based theory is possible. The second reason is also the main purpose of this paper: we would expect input saving to respond to incentives and this idea can also be assessed, again in a preliminary manner, by inspection of the series and how they change as the price changes. Thus, under the assumption of perfect competition in input markets, marginal products equal factor prices, so that labor's and energy's shares of income are given by

$$l_t^{share} = (1-\alpha)(1-\gamma)\left[\frac{A_t k_t^{\alpha} l_t^{1-\alpha}}{q_t}\right]^{\frac{\varepsilon-1}{\varepsilon}} \tag{2}$$

and

$$e_t^{share} = \gamma\left[\frac{A_{et} e_t}{q_t}\right]^{\frac{\varepsilon-1}{\varepsilon}}, \tag{3}$$

respectively. Equations (2) and (3) can be rearranged and solved directly for the two technology trends A_t and A_{et}. This delivers

$$A_t = \frac{q_t}{k_t^{\alpha} l_t^{1-\alpha}}\left[\frac{l_t^{share}}{(1-\alpha)(1-\gamma)}\right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{4}$$

and

$$A_{et} = \frac{q_t}{e_t}\left[\frac{e_t^{share}}{\gamma}\right]^{\frac{\varepsilon}{\varepsilon-1}}. \tag{5}$$

It is clear here that with ε and γ given, and with data on q_t, k_t, l_t, e_t, l_t^{share}, and e_t^{share}, equations (2) and (3) give explicit expressions for the evolution of the two technologies. Clearly, however, the parameter γ is a mere shifter of these time series and will not play a role in the subsequent analysis. The key parameter, of course, is the elasticity ε.

Now let us use these expressions for a preliminary evaluation: let us use a low value for ε (in order to obtain the strong covariation between the energy share and the price without severely fluctuating technology series) and inspect the results. In Figure 2 below we thus assume a very low elasticity: $\varepsilon = 0.02$.

The figure shows the path for fossil energy-saving technology A_e. Two points are noteworthy in this graph. First, we observe a weakly increasing, and overall reasonable-looking

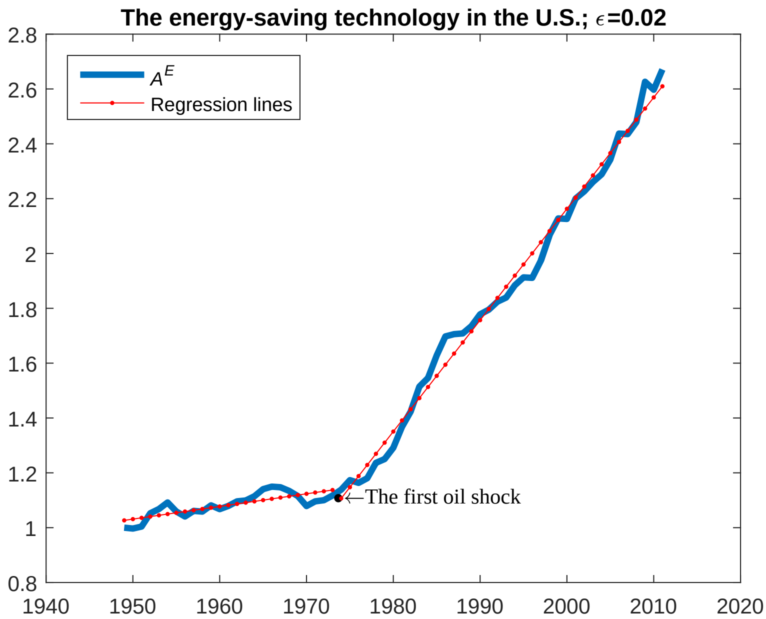

Figure 2: Energy-saving technology with an elasticity of 0.02.

graph for fossil energy-specific technology. The mean growth rate is 1.47 percent and the standard deviation is 2.25 percent. Second, there is a very striking kink in the series in the beginning of the 1970s: after a virtually flat technology path up to this moment, the series takes off at a significant rate. The figure thus also shows separate trends lines before and after the first oil-price shocks: 1949–1973 and 1973–2009. Clearly, the technology series appears to have a kink around the time of the first oil price shock; the growth rate is 0.1 percent per year up to 1973 and 2.54 percent per year after 1973.

What does a low substitution elasticity imply for the evolution of the capital/labor-augmenting technology? The series for A is plotted as the solid line in Figure 3, alongside the A_e series. A, like A_e, is rather smooth and increasing and very much looks like the conventional total-factor productivity (TFP) series. The mean growth rate in A is 1.28 percent and the standard deviation is 1.63 percent.

Thus, in summary, we note that an aggregate production function of the sort we have used produces paths for the input-saving technology levels that look rather reasonable. In terms of interpretation, let us look at how the two series correlate. When the energy price rises, energy saving responds, at the same time as the input saving in capital/labor slows down. Thus, from our preliminary assessment here, it does look like technical change directs itself toward the input on which it is profitable to save. This observation is the basis on

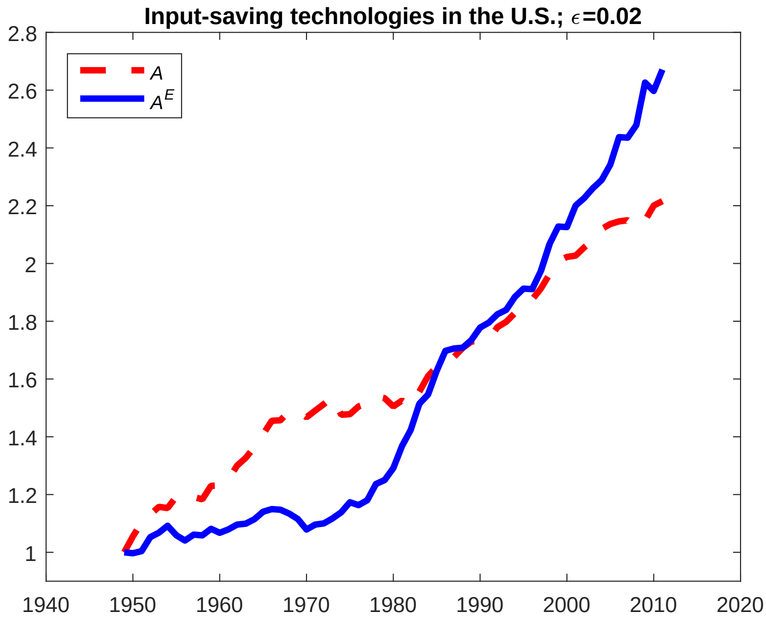

Figure 3: Energy- and capital/labor-saving technologies compared.

which we will build a model where input saving does respond to incentives.

Before presenting the theory, let us note that a possible alternative explanation for the kink around 1973 could be structural transformation, e.g., the expansion of the service sector relative to the manufacturing sector. Specifically, if the production of services requires relatively less energy, then such a process could be mistaken for energy-saving technical change. At the same time, however, if services in fact do require relatively less energy than manufactured goods then this type of transformation could also be an endogenous response to the oil price shocks. Regardless of the direction of causation, Figure 6 in Appendix A.2 shows the energy-saving technology in the manufacturing sector, and it is qualitatively very similar to that in the aggregate; the kink is somewhat less pronounced, but the interpretation is that there appears to have been a drastic increase in growth rate of energy efficiency in both the manufacturing and the service sectors.

Also, note that although the backed-out series for input-saving technologies depend on the observed prices, the fundamental reasons behind the price fluctuations do not play in directly in our measurements. This is important, as there are competing hypotheses as to what drives fossil-fuel prices. Similarly, whether the supply side is, or has been, characterized by monopoly or oligopoly, e.g., through the Texas Railroad Commission or OPEC, is not of immediate consequence, as long as the demand side is competitive, because we use prices only to back out features of demand.[15]

The conclusions above are robust to values for the elasticity ε in a range from 0 to about 0.05. For larger values, higher volatility in the technology series are obtained, especially near 1 where the technology fluctuations up and down are very dramatic.[16] However, all of these observations only represent simple correlations. Especially given the idea that technical change responds to incentives, it would be important to factor this endogeneity into an estimation of ε. We perform such a structural estimation later in the paper, precisely using a model with directed technical change. To move toward this estimation, we present our theory in the next section.

4 The model

In this section, we formulate a model with the aim of allowing technical change to respond endogenously to changes in the economic environment, allowing us in particular to evaluate the predictions for future energy dependence. The model will have the features suggested by the above data: a very low short-run elasticity between energy and other inputs but a significantly higher one over longer horizons, engineered by directed technical change that saves on expensive inputs. We parameterize the structure so as to allow a quantification of the future energy dependence and its implications for the overall growth rate of the economy. Our analysis of directed technical change allows us to address the future path for energy use. Clearly, fossil-fuel use has increased over time and since this resource is in finite supply it must reach a maximum at some point and then fall. A challenge here is that standard models do not predict this pattern: they predict *falling* resource use from the beginning of time. By standard models, here, we refer to settings relying on the classic work on nonrenewable resources in Dasgupta and Heal (1974) and with high substitutability between fossil fuel and capital/labor. As will become clear, a lower elasticity of substitution between fossil fuel and capital/labor is a potentially important factor behind this phenomenon; in particular it is important to include physical capital in the analysis. We will look at these issues more

[15] For these issues, in the case of oil, see Barsky and Kilian (2002), Kilian (2009), and Bornstein, Krusell, and Rebelo (2017).

[16] One can operationalize the hypothesis that the A and A^e series are technology variables by selecting a value for the elasticity that minimizes a metric like the sum of squares of their growth rates and this procedure produces a value close to that on which the above graphs are based.

carefully after laying out the basic model.[17]

The model will be constructed in steps. First, in Section 4.1, we specify a standard neoclassical dynamic macroeconomic model with an energy input, as in Dasgupta and Heal (1974). Relative to Dasgupta and Heal's work, we generalize to the kind of nested CES production function used in the previous section, a formulation on which we will also base the main analysis below. In addition, we consider exogenous input-saving technical change in this section. The key result in the present section is a characterization of the asymptotic balanced growth paths, i.e., paths where output grows at a constant rate but where one of the factors of production may have lost its importance in the sense of commanding a zero cost share. In particular, we show that balanced growth with balanced cost shares is only possible as a knife-edge case; this case is also consistent with an exact balanced growth path. If, in particular, the elasticity between the inputs is low(er than unity) and the energy-saving technology does not grow fast enough, the energy share will converge to 100%.

