ENERGY AND THE DIRECTION OF GROWTH: WILL AUTONOMOUS ENERGY EFFICIENCY IMPROVEMENTS CONTINUE?

Akshay Shanker[†] & David Stern[⋆‡]

Crawford School of Public Policy, Australian National University, 132 Lennox Crossing, ACT 2601, Australia

17 October, 2017

World and U.S. energy intensity has declined over the past century. Because the decline in energy intensity has persisted through periods of stagnating or even falling energy prices, energy economists have suspected factors related to the growth process, independent of prices, have played a role in historical energy efficiency improvements. In this paper, we study the future potential of non-price improvements in energy efficiency using directed technical change theory. We incorporate labor and energy augmenting technical change in a Neo-Schumpeterian endogenous growth model; the model allows us to study the economic trade-offs innovating firms make as they decide whether to pursue energy or non-energy augmenting innovation. We compare two innovation regimes, one with strong scale effects, where the existing body of knowledge enhances the productivity of research and one without scale effects. Our main finding is that whether or not non-price energy efficiency improvements continue in the future depends on the extent to which labor augmenting innovation crowds out energy augmenting innovation, and in the regime with scale effects, energy augmenting innovation eventually declines to zero. By contrast in the regime without scale effects, energy innovation continues albeit at a slower rate than non-energy innovation.

JEL Classification: O33, O41, Q43

[†]Electronic address: `akshay.shanker@anu.edu.au`.

[⋆]Electronic address: `david.stern@anu.edu.au`.

[‡]We thank Sofia Henriques, Astrid Kander, Per Krusell, Warwick McKibbin, and seminar participants at the Scandinavian Energy Transitions Workshop for discussions that have guided the paper. We also thank the Department of Economic History at Lund University for hosting the first author's research visit and gratefully acknowledge funding from the Svenska Handelesbanken Foundation grant on energy and long-run growth.

1. INTRODUCTION

This paper uses the theory of directed technical change to understand the potential future of 'non-price' energy efficiency improvements. The past century has seen a persistent decline in energy intensity across the world. The United States (U.S.) exemplifies the trend, Figure 1 shows U.S. energy intensity and the U.S. real energy price from 1900 to 2015. While the energy price has fluctuated, there is no clear increasing trend. Moreover, even during periods of falling prices, particularly between 1980 and 2000, energy intensity continued decline. As such, many energy economists have suspected non-price, autonomous, factors tied to growth have played role in the decline of energy intensity.

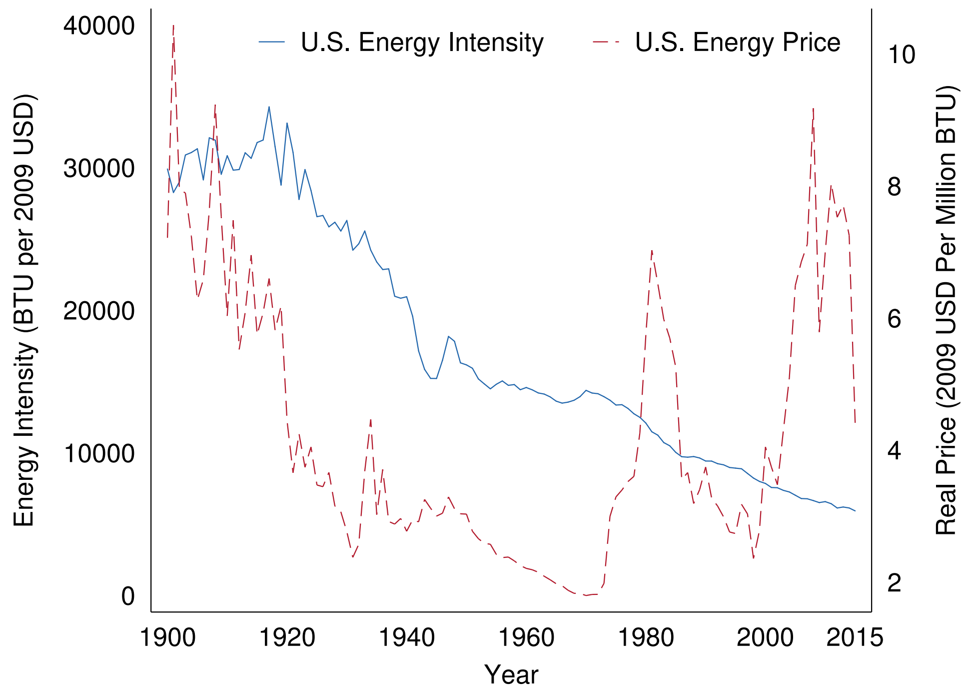

FIGURE 1.— U.S. energy intensity (including traditional fuels) and real price index of fuels between 1900 and 2015. We detect no significant trend in energy prices (see Appendix). Energy intensity fell at an average rate of 1.2 percent per year between 1900 and 1950 and 1.5 percent per year between 1950 and 2015. See the appendix for a description of data sources.