We then endogenize the input-saving technology. In Section 4.2.1, first, we look at the simplest possible framework for directed technology choice: a static model that allows us to derive an endogenous energy share under technology choice. As it turns out, the key properties of the static model will be inherited by the steady state of the main dynamic model to be estimated. The dynamic model of endogenous technical change is studied, in Section 4.2.2, first in a general formulation and then for special cases relevant to the estimation and further analysis. Here we find that, if the inputs are sufficiently complementary in the short run, the long-run cost shares will robustly be balanced, i.e., when technology choice is directed, the long-run share of energy will neither go to zero nor to 100% but to a nontrivially determined intermediate value. This model thus features different input substitution elasticities at different time horizons.

4.1 Exogenous technical change

We begin by studying a standard neoclassical macroeconomic model with an energy input and with exogenous paths for input-saving technology levels. We admit any preferences consistent with exact balanced growth (the period utility flow is a power function of consumption). The production function is of the CES variety considered in Section 3 above. We use the simplest possible formulation for the natural resource: we assume costless extraction. This is not an

[17] A number of explanations for increasing resource use, such as new discoveries, imperfect foresight, and falling marginal costs of production, have been proposed. The mechanism we use here should of course be viewed as complementary with these.

essential feature of the framework but simplifies the exposition.[18]

Let us now describe the formal setting. There is one consumption good each period and we consider an infinitely lived household with the following utility function.[19]

$$\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma} - 1}{1-\sigma}. \tag{6}$$

Output, y, is produced with capital, k, labor, l, and a natural-resource input, e, as inputs, exactly as in the specification in the empirical section above: y_t is given by $F(A_t k_t^{\alpha}, A_t^e e_t)$, where F is defined in equation (1) above; the one difference is that we keep labor input equal to 1 here. Again, A and A_e are measures of input-saving: the level of the capital/labor-augmenting technology and the natural resource-, or as in our main application, energy-augmenting technology, respectively. In this section, we assume constant growth rates for both of the input-saving technologies:

$$A_t = g_A^t$$

and

$$A_{et} = g_{A_e}^t.$$

We assume that our (gross) rates g_A and g_{A_e} are strictly greater than 0 and finite.

The period resource constraint, as in standard one-sector growth models, is given by

$$c_t + k_{t+1} = y_t + (1-\delta)k_t, \tag{7}$$

with the depreciation rate $\delta \in (0, 1)$, as in standard one-good models.

The size of the natural resource stock at time zero is R_0 and the following constraint

[18] Oil is available from different sources, each associated with a specific (non-zero) unit cost of extraction. Although the marginal cost of most oil in Saudi Arabia is close enough to zero, it is not close to zero in the North Sea. Moreover, the extraction costs can also be affected by R&D and may be stochastic (shale gas and tar sand are examples of recent innovations of this nature). A full quantitative treatment of oil hence needs a richer structure. It also needs the inclusion of other energy sources, fossil as well as non-fossil. Our simplifying assumptions allow us to uncover mechanisms rather clearly and to characterize long-run outcomes without resorting to numerical analysis. However, for all the reasons above, it would clearly be worthwhile to go beyond the simple assumptions we use here and look at their robustness, a task that we judged to be outside the scope of the present analysis.

[19] We assume that β times the growth factor for consumption raised to $1-\sigma$ is less than one; consumption's growth rate is a nontrivial—but easy-to-determine—factor of the underlying parameters, as we will see below.

must be satisfied:

$$\sum_{t=0}^{\infty} e_t \leq R_0, \tag{8}$$

where R_t is the remaining stock of resource in ground in the beginning of time t; we could equivalently write $R_{t+1} = R_t - e_t$ for all t. We impose that e and R be non-negative at all times; these constraints will not bind for any cases of interest so they are omitted for brevity.

As is clear from the above equations, we restrict labor supply so that $l_t = 1$ for all t. This can be thought of as "full employment", an assumption that we think is an appropriate approximation for the long run in the absence of changes in the tax system.[20]

In the analysis of endogenous technical change below, there are potential differences between the optimal allocation and the equilibrium allocation we consider. However, in the present section, where technology paths are taken as given, the straightforward competitive equilibrium decentralization produces an optimal allocation. Therefore, we only study the planning problem here.

We define an *Asymptotic Constant Growth Path* (ACGP) as a limit solution to the planner's optimality conditions—including transversality conditions—where all variables grow at constant, though possibly different, rates. An *Exact Balanced Growth Path* (EBGP) is a solution to the same conditions that, by appropriate choice of initial conditions, features exact balanced growth at all points in time with (i) identical growth rates for capital, consumption, and output and (ii) constant positive cost shares for the two inputs, i.e., the energy share $s_{et} \equiv A_{et}e_t F_2(A_t k_t^\alpha, A_{et}e_t)/F(A_t k_t^\alpha, A_{et}e_t)$ is strictly between zero and one. Hence, an EBGP is a special case of an ACGP, and an ACGP does not necessarily deliver balanced cost shares nor can it necessarily feature exact constant growth. For either case, here and in the analysis below on endogenous technical change, we do not consider the possibility of limit behavior that is not asymptotically of the constant-growth variety (such as cycles, chaos, or exploding paths); it is possible to show convergence to the solutions we focus on for special cases, either analytically or with numerical methods.

The following theorem communicates how, under exogenous growth, relative scarcity—captured by $\tilde{\beta}\tilde{g}$ below—has drastic implications for shares when the elasticity of input substitution is less than one.

Theorem 1 *Suppose $\varepsilon < 1$ and define $\tilde{\beta} = \beta g_A^{\frac{1-\sigma}{1-\alpha}}$ and $\tilde{g} = g_{A_e}/g_A^{\frac{1}{1-\alpha}}$.*

[20]Boppart and Krusell (2018) argue that a better approximation is that hours fall at a small constant rate that is proportional to the rate of productivity growth. No central results here would change under such an alternative framework.

1. *If $\tilde{\beta}\tilde{g} > 1$, then there is no EBGP and there is a unique ACGP where the energy share is zero and where output, consumption, and capital all grow at the same rate and where the ratio $Ak^\alpha/(A_e e)$ goes to zero.*

2. *If $\tilde{\beta}\tilde{g} = 1$, then there is a unique ACGP and it is an EBGP, where output, consumption, and capital all grow an equal, unique rate and where the ratio $Ak^\alpha/(A_e e)$, along with the energy share, are finite and positive and determined by the initial condition on $A_e R/(Ak^\alpha)$.*

3. *If $\tilde{\beta}\tilde{g} < 1$, then there is no EBGP and there is a unique ACGP where the energy share is one and where output and consumption grow at the same rate but capital grows at a lower rate and where the ratio $Ak^\alpha/(A_e e)$ goes to infinity.*

The key parameter expression $\tilde{\beta}\tilde{g}$ captures the roles of technology growth and energy scarcity: if its value is high, although energy is scarce, it is not (chosen) to go to zero fast enough, relative to the growth of the factor-augmenting technologies, to prevent the long-run energy share from going to zero. The growth rates of the factor-augmenting technologies are, of course, key, because they gauge the relative scarcities of the inputs. The discount factor, β, appears because it is key determinant of how fast energy use is chosen to go to zero (under logarithmic preferences, energy goes to zero at rate β exactly). The theorem also says that balanced input shares will only result under a knife-edge condition.

The proof of this theorem, along with the proofs of our other formal propositions, can be found in our online appendix. The proof is straightforward, but it is worthwhile to describe some of its elements here. It is convenient, first of all, to transform the problem into a potentially stationary one by transforming consumption, output, and capital by dividing by $A_t^{\frac{1}{1-\alpha}}$ and by defining a new discount rate as $\beta g^{\frac{1-\sigma}{1-\alpha}}$. This makes the first element of the production function simply the transformed level of capital, without any input-augmenting technology appearing. The second element of the production function—the energy input in efficiency units—can also be defined to be a potentially stationary variable: $\tilde{e}_t = e_t A_{et}/A_t^{\frac{1}{1-\alpha}}$. That leaves a planning problem that is nonstationary only in that the energy resource constraint now reads

$$\sum_{t=0}^{\infty} \frac{\tilde{e}_t}{\tilde{g}^t} \leq R_0.$$

Given this transformed problem, one can analyze the first-order conditions case by case and establish the claims in the theorem. In the proof, one sees that in the case where the energy share goes to zero, output asymptotically becomes linear in the first production input; here, capital is the "bottleneck", because energy-saving technical change grows so fast. When the energy share goes to one, instead, output becomes linear in the energy input and although Ak^α grows faster than $A_e e$, capital grows more slowly than output and the asymptotic capital-output ratio is actually zero.

Clearly, in the transformed problem planning problem, the case $\tilde{g} = 1$ stands out in that it makes the problem entirely stationary. Therefore, this case also formally coincides with the setting analyzed in Dasgupta and Heal's work (no input-saving technical change, as captured by $g_A = g_{A_e} = 1$), though the stationarity here really derives from a case where the two factor-augmenting technologies grow so as to be fully offsetting. Also, notice that whenever $\tilde{g} > 1$, it is feasible to make the (transformed) energy input into production constant: by selecting it to be sufficiently low ($\tilde{e} = \frac{\tilde{g}-1}{\tilde{g}} R_0$). However, it is only when $\tilde{g}$ reaches $1/\tilde{\beta}$ that this is an optimal choice.

Theorem 1 also says that when an exact balanced growth path exists, although its growth rates are pinned down uniquely by $\tilde{\beta}$ and $\tilde{g}$, the long-run energy share depends on initial conditions. This is in sharp contrast to the results below under endogenous technology, where the long-run share is always uniquely determined, independently of initial conditions.

The requirement that ε be less than one of course captures "sufficient complementarity" and is critical: it implies unique ACGEs with the energy share going to 100% unless the energy-saving technology grows fast enough relative to the capital/labor-augmenting technology. If they both grow at the same rate, we obtain that $\beta g_A^{\frac{1-\sigma}{1-\alpha}+1-\frac{1}{1-\alpha}} = \beta g_A^{\frac{1-\alpha-\sigma}{1-\alpha}}$, which may be larger or smaller than one depending on parameter values, and hence it is nontrivial even in this case whether the long-run energy share goes to zero or one. As we pointed out above, however, in the Dasgupta-Heal case where technology does not grow at all, energy becomes the bottleneck input eventually, dragging output downward at a constant rate determined only by preferences and with a 100% energy share. This outcome points to strong incentives to improve on energy efficiency (or similarly for the case where the long-run energy share goes to zero, to improve on capital/labor efficiency). In the next section, we allow for such a channel: endogenous factor-augmenting technical change.

Comparison to Uzawa (1961) It is useful here to compare to the celebrated Uzawa (1961) result: balanced growth is only possible when the production function features labor-augmenting technical change and no capital-augmenting technical change. Uzawa's result is also a knife-edge case. His result is different in that it mainly discusses the feasibility of constant growth. It is also different, of course, because it considers a broader class of utility functions. Uzawa's knife-edge condition, moveover, only involves a technology parameter (that $g_a = 1$), whereas ours, $\tilde{\beta}\tilde{g} = 1$, also involves preference parameters β and σ.[21] Finally, importantly, in the knife-edge case considered by Uzawa, the long-run input shares are uniquely pinned down, independently of initial conditions, whereas we show that the long-run energy share will depend on initial conditions. In the Uzawa case, the real interest rate is pinned down by the Euler equation, and it implies a value for the ratio $k/(Al)$, where l is labor, and hence the capital and labor shares are determined. Our real interest rate, in contrast, does not pin down the share because the marginal product of capital does not just depend on the ratio of the inputs in the production function.[22] Thus, balanced growth in our model is not consistent with a unique long-run share between zero and one—as observed in data—other than as a result of initial conditions.

4.2 Endogenous technical change

Let us now consider endogenous technology: technology that can be directed toward scarce inputs, if it is in the economy's interest. We will show below that under relatively mild conditions, and unlike in the case of exogenous technical change, an EBGP will exist and be the only ACGP. That is, the long-run outcome will in general be balanced growth and will, in fact, feature $\tilde{\beta}\tilde{g} = 1$, the knife-edge restriction on technology growth rates in Theorem 1. I.e., g_A and g_{A_e} will endogenously adjust to values satisfying this equality. To illustrate the mechanism, we first look at a static model of technology choice and then incorporate it into the dynamic setting studied above.

4.2.1 Static model

We assume the same technology for producing output as in the previous section and we add a frontier from which (A, A_e) can be chosen. We consider a given amount of capital and energy: k and R, respectively. We begin by studying the planner's problem and then turn to perfectly competitive decentralization with joint input and technology choice on the firm level. The latter illustrates the core of our decentralization in the dynamic model.

[21]If one considered exogenous (negative) growth for e, these parameters would not appear.
[22]It depends on $Ak^\alpha/(A_e e)$ but also on $\alpha A k^{\alpha-1}$.

The planner The problem here is thus to maximize

$$\left[(1-\gamma)\left(Ak^{\alpha}\right)^{\frac{\varepsilon-1}{\varepsilon}}+\gamma\left(A_{e}R\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

by choice of A and A^e subject to the technology constraint:

$$G(A, A_e) = 0.$$

We will sometimes refer to G as specifying the *technology technology.* Here, G is strictly increasing in both arguments and hence the choice to select a high level of one of the input-saving technologies comes at the expense of the other. We will assume that G has constant returns to scale and is quasi-concave and twice differentiable. Different assumptions on its curvature will then, as we shall see, deliver qualitatively different outcomes. In the dynamic section we will present a straightforward extension to the case where the technology levels evolve over time and depend on their past values.

Our aim is to consider cases where there is an active tradeoff between the two forms of input saving. This outcome is not a foregone conclusion in this model and we will indeed give examples for which a corner solution is obtained. To make sure that the first-order condition for the technology choice is satisfied with equality, we must first establish a result to that effect. We have the following assumption and result.

Proposition 1 *Suppose that $\varepsilon < 1$. Then the technology choice is unique and interior.*

The proposition is straightforwardly proven, given that G is quasiconcave. It uses fixed levels of the inputs k and R. In a dynamic model, these are both chosen and given the multiplicative nature of the input arguments Ak^{α} and $A_e e$, here specialization can be more attractive. However, this possibility also appears in the competitive equilibrium version of the static model to which we now turn.

Competitive equilibrium We consider a representative firm choosing (k, R, A, A_e) taking prices for the inputs, r and p, respectively, as given. For convenience, we will focus on the cost-minimization version of the firm's problem. Hence consider the cost function

$$C(r, w, p, y; A, A_e) = \min_{k,l,e,A,A_e} rk + wl + pR$$

subject to $F(Ak^{\alpha}l^{1-\alpha}, A_eR) \geq y$ and $G(A, A_e) = 0$, where F again is CES with elasticity parameter ε. Now the main result is the following:

Proposition 2 *Suppose $\varepsilon < 1$. Then the equilibrium allocation coincides with that of the efficient solution and the energy share, e^{share} is given by*

$$\frac{1 - e^{share}}{e^{share}} = \frac{AG_1(A, A_e)}{A_eG_2(A, A_e)}.$$

The proof uses the fact that the cost function for a CES production function resulting from choosing quantities can be derived in closed form, and using this closed form it is straightforward to show that the isoquants in (A, A_e) space are strictly concave. Hence, given a quasiconcave G, the problem is well-behaved. The derivation of the energy share expression is easy. The relative share of the capital/labor composite and energy is given by $AF_1k^{\alpha}/(A_eF_2R)$. The first-order condition for A and A^e in the profit-maximization problem, on the other hand, implies

$$\frac{k^{\alpha}F_1}{G_1} = \frac{RF_2}{G_2}.$$

Hence the relative share becomes the object stated. Notice that this derivation holds regardless of the form for F (assuming constant returns to scale); a specific F—CES with a restriction on the elasticity parameter—is used in the theorem in order to ensure sufficiency of the first-order conditions.

Specialization Violations of the assumptions underlying the theorem can lead to corner solutions and specialization. For example, assume that production is entirely symmetric, with $\alpha = 1$ and a linear F ($\varepsilon = \infty$): output is then proportional to $Ak + A_ee$. Moreover, let G be symmetric and linear. Then one obtains full specialization and which factor is chosen depends on the available amounts of capital and energy (or alternatively, from the firm's perspective, on the prices of these two inputs). The key is to note that the technology levels are endogenous and multiply the input levels in production. In the dynamic model, specialization is of course also possible, but we will focus on the cases where specialization does not occur.

Notice for the static model that the energy share will be given by the simple expression in Proposition 2, which in turn is pinned down by A/A_e, but the optimal value of A/A_e will in general depend on all the parameters of the model. An exception is simple case where G is log-linear, which implies that G is not quasiconcave but this case is nevertheless well-behaved so long as F is CES with a substitution elasticity less than unity; here, the share will just be a function of the exogenous coefficients in G. In the dynamic model, as we shall see, (i) the balanced growth rate share will be pinned down exactly as in the proposition here and (ii) its determinants will be a function of only a small number of parameters.

4.2.2 Dynamic model

We now formulate a dynamic model that includes endogenous technology choice. The extension of our static technology choice to a dynamic one is the following:

$$G(A_{t+1}/A_t, A_{e,t+1}/A_{et}) = 0. \tag{9}$$

I.e., we consider the same function G only with growth rates, not levels, as arguments.

We maintain a production function F with a substitution elasticity that is less than unity and characterize the exact balanced growth path (EBGP) for it, including its long-run energy share. Then we show that the EBGP is the only asymptotic constant-growth path (ABGP). We also look at two examples—special cases of preferences and the technology technology G, one of which is used in our estimation section below—as well as discuss robustness.

In the main analysis based on the general model, we also focus on the planning solution and sidestep any issues coming from suboptimal policy. For the special cases we look at, we also consider a decentralized model. As pointed out above, for the special cases there, we opt for as simple a version as possible, one that builds on learning-by-doing externalities. One of the points of this analysis is to show that, under some conditions, the equilibrium is actually optimal, a result that is somewhat special but captures an important aspect of the setup of the model with directed technical change without an overall choice of the amount of research/technology growth. In any case, the key results in this paper do not hinge on the market version of our economy. Throughout our analysis, we also abstract from the (global) climate externality.[23] These could also straightforwardly be included in the analysis.[24]

Thus, the planning problem here reads

$$\max_{\{c_t, k_{t+1}, e_t, A_{t+1}, A_{e,t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma} - 1}{1-\sigma} \tag{10}$$

[23]See Nordhaus and Boyer (2000) and Golosov et al. (2014).
[24]See Hassler, Krusell, Olovsson, and Reiter (2019).

subject to

$$c_t + k_{t+1} = F(A_t k_t^\alpha, A_{et} e_t) + (1-\delta)k_t$$

and

$$G(A_{t+1}/A_t, A_{e,t+1}/A_{et}) = 0$$

for all t and

$$\sum_{t=0}^{\infty} e_t \leq R_0.$$

Notice here that, under the assumptions made on G above (essentially, that it is increasing in both arguments), one can define an intermediate variable n and functions f and f_e such that the following equation system describes the same technology:

$$A_{t+1}/A_t \equiv g_{A,t} = f(n_t) \tag{11}$$

$$A_{e,t+1}/A_{et} \equiv g_{A_e,t} = f_e(1-n_t). \tag{12}$$

With this formulation, which we will use for much of the analysis below, we interpret n as the share of a fixed amount of R&D resources that is allocated to enhancing the efficiency of the capital/labor bundle; $1-n$ is the fraction allocated toward energy-saving. We use this alternative formulation in the estimation section below, along with specific functional forms (where both f and f_e are increasing and have the same kind of curvature). We also use it to specify the form of externalities in the decentralized model.