The climate policy modelling literature refers to the non-price decline of energy intensity as Autonomous Energy Efficiency Improvement (AEEI). Though extensively debated, estimates AEEI range from .5% to 2% per year (See Williams et al. (1990), Löschel (2002), Stern (2004), Sue Wing and Eckaus (2007) and Webster et al. (2008)).[1] And climate policy models,[2] and projections of future energy intensity, for example by the International Energy Association (IEA, 2016) and the U.S. Energy Information Administration (EIA, 2017), both assume the secular decline in energy intensity continues in the future. However, despite the importance of energy intensity projections to estimating the response of carbon emissions to policy and despite the vast literature incorporating endogenous technical change in models estimating the costs of climate policy (Goulder and Schneider, 1999; Nordhaus, 2002; Jakeman et al., 2004; Popp, 2004), there has been no systematic theoretical analysis of the economic drivers and future potential of autonomous energy intensity improvements.

Here, we use the theory of directed technical change to understand non-price energy efficiency improvements; if any energy augmenting innovation occurs, the premise of our analysis is that it is due to purposeful inventive activity and firms face a trade-off between energy augmenting innovation and innovation that augments other factors. Our main questions are as follows. What characterizes the trade-off between energy augmenting and non-energy augmenting innovation as an economy grows? In the absence of a sustained increase in energy prices or energy scarcity, when would there be incentive for energy augmenting innovation? And, should we expect a market-driven non-price decline in energy intensity to continue?

Our analysis uses a modern treatment of Schumpetarian endogenous growth with quality improvements (Aghion and Howitt, 1992) due to Acemoglu (2009), ch. 14; we incorporate directed technical change (Acemoglu, 1998, 2002; Acemoglu et al., 2015) into the Schumpeterian framework. There are two factor inputs, energy and labor, and profit incentives drive the decision to improve the quality of energy services production or labor services production.

We characterize growth paths where quality improvements of capital that delivers

[1] Newell et al. (1999), who use micro-level data on efficiency of air-conditioners, find price increases induce technical change but "autonomous drives of energy efficiency explain up to 62% of total changes in energy efficiency". On the other hand, Kaufmann (2004) using a co-integration analysis finds no left-over deterministic trend.

[2] See Popp et al. (2010) for an overview of the use of AEEI in models with exogenous technical change. Some models with induced technical change also incorporate and exogenous efficiency trend, for example, Popp (2004).

energy services (energy augmenting technical change) gives rise to a non-price decline in energy intensity. Along a growth path, the relative profitability between energy and labor augmenting research depends on price and market size effects. As an economy grows, if energy prices do not change, energy use grows[3] and the price of energy relative to labor falls, decreasing incentives for energy augmenting research (price effect). At the same time, the expansion of energy use increases the potential market for energy services, increasing incentives for energy augmenting research (market size effect). Because the elasticity of substitution between energy and labor is less than one (see Stern and Kander (2012) and also Van der Werf (2008)), the price effect dominates; labor is relatively scarce and there are always stronger incentives to undertake labor augmenting research (Acemoglu, 2002)).