An increase in A (A_e) is equivalent to a decrease in the input requirement coefficient for the capital/labor (energy) bundle. By changing n, the planner can direct technical change to either of the two activities. When A grows at a different rate than A_e, the required amount of energy relative to that of the capital/labor bundle changes, for any level of output. Thus, this is a source of factor substitutability in the long run and the longer the time available to adjust the technology levels, the more flexibility/substitutability there is. One of our special cases below will be one where F is Leontief, i.e., where there is no substitutability at all in the shortest run but where due to the endogenous technology choice there is significant substitutability in the long run.

The long-run cost share of the natural resource We begin with the central focus of our paper: long-run natural resource dependence (energy dependence in our application). Theorem 2 below thus shows that the long-run energy share in an EBGP—defined as in

Section 4.1—of such an economy depends exclusively, through the R&D tradeoff, on a small set of model parameters. Thus, if we restrict attention to exact balanced growth, the energy share does not depend directly on the elasticity of substitution between capital/labor and energy, nor on the stock of fossil energy. The theorem also determines the rate at which energy use is chosen to go to zero on the EBGP.

Theorem 2 *On an exact balanced growth path (EBGP) with an interior choice for technology, the following features must hold:*

1. *The two arguments of the aggregate production function, $A_t k_t^\alpha$ and $A_{et} e_t$, both grow at the rate of output g.*

2. *Energy use falls at a constant rate: $\frac{e_{t+1}}{e_t} = \beta g^{1-\sigma}$.*

3. *Technology effort n and the consumption growth rate g are determined by $f_e(1-n)\beta = f(n)^{\frac{\sigma}{1-\alpha}} = g^\sigma$.*

4. *Energy's share of income is exclusively determined by how costly it is to enhance energy efficiency in terms of lost capital/labor efficiency. Specifically, the long-run energy share is implicitly given by equation (13):*

$$\frac{1 - e^{share}}{e^{share}} = -\frac{dg_{A_e}}{dg_A}\frac{g_A}{g_{A_e}}. \tag{13}$$

The proof simply involves working out implications of first-order conditions and constraints under the assumption of exact balanced growth. The first statement follows straightforwardly from imposing exact balanced growth. The second statement, asserting that energy use has to fall at a rate determined by preference parameters only is connected to the Hotelling (1931) theorem: the marginal value of a finite resource should rise at the real interest rate. The gross interest rate will thus both equal $\frac{g^\sigma}{\beta}$, from the Euler equation, and $A_{e,t+1}F_{2,t+1}/(A_{et}F_{2t})$, from the Hotelling equation. The latter must, because of the first statement in the theorem, equal g/g_e on a balanced growth path; hence the second statement follows.

The third statement of the theorem follows directly from the first and second statements: in growth rates, using the functions determining the technology growth rates, the first statement now reads $f_e(1-n)\beta g^{1-\sigma} = f(n)g^\alpha = g$. Rearranged, this delivers the third statement and it allows us to solve for n and g, and thus the key remaining growth rates. The fourth statement, finally, is a dynamic extension of Proposition 2.

Notice that the theorem implies that the steady-state income share of energy is unique and determined independently of initial conditions—unlike in the corresponding EBGP case under exogenous growth. Key for this result is the constant-returns assumption on F, which means that the two arguments of the production function need to grow at the same rate on any exact balanced growth path. This is also a reason why the elasticity between inputs is not relevant in the theorem or in the determination of n and the energy share.

In order to quantify the long-run energy share for a calibrated economy, we thus need to estimate the nature of the tradeoff $\frac{dg_{A_e}}{dg_A}\frac{g_A}{g_{A_e}}$. Specifically, the evolutions of the two technology trends A and A_e need to be separately identified, and although we backed out these series in Section 3, we simply assumed a value for the substitution elasticity between the capital/labor composite and the energy inputs, ε. In Section 5 below, we formally estimate a full model, i.e., taking into account the endogeneity of technology, given assumed parametric forms for f and f_e.

Theorem 2 focuses on EBGPs but leaves open whether there may be long-run constant-growth paths, ABGPs, that do not feature exact balanced growth. We now turn to this issue.

Asymptotic constant-growth paths Let us now assume that the production function is of the CES variety with elasticity of substitution less than one. Let us also place bounds on the relative growth rates on the two technologies. We first define the key objects: $\bar{g}_A \equiv f(1)$ and $\underline{g}_{A_e} \equiv f_e(0)$, along with $\bar{\tilde{\beta}} \equiv \beta \bar{g}_A^{\frac{1}{1-\alpha}}$ and $\underline{\tilde{g}} = \underline{g}_{A_e}/\bar{g}_A^{\frac{1}{1-\alpha}}$; and $\underline{g}_A \equiv f(0)$ and $\bar{g}_{A_e} \equiv f_e(1)$, along with $\underline{\tilde{\beta}} \equiv \beta \underline{g}_A^{\frac{1}{1-\alpha}}$ and $\bar{\tilde{g}} = \bar{g}_{A_e}/\underline{g}_A^{\frac{1}{1-\alpha}}$. The key assumption is as follows.

Assumption 1 $\bar{\tilde{\beta}}\underline{\tilde{g}} < 1$ *and* $\underline{\tilde{\beta}}\bar{\tilde{g}} > 1$.

We now show the following for the dynamic model with endogenous technology.

Theorem 3 *Suppose that* $\varepsilon < 1$ *and that Assumption 1 holds. Then there is a unique ACGP and it is an EBGP.*

The theorem says that with endogenous technology, any asymptotic constant-growth path has balanced input shares. This is quite the reverse of what occurred under exogenous technical change, where balanced growth only obtained under a knife-edge condition on the technology growth rates. That is, the present theorem says that the technology growth rates are in fact chosen so as to satisfy the knife-edge condition, regardless, for example, of the elasticity of substitution ε, so long as it is less than one. The assumptions stated in Assumption 1 ensure ruling out corner solutions where all of the technical change is concentrated on one of the inputs.[25] This result is also reminiscent of Acemoglu's work on technical change (see, e.g., Acemoglu, 2003 and how technical change in a capital vs. labor context will be chosen to be labor-augmenting), though of course here both inputs are changing at endogenous rates.

Below, we supplement the analysis with special cases of our setting where we also compute transition paths and verify that the economy also converges to the exact balanced growth path.[26] First, however, let us discuss decentralized equilibria for our economy.

Competitive equilibrium with learning externalities Recall from our analysis of the static model in Section 4.2.1 that firms were depicted as both acquiring inputs in perfectly competitive markets and choosing their two input-saving technologies so as to maximize profits. In such an equilibrium profits were also zero; conditional on technology levels, firms operate in a standard constant-returns environment and make zero profits, and the directed nature of technology choice that was then added did not change this conclusion. The equilibrium was optimal. In the present section, we formulate a dynamic version of firm competition that has the same features: the firm is price-taking and has a static choice, as a result of which equilibrium profits will still be zero. However, we will consider externalities here and potential sources of inefficiency.

To make the firm's choice static, the assumption is that the dynamic effects of technology choice are not taken into account by firms: they are spillovers. In particular, a firm operating at time $t+1$ will choose the technology it operates at the end of period t but this decision is not dynamic. That is, when a given firm chooses a higher A_{t+1}, at the expense of $A_{e,t+1}$ given the constraint $G(A_{t+1}/A_t, A_{e,t+1}/A_{et}) = 0$, it influences the profits it will make at $t+1$ but the decision does not involve any market transactions at t. The interpretation, in line with that in the static model, is that $G(A_{t+1}/A_t, A_{e,t+1}/A_{et})$ describes a menu of available technologies $(A_{t+1}, A_{e,t+1})$ to operate at $t+1$ given what technologies have been

[25]It is straightforward to study corner solutions appearing when Assumption 1 is not met; we omit this case for brevity.

[26]Acemoglu (2003) provides analytical convergence results for his model in the case of linear preferences; one can provide such results here as well but our quantitative focus in this paper is on utility functions with a preference for consumption smoothing.

in used in the past. Formally we can think of A_t here as the average value of capital/labor-augmenting technologies chosen by firms for operation at t, and similarly for A_{et}: we can write the constraint as $G(A_{t+1}/\bar{A}_t, A_{e,t+1}/\bar{A}_{et}) = 0$, where it is made explicit that a firm chooses its technologies to use at $t+1$, it has only infinitesimal impact on the availabilities of technologies at $t+2$. Yet, firms are all the same and, given our assumptions on F and G, will make the same technology choices and make zero profits, just as in the static model in Section 4.2.1. Our equilibrium definition, formally, is that used in Romer (1986), but it operates as a dynamic externality here and it is directed: by choosing more of A in the current period, firms today make it possible to choose even higher values one period hence, all at the expense of the A_e technology in the future. Thus, it involves a positive spillover for one type of input-saving but, by the very same token, a negative one for the other.

We will omit the formal equilibrium definition here but now use it in a special case.[27]

4.2.3 Example: a log-linear technology technology

Suppose G is log-linear, so that $G(x, y) = 0$ can be written $\log y = a - b \log x$, where a and b are positive constants. We do not go through the equations in detail but merely summarize the main conclusions from the analysis.[28] The case is interesting for three reasons. One is an a priori reason: the log-linear case is, within the class of CES functions for G, the largest departure from a convex maximization problem for the firm in choosing its technology and input levels. The finding that this model leads to an easily analyzed and well-behaved case is thus comforting.

A second attractive feature of the case with a log-linear G is that it generates a closed-form solution for the production function, after the technology levels have been maximized out. More precisely, when the firm chooses A_{t+1} and $A_{e,t+1}$ foreseeing that they will later choose the levels k_{t+1} and e_{t+1}, the reduced-form production function is a Cobb-Douglas function in k_{t+1} and e_{t+1} with the exogenous starting levels A_t and $A_{e,t}$ appearing only in the form of a TFP factor (which itself is Cobb-Douglas).[29] Thus, the model of endogenous technology choice reduces to one of exogenous technical change with well-known and well-behaved properties.