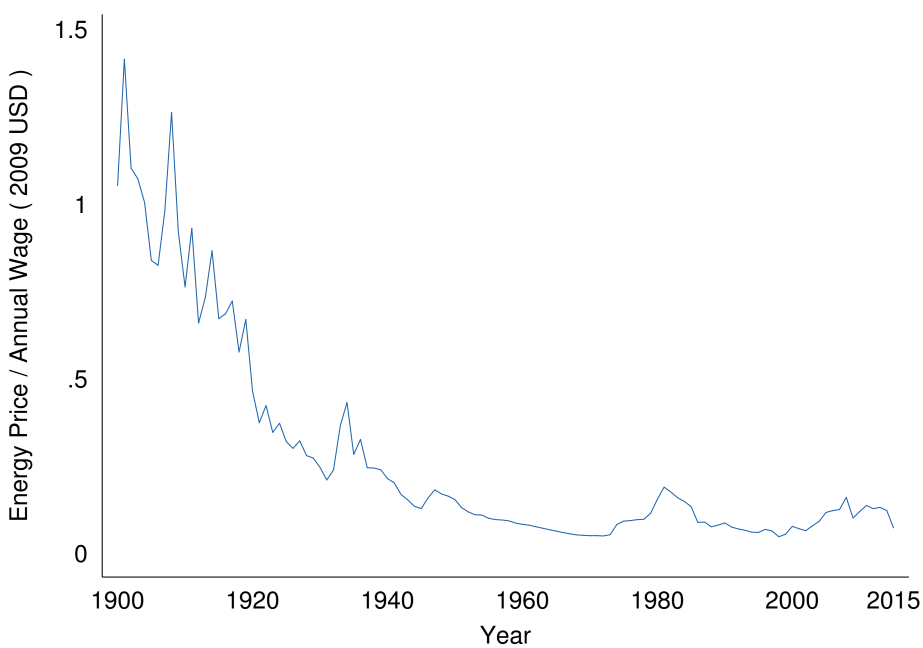

FIGURE 2.— The ratio of U.S. real energy price to real annual wage per employee shows a declining trend.

[3]Initially, the growth of energy occurs both because of energy augmenting innovation and labor augmenting innovation. However, we show once the cost share of energy innovation is small enough, energy augmenting innovation reduces energy use, *ceteris paribus*. That is, the rebound effect is less than unity. However, energy use still grows as labor augmenting technologies grow faster than energy augmenting technologies.

Our key finding is that the possibility of future sustained non-price induced energy intensity improvements depends on whether the price effect leads to non-energy innovation crowding out energy innovation *completely* or whether it only lowers the growth rate of energy innovation. The extent of crowding out depends on path dependence — existing level of knowledge for a particular type knowledge enhancing the productivity of research for that knowledge type (Aghion et al. (2016) find evidence of path dependence for 'clean' and 'dirty' technologies in the automotive industry). Under path dependence, the relatively stronger incentive to undertake non-energy research arising from price effects are compounded. And, under path dependence, the incentives to undertake labor augmenting research are strong enough to eventually crowd out all energy augmenting research, leading to an eventual stop in the non-price decline of energy intensity. On the other hand, without path dependence, energy intensity continues to decline, albeit at a rate slower than output growth.

A feature of growth with non-price energy efficiency improvements is that it is *inconsistent with balanced growth*. Rather, non-price energy efficiency improvements are associated with structural change where energy services grow faster than labor services, but the fall in the price of energy services results in a fall in the cost share of energy and energy services. While we rule out balanced growth along a path with non-price energy efficiency improvements, the economy converges asymptotically to a balanced growth path (BGP). Under path-dependence, the BGP features a constant share of energy and under no path dependence, the BGP energy share is zero.

The directed technical change model agrees with past trends of energy intensity around the world, with or without path dependence. In particular, the fact that energy intensity declines faster in countries with higher energy intensity, leading to convergence of energy intensity for given levels of output (see Stern et al. (2016) for evidence of convergence). Notwithstanding, to explain the persistent decline in energy intensity, if we assume path dependence, we also have to assume that most countries initially are away from the BGP and are sufficiently energy inefficient.

The data on world wide energy intensity trends, however, is not conclusive on whether energy efficiency improvements will continue if prices did not rise in the future. For instance, we see some countries with low energy intensities such as Spain, Argentina and Italy that have seen no decline in energy intensity for the past 50 years, suggesting at first a path dependence regime. However, this observation does not necessarily imply a path dependence scenario; under no path-dependence, countries may feature no decline in energy intensity if energy intensity is low enough and research resources are too low.

One piece of evidence favouring a path dependence regime is the relatively high elasticity of energy intensity with respect to output in the cross-section for high income countries — energy intensity relative to output 'falls' at a relatively slow rate as we move across high income countries. The high elasticities can only be supported in a model with persistent AEEI if the elasticity of substitution is higher than previous estimates in the literature (see Stern and Kander (2012) or Hassler et al. (2016)) and higher than what is required for a smooth evolution of energy augmenting technologies in the data (we discuss these issues in detail in section 4).

Related Literature

Our paper complements Hassler et al. (2016) and André and Smulders (2014) who also analyze models with endogenous energy and labor augmenting technical change. Hassler et al. (2016) provide a calibrated and tractable framework to forecast the role of energy efficiency improvements in generating future possibilities of growth where energy is scarce. André and Smulders (2014) uses a directed technical change model to account for stylized facts in U.S. energy use. However, both authors do not explicitly consider the role and potential for autonomous drives of energy intensity as we do here. Both papers also only study innovation under path-dependence and use Hotelling rule featuring exponentially increasing prices, which become the driver of long-run energy efficiency improvements.

The concept of asymptotic convergence to balanced growth we use is similar to Acemoglu (2003). Acemoglu (2003) shows how long-run technical change can only be labor augmenting, rather than capital augmenting for interest rates to remain constant. Acemouglu's argument features path dependence, our model under path dependence is similar to Acemoglu's, except with capital replaced by energy. A specification with no path dependence effects is ruled out by Acemoglu precisely because it would lead to a fall in capital intensity given constant interest rates.

Finally, our paper relates to empirical work by Sue Wing (2008) showing most of the decline in U.S. energy intensity before the 1970s oil shock can be attributed to structural change towards sectors with lower energy intensity, while falls in intensity during the period of increasing prices can be attributed to within sector declines in intensity. Mckibbin et al. (2004) take a similar view, arguing the relationship between output and energy use depends on changes in the shares of different sectors in an economy, rather than just on an energy efficiency trend within sectors.

These arguments may at first seem inconsistent with our argument, where structural change implies an increase in the quantity share of energy services to output. However, energy augmenting technical change in models with high aggregation such as ours should be seen as any change in the economy that allow more output per unit of energy, including technical change leading to structural change,[4] than strictly as 'energy efficiency' within specific industries. The results by Sue Wing (2008) also point to an important difference between price induced technological change non-price induced change, whereas in our aggregated model, both are captured by the same variable.

2. TRENDS IN ENERGY INTENSITY — ENERGY INTENSITY DECLINES FASTER IN COUNTRIES WITH HIGHER ENERGY INTENSITY

Stern et al. (2016) establish a stable relationship between growth and declining energy intensity. The relationship holds both as an economy grows and in a cross section. Stern et al. (2016) also show evidence of convergence of energy intensities both conditional and unconditional on income. In this section, we build on the results, in particular we argue energy intensity declines faster in economies with higher energy intensity; we will use the observation to guide our analysis of transition paths in the theoretical model.

Figure 2 shows the relationship between energy intensity and output between 1900-2010 for seven countries. A visual inspection of the graph suggests energy intensity converging for a given level output. The elasticity of energy intensity to output is highest for the United States, which starts off with the highest energy intensity. Between 1900 and 1950, energy intensity declined by .52 percentage points for each percentage point of output growth, an elasticity of -.52, and between 1950 to 2015, the elasticity was -.72. Countries such as Sweden, which are initially less energy intensive, show a modest decline (elasticity of -.01 for 1900 and 1950 and -.005 between 1950 and 2015); on the other hand, Argentina, Spain and Italy have almost no change in their level of energy intensity along the growth path. For example, in Spain, the elasticity was positive, at .3 and between 1950 and 2010, modest at -.001.

The Appendix also examines the relationship between output and energy intensity for cross-sections of countries through time; we show the relationship is stable,

[4] The definition of energy services in an aggregated directed technical change model does not map any particular sector in the economy, but may resemble a notion broader than "useful work" within an economy, which has shown an increasing trend, see Warr et al. (2010).