[27] Another interesting special case is studied in our online appendix: the ex-post production function is Leontief, in which case one can obtain analytical solutions for most objects of interest. For an application of this model, see Casey (2019). Finally, we solve a more general specification of the model numerically for transitional behavior in the section below on estimation.

[28] An online appendix contains full derivations for the case where the utility function is logarithmic and there is full depreciation.

[29] A version of this case is studied in Jones (2005).

Third, in the log-linear technology case the competitive equilibrium with dynamic externalities is optimal. That is, the positive dynamic spillover exactly cancels with the negative one because whatever tradeoff there is in the current period in the choice between A_{t+1} and $A_{e,t+1}$ remains in future periods: these two variables appear in TFP terms at $t+2$, $t+3$, and so on, but when their marginal effects on welfare in the current period are equalized, they also become equal in all future periods.

4.2.4 Discussion

In this section, we discuss alternative sources of energy and natural-resource prices, each suggesting interesting extensions.

Alternative energy sources The general production function posited here—$F\left(Ak^{\alpha}l^{1-\alpha}, A_e e\right)$, where F has constant returns to scale—can be thought of more generally than from our limited-resource example. In particular, "e" can be any source of energy, or it could be a function of multiple energy sources. So what if we consider an alternative to fossil fuel: what are the implications then for the future energy share?

Because we focus on balanced growth paths where *both* the capital-labor composite and energy are actively used and command constant (positive) income shares, we can immediately focus on the following equations:

$$g = g_A g^{\alpha} = g_{A_E} g_e.$$

The second of these equations states that the two arguments of the production function, due to constant returns to scale, have to grow at the same rate; the first equation says that this rate also has to equal the growth rate of output. From the first equation, we can deduce that $g = g_A^{\frac{1}{1-\alpha}}$. Using the assumptions on the research technologies, we then have that

$$f(n)^{\frac{1}{1-\alpha}} = f_e(1-n)g_e$$

must hold. That is, given a growth rate for energy, this equation determines how the research input (n) must be allocated. Above, we saw that in the zero marginal cost case (conventional oil), g_e is endogenous and given by a simple function of g, σ, and β. Suppose, instead, that an alternative energy source were considered, and let us look at some different possibilities. First, consider "solar power", and let us treat solar power—when fully developed in a costeffective way—as providing a constant energy flow (fundamentally given by the amount of sunlight reaching earth per unit of time). I.e., for solar power $g_e = 1$ in the long run. Similarly, wind power and power generated by ocean waves arguably involve $g_e = 1$ in the long run. When it comes to other resources in finite supply, such as coal or nuclear power, their long-run values must be the same as that derived for the zero marginal cost case.[30]

Recall that the long-run energy share in our setting will be pinned down by (decreasing in) the balanced-growth value of $-\frac{dg_{A_e}}{dg_A}\frac{g_A}{g_{A_e}}$, a result that holds regardless of the energy source. As already pointed out, this expression only depends on n, since we can write $g_{A_e} = f_e(1 - f^{-1}(g_A))$. How, then, is n (and, thus, the energy share) affected by the type of energy source considered? In our analysis, we focus on assumptions on f and f_e such that $g_{A_e} = f_e(1 - f^{-1}(g_A))$ describes a concave function in the positive orthant: our technological possibility frontier for growth rates. Thus, a higher n (or g_A) implies a higher derivative $\frac{dg_{A_e}}{dg_A}$ in absolute value as well as a higher g_A/g_{A_e}, and hence a higher $-\frac{dg_{A_e}}{dg_A}\frac{g_A}{g_{A_e}}$. In sum, the higher is n, the lower must the long-run energy share be.

Putting this insight together with that above, we conclude that a higher g_e, which obtains to the extent the energy source is not based on a resource in finite supply, must attract R&D away from energy-saving, so n goes up and hence the energy share falls. The extent to which it falls depends on the global properties of f and f_e. This is how the "finiteness" matters for long-run income shares in this model. Of course, there can be long transition periods, and indeed the case we look at below displays a long transition to the steady-state value of the energy share.

Natural-resource prices In our simple model, where the resource is assumed to be costless to extract, optimal/market behavior is in line with the analysis in Hotelling (1931). In this section, we briefly discuss the role of this model element as well as its empirical support.

Our main analysis—that on endogenous, directed energy-saving—does not rely on an assumption that fossil energy is extracted in line with Hotelling's arguments. A finite resource has to eventually have its use converge to zero and an alternative assumption for us would have been to simply take as given some exogenous extraction path, say, asymptotically having its rate go to zero at rate g_e. In fact, none of our main results would change under these assumptions: Theorem 1 would apply again, with the exogenous g_e in place of the

[30]Here as well, these energy sources may involve $g_e > 1$ for a long period of time, as the relevant technologies experience efficiency gains. Arguably, new types of nuclear power, like thorium breeders and fusion power reactors, may have the potential to imply $g_e > 1$ for a period at least as long as the time over which we have seen increasing fossil fuel use.

rate at which energy is optimally extracted in the model (which depends on the preference parameters β and σ). Theorem 2 would also go through, as would Theorem 3. Our choice—to include optimizing behavior for extraction—was mostly dictated by completeness: in a model of the long-run role of energy in the economy it seems desirable to include this choice.

It should also be noted that Hotelling's model of extraction—which is the natural setting for extraction choice—can be criticized, especially for its price implications. The logic in Hotelling's price formula is very powerful: if the resource is extracted at both t or $t+1$, the resource producer has to be indifferent between producing in the two periods, implying—from indifference—that the price at $t+1$ is equal to the price at t times the gross real rate of interest. More generally, the Hotelling price predicts that the marginal revenue—price minus marginal cost—will rise at the real rate of interest. Yet natural-resource prices, that for oil included, tend to have been rather falling, or stationary, in the data over longer periods of time, which is a well-known challenge for theory to explain; see, e.g., Smith (1981) for data and a discussion.

A number of explanations for solving the Hotelling puzzle have been proposed. One is falling marginal costs of extraction; this hypothesis is not unreasonable given significant technical progress, but marginal costs are not directly observable, making evaluation difficult. An explanation consistent with a stationary price also obtains if one regards the real interest rate as roughly equal to zero; arguably the real interest rate has been near zero over significant periods of time.[31] However, significant volatility in resource prices suggests a premium over a riskless rate and hence an upward trend in prices. Restrictions to asset markets can directly invalidate the Hotelling logic; a simple case is that we employ in Hassler and Krusell (2012). Surprise new resource findings—e.g., the discovery of a new large oil field—will exert downward pressure on the price; in general the initial price is higher, the lower is the remaining, unexploited amount of the resource. However, for this hypothesis to explain a prolonged period of stationary or falling prices, one would likely need to appeal to a departure from rational expectations. A departure from rationality can more generally help explain arbitrage-based price theory, but it is still a challenging path for a variety of reasons; for example, the existence of even a small set of rational investors might suffice to reinstate Hotelling pricing. As a summary, in our view it is not obvious how all these suggested factors together might explain the bulk of the departures from the basic Hotelling theory. As for the (stark) price volatility in natural resource markets, at least for oil some

[31] Indeed it becomes zero asymptotically in the above model with exogenous technical change where $\alpha = 1$ and energy-saving technical change is slow enough.

recent quantitative theory suggests plausible explanations.[32] However, it is an open question whether this kind of theory can be combined with long-run (Hotelling-like) finite-resource theory and whether fluctuations in the prices of other natural resources (e.g., most metals) can be made consistent with the same kind of theory. Surely there are very interesting and open quantitative questions in this broad area waiting to be addressed.

5 Estimation

In Section 3 we took a preliminary look at the data from the perspective of an aggregate production function and perfectly competitive input markets.

One could go one step further and formalize a curve-fitting procedure whereby the key parameter—the input elasticity parameter ε—is selected in order to minimize the "fluctuations" in the two latent unobservables, A and A_e. The idea here would be that these variables are technology trends and should not exhibit large short-run movements, in particular not large downward movements. Such a procedure rather straightforwardly implies a tightly estimated value of ε close to zero: without near-Leontief behavior, it is hard to account for the lock-step movements in the fossil share and the fossil price exhibited in Figure 1.

In the present section we instead use the full structure of our dynamic model to estimate the elasticity parameter along with some key R&D parameters. The structural estimation, to be described below, is not only very different in terms of the econometric technique, but it also uses different equations, since it essentially relies on dynamic first-order conditions from resource use and directed technical change. As we shall see, however, the resulting estimate for ε is again robustly near zero.

Our full model is that specified above in (10). To carry out the estimation, however, three changes have to be made to the model relative to the original specification. First, functional forms for the R&D functions $f(n)$ and $f_e(1-n)$ must be specified. Second, shocks need to be added to the model. Here, we consider two shocks: one to the depreciation rate for capital and one to the productivity of the energy-saving technology. Third and finally, the growth model must be made stationary.

Regarding functional forms for $f(n_t)$ and $f_e(1-n_t)$ we assume

$$A_{t+1}/A_t \equiv f(n_t) = 1 + Bn_t^{\phi}, \tag{14}$$

[32] Bornstein, Krusell, and Rebelo (2017) argue that a combination of demand and supply, high short-run price elasticity of demand, and high costs of adjusting quantities in the short run can account for the data.

and

$$A_{e,t+1}/A_{e,t} \equiv f_e\left(1-n_t\right) = 1 + \exp\left(\chi_{e,t}\right) B_e \left(1-n_t\right)^{\phi}, \tag{15}$$

where B, B_e, and ϕ are parameters and $\chi_{e,t}$ is the shock to the productivity of the energy-saving technology. As explained in Section 4.2.2, n denotes the share of a fixed amount of R&D resources (e.g., researchers) that is allocated to enhancing the efficiency of the capital/labor bundle, implying that the share $1-n$ instead is devoted to improving energy efficiency.

To make the model stationary, we define $x_t \equiv k_t^{\alpha} A_t$ to be the first argument of the aggregate production function and then transform by dividing through by it: $\hat{c}_t = c_t/x_t$, $\hat{k}_t = k_t/x_t$, and $\hat{y}_t = A_t^e R_t/x_t$, and we define $\hat{e}_t = e_t/R_t$. Moreover, let $g_{x,t+1} = x_{t+1}/x_t$ and $g_{\hat{y},t+1} = \hat{y}_{t+1}/\hat{y}_t$. Denoting the shock to the depreciation rate for capital by $\chi_{\delta,t}$, the dynamic problem can be restated as

$$\max_{\left\{\hat{c}_t, \hat{k}_{t+1}, g_{x,t+1}, g_{\hat{y},t+1}, \hat{e}_t\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \frac{\left(\hat{c}_t x_t\right)^{1-\sigma}}{1-\sigma}$$

subject to

$$\hat{c}_t + \hat{k}_{t+1} g_{x,t+1} = F(1, \hat{y}_t \hat{e}_t) + (1 - \delta \exp\left(\chi_{\delta,t}\right)) \hat{k}_t, \tag{16}$$

and

$$G\left(g_{x,t+1}^{1-\alpha}\left(\frac{\hat{k}_{t+1}}{\hat{k}_t}\right)^{-\alpha}, \frac{\frac{g_{x,t+1} g_{\hat{y},t+1}}{1-\hat{e}_t} - 1}{\exp(\chi_{e,t})}\right) \equiv \frac{\frac{g_{x,t+1} g_{y,t+1}}{1-\hat{e}_t} - 1}{\exp(\chi_{e,t})} - B_e\left(1 - \left(\frac{g_{x,t+1}^{1-\alpha}\left(\hat{k}_{t+1}/\hat{k}_t\right)^{-\alpha} - 1}{B}\right)^{\frac{1}{\phi}}\right)^{\phi} = 0. \tag{17}$$

Note here that x and $\hat{y}$ are state variables—they are predetermined—and that there are four control variables. Estimation of the model thus amounts to taking first-order conditions for this problem and letting the implied equations confront data. These equations are laid out in the online appendix.

The model is estimated with two methods: a standard Kalman filter on a linearized version of the model and a nonlinear particle filter. The first method is standard in the empirical macroeconomic literature, but the method relies not only on linearization being accurate but also on normally distributed shocks. Particle filters are often applied in cases where the objective is to estimate latent states of a stochastic process; here the latent states are the growth rates of A and A_e.

As observables we use data on the growth rate of output (as defined in the paper) and the growth rate of fossil fuel use to identify the two shocks.[33]

With the linear model, we estimate the parameters ε, B, B_e, and ϕ. With the particle filter, we only estimate ε and B_e, thus implying that we then need to calibrate B and ϕ.[34] As we will see below, the estimates for ε and B_e are very similar for the two estimation procedures.

Three parameters are calibrated in both estimations. We impose $\alpha = 0.30$ (in order to fit the data on the relative capital/labor shares), $\beta = 0.985$ (which implies a growth rate for fossil-fuel use per person of 0.985, which is roughly consistent with the observed growth rate over the last 30 years), and $\gamma = 0.05$ (γ plays very little role and the estimates of φ and B_e only barely depend on it).

The first column of Table 1 presents the priors for the coefficients, indicating the density, mean, and standard deviation. In the benchmark (linear) specification, we use a beta distribution with a relatively large standard deviation for ε. The mean of the distribution is 0.2, which is based on previous studies on the (short-run) elasticity of substitution between capital and energy.[35] We also considered an inverse Gamma distribution that support values up to and over 1 for ε, but this does not change the results in any important way.[36] We also choose a Beta distribution for B and we set the mean to be a little below two percent, which is consistent with the average TFP growth over the period considered.

It is harder to find priors for B_e and ϕ. We choose a mean of 0.20 for B_e, which allows for a relatively low energy share, as observed historically. The posterior estimates do not seem to be particularly sensitive to this value. In the benchmark estimation, we choose a Beta distribution with a value close to one for ϕ.

5.1 Results

The results from the estimation are presented in Table 1.

In general, the estimated parameters are similar in all estimations. In particular, the posterior mean for ε is low. The standard deviations for the shocks χ^e and χ^δ are substantially lower in the non-linear estimation. The estimated research technology parameters will be

[33]The data is discussed in Section 3 as well as in the appendix; as pointed out there, we use a broad fossil-fuel index. We also looked at oil alone and obtained very similar estimates.

[34]The non-linear algorithm has difficulty finding the mode with more parameters, yielding unstable estimates.

[35]See, for instance, Berndt and Wood (1975).

[36]With the non-linear estimation, a somewhat tighter prior is needed for epsilon for the algorithm to find the mode.

Table 1: Prior densities and posterior estimates

Linear estimation						
	Prior			Posterior		
Coefficient	Prior density	Mean	Sd	Mean	Sd	$[10, 90]$
ε	Beta	0.2000	0.1500	0.0541	0.0447	$[0.0000, 0.1159]$
ϕ	Beta	0.9700	0.0200	0.9142	0.0267	$[0.8722, 0.9571]$
B_e	Beta	0.2000	0.0300	0.1762	0.0301	$[0.1261, 0.2244]$
B	Beta	0.0170	0.0010	0.0161	0.0006	$[0.0151, 0.0172]$
$std\left(\chi^e\right)$	Inv. Gamma	0.1000	0.0200	0.1708	0.0286	$[0.1237, 0.2166]$
$std\left(\chi^\delta\right)$	Inv. Gamma	0.6000	0.1000	0.9914	0.0797	$[0.8615, 1.1207]$

Non-linear estimation						
	Prior			Posterior		
Coefficient	Prior density	Mean	Sd	Mean	Sd	$[10, 90]$
ε	Beta	0.1000	0.1000	0.0051	0.0041	$[0.0000, 0.0106]$
B_e	Beta	0.2000	0.0150	0.1932	0.0151	$[0.1684, 0.2177]$
$std\left(\chi^e\right)$	Inv. Gamma	0.1000	0.0200	0.0878	0.0068	$[0.0766, 0.0985]$
$std\left(\chi^\delta\right)$	Inv. Gamma	0.6000	0.1000	0.6638	0.0481	$[0.5842, 0.7397]$

The linear estimation features a a Kalman filter and the non-linear estimation features a particle filter. In both cases, the posterior estimates are from 5 chains with 100,000 draws, where we discard the initial 50,000. The acceptance ratio is around 30 percent in both estimations.

discussed below.

5.2 Long-run implications of energy-saving technical change

The first, and perhaps key, observation above is a robustly estimated ε close to zero: the short-run (annual) substitution elasticity between the capital/labor composite and fossil energy is near zero. This observation also resonates well with the preliminary inspection of the data in Section 3. In this section we move to the long-run implications for energy dependence, which depend heavily on the presence of directed technical change in the saving on inputs. Thus, we begin by discussing the parameter estimates in our R&D specification and how our historical data have influenced them, and we then project forward.

Our estimates of the technology technology frontier—that describing the menu of choices for the growth rates of A and A_e, i.e., capital/labor-saving and energy-saving technology growth, respectively—are not as precise as that for the short-run elasticity parameter ε in terms of magnitudes but they show the clear presence of a tradeoff. First recall Figure 3, which shows the A and A_e series based on a substitution elasticity very close zero, i.e., close to

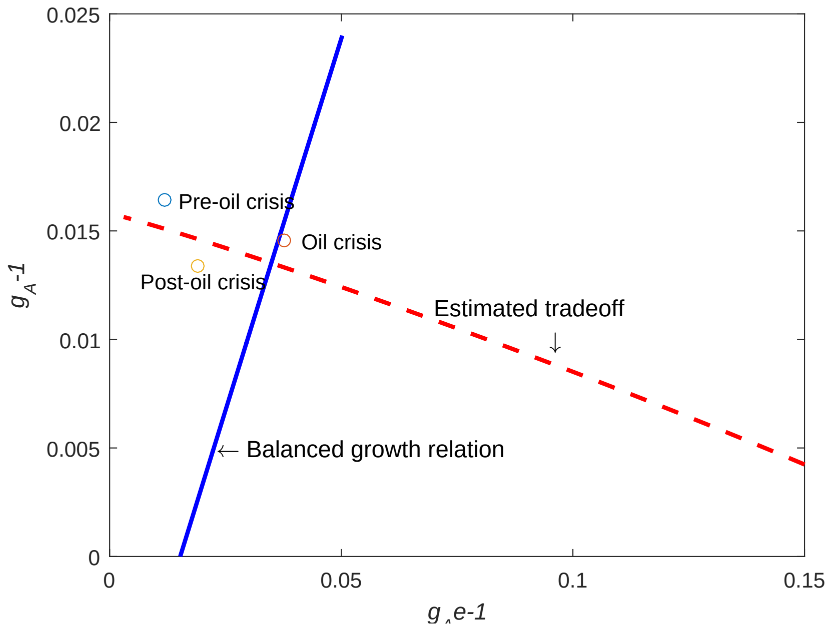

Figure 4: Directed technology tradeoff and balanced-growth relation.

that just estimated: the two series nearly mirror each other, indicating that there is a tradeoff in the direction of technology choice. Thus, in the beginning of the period, the capital/labor-augmenting technology series grows at a relatively fast rate, whereas the growth rate for the energy-augmenting technology is relatively slow. This goes on until around 1970, i.e., somewhat just before the first oil price shock. After 1970, the energy-augmenting technology grows at a faster rate and the growth rate for the capital/labor-augmenting technology slows down. This continues up to the mid-1980s. Hence, the much-discussed productivity slowdown coincides with a faster growth in the energy-saving technology. Note also that this interpretation indicates that there are substantial costs associated with improving energy efficiency, since a higher energy efficiency seems to come at the cost of lower growth of capital/labor-efficiency.

As for the parameters driving this technology tradeoff (B, B_e, ϕ), the formulation in our estimated structure is

$$g_{A^e} = 1 + B_e \left[1 - \left(\frac{g_A - 1}{B} \right)^{\frac{1}{\phi}} \right]^{\phi}. \tag{18}$$

The parameter point estimates imply the line plotted in Figure 4.

The figure also plots the average growth rates for the two technologies during three specific time periods: the pre-oil crisis (1949–1973), the oil crisis (1973–1983), and the post-oil crisis (1983–2011). As can be seen, these points also display a negative relation and they are scattered around the estimated relationship. If one were to estimate a technology tradeoff directly using these three points, thus relying on medium-run fluctuations to identify the relevant tradeoff, the slope would be higher: the scope for energy-saving technical change would be smaller, implying a stronger energy dependence. Still a different method for computing the tradeoff that also would emphasize the medium-run features of the data would be to apply an HP filter to the two series estimated growth series, take out the cyclical component, and then regressing the trend growth of energy-augmenting technology on the trend growth of the capital-augmenting technology. This procedure gives even less scope for energy-saving technical change (in terms of the figure, it gives a higher slope). From the perspective of using lower-frequency movements in the technology series, therefore, our baseline estimate of scope for energy-saving technical change should be viewed as a lower bound.[37]

As for the long-run share implications, we know from the theory developed in Section 4.2.2 that the share can be computed directly from the slope $-\frac{dg_{A_e}}{dg_A}\frac{g_A}{g_{A_e}}$ evaluated on the balanced long-run growth path (recall that the growth rates here are in gross terms). The relation between these gross growth rates is not exactly log-linear, and hence one needs to know at which point to evaluate the derivative. To this end, we find the intersection between the technology tradeoff line with that characterizing balanced growth: on a balanced path, the two production inputs (the capital/labor composite and energy) need to grow at the same rate. The implied relation is $g_{A^e} = \beta^{-1} (g_A)^{\frac{1}{1-\alpha}}$, and it is also plotted in Figure 1 (where we set $\alpha = 0.3$ and $\beta = 0.985$ as in the estimation process). Thus, the long-run equilibrium is found at the intersection of the two lines, which occurs at $g_A = 1.0135$ and $g_{A^e} = 1.0349$. This implies a long-run growth rate of consumption of 1.94 percent per year.

We now compute the required slope by differentiation of equation (18) and evaluation using the obtained long-run growth rates. This implies that $-\frac{dg_{A_e}}{dg_A}\frac{g_A}{g_{A_e}} = 13.7071$, which in turn delivers a long-run energy share of income $e^{share} = \frac{1}{13.7176+1} = 0.0680$. Hence, our findings suggest that energy will earn a higher scarcity rent in the future than now. The estimate of a little below 7% rises if we use the medium-run comovements in technologies as a basis for assessing the long-run tradeoff: the share rises to around 14%. Note, finally, that resource scarcity and the higher energy share do not appear very harmful for economic

[37] Our findings of a negative relation adds macroeconomic support to the findings in Popp (2002), who uses patent data from 1970–1994 to estimate a long run price elasticity between energy prices and energy patents of 0.35. Even though Popp's findings have implications for the impact of factor prices on the direction R&D will take, he does not explicitly compute the tradeoff between the two growth rates for the technologies.

growth, which will be somewhat lower than historically but not by a large amount.

5.3 Transitional dynamics

In this section we complement the long-run analysis just undertaken with a study of the transition path for our economy. One challenge in many standard models of finite resources is that they predict falling resource use from the beginning of time. The model here has the ability to instead generate a rising path initially, namely when the state variables of the system—chiefly the initial levels of technology and capital—are such that capital is relatively scarce (k and A are low relative to A_e). Then, most of the improvements in technology occur by pushing A upward, and over time the energy input hence needs to rise in lock-step, with little need to improve its efficiency. To what extent this occurs, however, is a quantitative question, and the purpose here is to provide a quantitative answer.

To compute the transition path, we use our estimated model and set initial conditions so that $\hat{y}_t$ (the state variable indicating the transformed level of energy efficiency) is 50 percent above its steady-state value whereas $\hat{k}_t$ is 5 percent below its steady state value. These values are selected mostly as an example to illustrate the quantitative magnitudes involved. The starting year is set to 2015 and the results are plotted in Figure 5 below.

The top left graph shows that without any shocks, we can expect 17 years of growing fossil-fuel use (the net growth rate of e is larger than zero during this period). During the transition, e has to grow fast to keep up with the quickly increasing capital and capital/labor-saving technology growth. The top right graph shows that energy's share of income is growing steadily from today's value up toward the long-run value of just below seven percent. The bottom-left graph, finally, shows that the growth rate of output will decline gradually to 1.94 percent/year.

As we have seen, during the transition path the growth rate of fossil-fuel use is quite different from its balanced-growth value. In contrast, the real interest rate does not depart much from its balanced-growth level: it is not off by more than around half a percentage point at any point in time. Thus, our model produces relatively slow convergence in one dimension—fossil-fuel use—while it behaves as a standard neoclassical model in terms of interest rates. This is in line with data, where the secular increase in fossil-fuel use has occurred without any apparent trend in the interest rate. An intuitive explanation for why convergence is faster in the capital-labor ratio than in the ratio of the two technologies is that capital depreciates but knowledge does not.

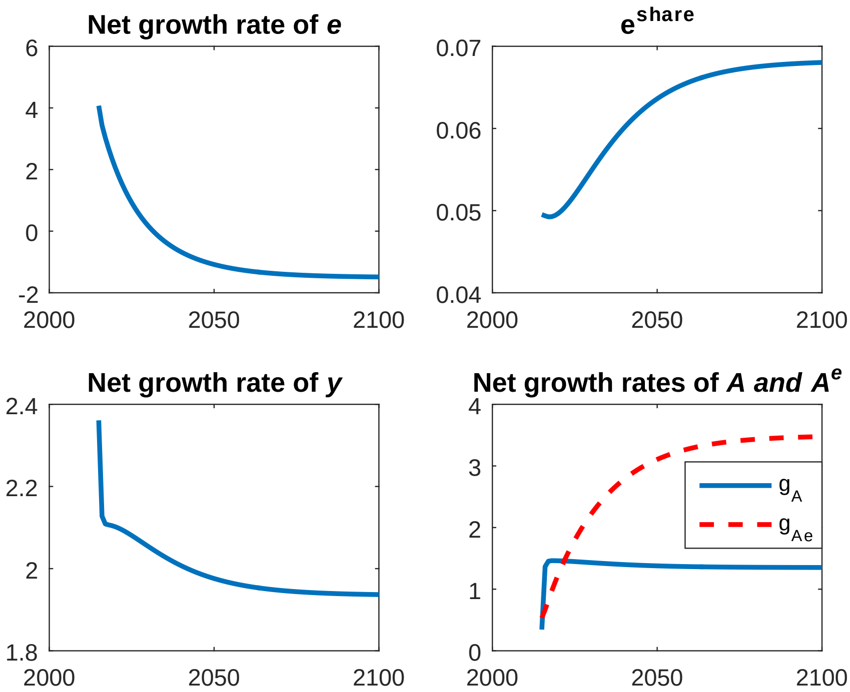

Figure 5: The transition path.

6 Concluding remarks

In this paper we propose a framework for thinking about technological change as an economy's response to the finiteness of natural resources. Through endogenous technical change directed at valuable resources that become scarce, our theory naturally builds in substitutability across (input) goods that is higher ex ante than it is ex post. The theory captures a tradeoff between directing the research efforts toward different inputs and, to the extent the result of these efforts can be measured in the data, we can obtain some insights about the nature of these tradeoffs. We formulate an aggregate production function that is rich enough to allow us to measure the levels of input-saving technologies given data on output, inputs, and prices, and we use it to make quantitative use of our framework.

Along these lines, we thus estimate an aggregate production function in capital, labor, and fossil energy on historical U.S. data in order to shed light on how the economy has dealt with the scarcity of fossil fuel. The evidence we find strongly suggests that the economy actively directs its efforts at input-saving so as to economize on expensive, or scarce, inputs. As outlined, we then use this evidence to inform estimates of R&D technologies, allowing us to make projections into the future regarding energy use and sustainability. Our conclusion is that we can expect fossil fuel to demand a significantly higher share of costs in the future than now; our projection suggests roughly 7 percent as a lower bound but values above 10% are not wholly implausible. The conclusion that the energy share will be significantly higher in the future in the absence of innovation into new sources of energy will not, however, have major implications for consumption growth. Our model implies a long-run growth rate of consumption that is only somewhat lower than in the past—almost 2% per year—so from this perspective the energy dependence looks less problematic.

Our framework is rather aggregate and stylized in nature and any results we derive of course will suffer from not introducing more detail, such as to the energy sector, where not only fossil energy is used but also a range of other energy sources. The setup is transparent and tractable, however, and we think of it as a possible blueprint for addressing sustainability issues more broadly in economics. Thus, it is ready to be applied in different contexts and to be made richer, as we see no conceptual or computational difficulties for most extensions of interest. If the resource in question has a high short-run complementarity with other inputs, much of the qualitative analysis in this paper applies, but the quantitative conclusions can of course differ markedly across different kinds of applications. What we currently regard as the most challenging issue in this area is understanding price trends: for finite resources, Hotelling's (1931) robust logic applies, which is that the marginal profit per unit extracted must grow at the real rate of interest, whereas in the data, natural-resource prices (less marginal costs) do not seem to follow an exponential trend, thus violating Hotelling's rather robust logic. We discuss this issue briefly in Section 4.2.4 of the paper and mention possible paths forward.

An important question in the area of natural resource management and exhaustibility of resources is whether there is a need for government regulation. Our model features no major market failures; in its decentralized version there is an R&D externality but given its directed nature, market outcomes are close to optimal (and exactly optimal in simple versions of the model): the externalities are proportional to both private costs and private benefits and hence nearly cancel. We believe this finding to be rather robust. We do not include any options to increase the overall resources used on R&D in our model; such a (standard, to the endogenous-growth literature) formulation would of course imply a need to subsidize research, at least in the absence of strong negative business-stealing externalities. Our focus here is on the directedness of technical change, however, and our main point is that subsidies may not be necessary for regulating the direction of technical change.

References

Abrell, Jan, Sebastian Rausch, and Hagen Schwerin, (2016), "Long-Run Energy Use and the Efficiency Paradox," Center of Economic Research at ETH Zurich, Working paper 16/227.

Acemoglu, Daron, (1998), "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics* 113, 1055–1090.

Acemoglu, Daron, (2002), "Directed Technical Change," *Review of Economic Studies* LXIX, 781–810.

Acemoglu, Daron, (2003), "Labor- and Capital-Augmenting Technical Change," *Journal of the European Economic Association* 1(1), 1–37.

Acemoglu, Daron, (2007), "Equilibrium Bias of Technology," *Econometrica* 5(75), 1371–1409.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hémous, (2012), "The Environment and Directed Technical Change," *American Economic Review* 102(1), 131–166.

Aghion, Philippe, and Peter Howitt, (1992), "A Model of Growth Through Creative Destruction," *Econometrica* LX, 323–351.

Aghion, Philippe, Antoine Dechezlepretre, David Hémous, Ralf Martin, and John Van Reenen, (2016), "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry," *Journal of Political Economy* 124(1), 1–51.

Atkeson, Andrew, and Patric J. Kehoe, (1999), "Models of Energy Use: Putty-Putty versus Putty-Clay," *American Economic Review*, 89(4), 1028–1043.

Barsky, Robert, and Lutz Kilian, (2002), "Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative," in *NBER Macroeconomics Annual 2001.* Vol 16, ed. Ben S. Bernanke and Kenneth Rogoff, 137–83. Cambridge, MA: MIT Press.

Berndt, Ernst and David Wood, (1975), "Technology, Prices, and the Derived Demand for Energy," *Review of Economic Studies* 57, 259–268.

Boppart, Timo, and Per Krusell, (2018), "Labor Supply in the Past, Present, and Future: A Balanced-Growth Perspective," *Journal of Political Economy*, forthcoming.

Bornstein, Gideon, Per Krusell, and Sergio Rebelo, (2017), "Lags, Costs, and Shocks: An Equilibrium Model of the Oil Industry," NBER Working Paper No. 23423.

Bosetti V., C. Carraro, M. Galeotti, E. Massetti and M. Tavoni, (2006), "WITCH: A World Induced Technical Change Hybrid Model," *Energy Journal*, Special Issue. Hybrid Modeling of Energy-Environment Policies: Reconciling Bottom-up and Top-down, 13–38.

Casey, Greg, (2018), "Energy Efficiency and Directed Technical Change: Implications for Climate Change Mitigation," mimeo, Brown University.

Dahl, Carol, and Thomas Sterner, (1991), "Analysing gasoline demand elasticities: a survey," *Energy Economics* 13 (3), 203–210.

Dasgupta, Partha and Geoffrey Heal, (1974), "The Optimal Depletion of Exhaustible Resources," *Review of Economic Studies* 41, 3–28.

Drandakis, Emmanuel, and Edmund Phelps, (1966), "A Model of Induced Innovation, Growth and Distribution," *Economic Journal* 76, 823–40.

Golosov, Michael, John Hassler, Per Krusell, and Aleh Tsyvinski, (2014), "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica* 82(1), 41–88.

Hassler John, and Per Krusell, (2012), "Economics and Climate Change: Integrated Assessment in a Multi-region world," *Journal of the European Economic Association* 10, 974–1000.

Hassler John, Per Krusell, Conny Olovsson and Michael Reiter, (2017), "Integrated Assessment in a Multi-Region World with Multiple Energy Sources and Endogenous Technical Change," Working paper, IIES, Stockholm University.

Hémous, David, (2016), "The Dynamic Impact of Unilateral Environmental Policies," *Journal of International Economics* 103(C), 80–95.

Hicks, John, (1932), *The Theory of Wages*, London: McMillan.

Hotelling, Harold, (1931), "The Economics of Exhaustible Resources," *Journal of Political Economy* 39, 137–175.

Jones, Charles, (2002), *Introduction to Economic Growth*, Second Edition, W.W. Norton.

Jones, Charles, (2005), "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics* 120 (2), 517–549.

Jorgenson, Dale, and Zvi Griliches, (1967), "Appropriate technology and balanced growth," *Review of Economic Studies* 34 (3), 249–283.

Karabarbounis, Loukas, and Brent Neiman, (2013), "The Global Decline of the Labor Share," *Quarterly Journal of Economics* 129 (1), 61–103.

Katz, Lawrence F., and Kevin M. Murphy, (1992), "1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics* 107 (1), 35–78.

Kennedy, Charles, (1964), "Induced Bias in Innovation and the Theory of Distribution," *Economic Journal* LXXIV, 541–547.

Kilian, Lutz, (2008), "The Economic Effects of Energy Price Shocks," *Journal of Economic Literature* 46(4), 871–909.

Kilian, Lutz, (2009), "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review* 99(3), 1053–1069.

Leon-Ledesma, Miguel, and Mathan Satchi, (2019), "Appropriate Technology and Balanced

Growth". *Review of Economic Studies* 86 (2), 807–835.
Manne, Alan, Robert Mendelsohn, and Richard Richels, (1995), "MERGE: A Model for Evaluating Regional and Global Effects of GHG Reduction Policies," *Energy Policy* 23 (1), 17–34.
Nordhaus, William and Joseph Boyer, (2000), *Warming the World: Economic Modeling of Global Warming*, MIT Press, Cambridge, Mass.
Popp, David, (2002), "Induced Innovation and Energy Prices," *American Economic Review* 92, 160–180.
Romer, Paul M., (1986) "Increasing Returns and Long-Run Growth," *Journal of Political Economy* 94(5), 1002–1037.
Smith, V. Kerry, (1981), "The Empirical Relevance of Hotelling's Model for Natural Resources," *Resources and Energy* 3(2), pages 105–117.
Solow, Robert, (1974), "Intergenerational Equity and Exhaustible Resources," *Review of Economic Studies* 41, 29–45.
Stern, David I. and Astrid Kander, (2012), "The Role of Energy in the Industrial Revolution and Modern Economic Growth," *Energy Journal* 33, 125–152.
Stiglitz, Joseph, (1974), "Growth with Exhaustible Natural Resources: Efficient and Optimal Growth Paths," *Review of Economic Studies* 41, 123–137.
Stiglitz, Joseph, (1979), "A Neoclassical Analysis of the Economics of Natural Resources", in Scarcity and Growth Reconsidered, Resources for the Future," (Ed.) Kerry Smith, Chapter 2, 36–66. Baltimore: The Johns Hopkins University Press.
Uzawa, Hirofumi, (1961), "On a Two-Sector Model of Economic Growth," *Review of Economic Studies* 29(1), 40–47.
Van der Werf, Edvin, (2008), "Production Functions for Climate Policy Modeling: An Empirical Analysis," *Energy Economics* 30(6), 2964–2979.

A Appendix

A.1 Data sources and construction of our variables

In the model, q is a final good used for consumption and capital investment. The inputs are capital, labor, and fossil fuel. By "fossil fuel" here we mean its energy equivalent and we take this measure to equal the fossil energy index (in Btus) from the U.S. Energy Information Agency. We take all the other data from the National Income and Product Accounts.

From the assumption that q is produced from a constant-returns function F, we obtain that the income shares of these inputs sum to unity—in the model. The production of fossil energy is assumed to be at zero cost and is hence treated as a pure rent—a part of capital income. Hence, abstracting from the fact that not all fossil energy produced is used as an intermediate good in domestic production, q will equal GDP, the sum of the payments to labor and capital plus a pure rent (which too can be thought of as capital income). Because fossil energy is also used as a final good by consumers and because some of it is net exported, q will not exactly equal GDP. Denoting these uses by e_c and e_x, respectively, GDP is equal to $q + p(e_c + e_x)$, where p is the price of fossil fuel. The energy share in data we would ideally use is the one corresponding to the share in producing y. So it should be $p\frac{e-e_c}{GDP-p(e_c+e_x)}$, where e is total domestic fuel use. However, given that we do not have data on e_c (except for a shorter period of time), we set $e_c = 0$ in the previous expression for the energy share. The data we do have suggests that the omission of e_c only has a minor level effect on the energy share and no effect on its movements.

A.2 Energy-saving in the manufacturing sector

Figure 6 plots the evolution of the level of the energy-saving technology for the manufacturing sector. The kink is less pronounced, but the energy-saving technology is clearly growing at a faster rate after the oil shocks than before.

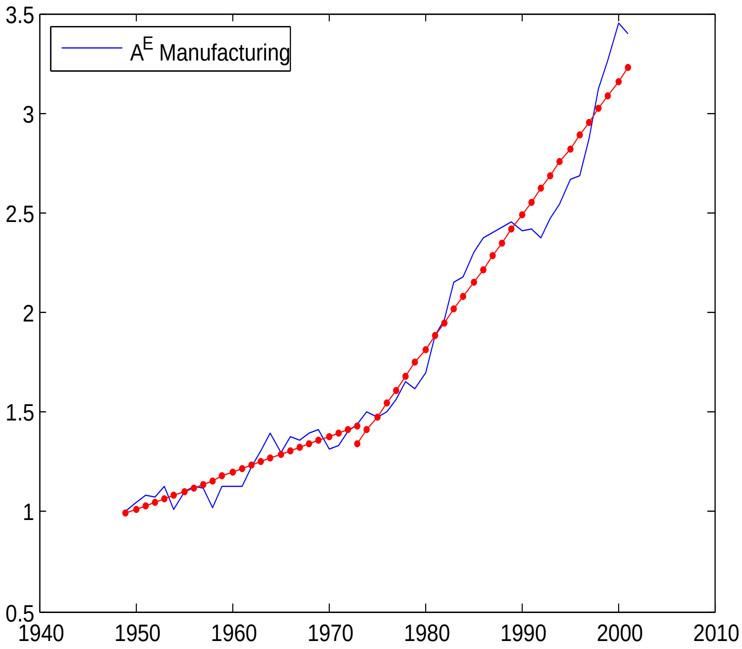

Figure 6: Energy-saving technology with an elasticity of 0.02. Data on industrial-sector energy consumption is taken from the U.S. Energy Information Administration, and the data on output in U.S. manufacturing is from the Board of Governors of the Federal Reserve System (i.e., the FRED database).