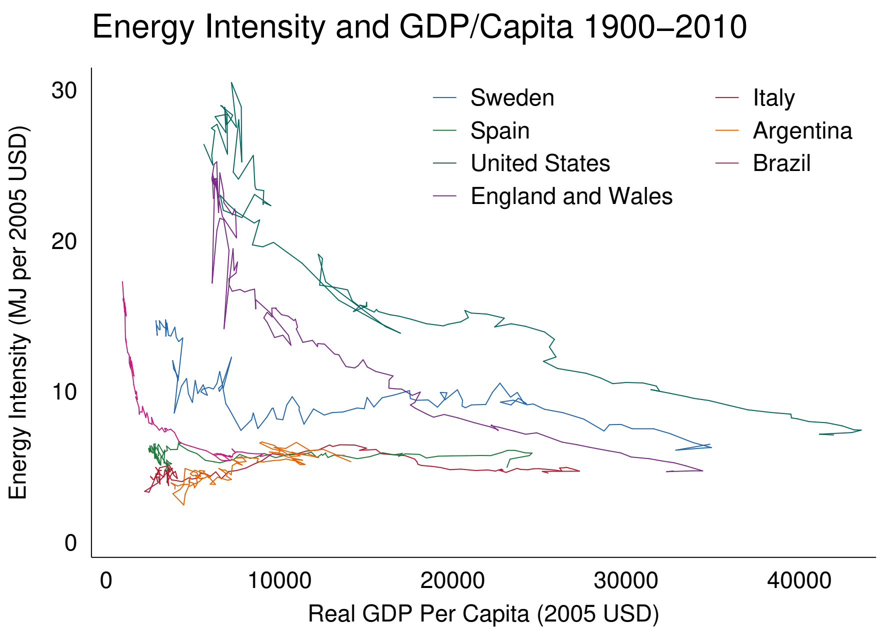

FIGURE 3.— Energy intensity and real GDP per capita for selected countries between 1900 and 2010.

higher output is associated with lower energy intensity. However, once again, for higher levels of output (and lower levels of energy intensity), the elasticity energy intensity respect to output falls close to zero. Moreover, note in figure 2, we see energy intensity does not fall for economies with very low energy intensity. As such, while we see a negative relationship between energy intensity and output, the data also suggests the possibility that the decline in energy intensity may converge to zero as energy intensity approaches a certain level (about 5 MJ/ 2005 USD) — however, it is not clear whether after more growth the countries with lower energy intensities will see falling intensities or whether all countries will converge to this level.

To confirm the intuition from our graphical inspection, we turn to data on a larger set of 100 countries, between 1971 and 2010. The details of the data sources are available in Stern et al. (2016). (Figure B.2 in the Appendix shows a the relationship between energy intensity and GDP per capita for key countries between 1971 and 2010.) Figure 2 shows the elasticity of the average annual rate of energy intensity decline with respect to average annual output growth for 100 countries between

1971 and 2010. The mean elasticity is -.28 and the negative relationship between elasticity and energy intensity in 1970 is statistically significant.

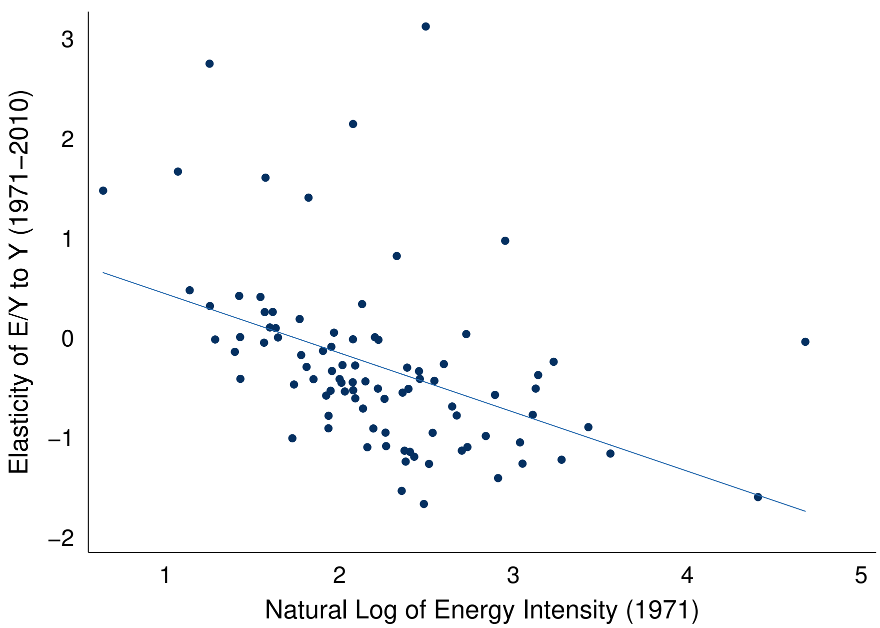

FIGURE 4.— Elasticity of average rate of decline in energy intensity to average rate of output growth between 1971 and 2010. The line of best fit has a slope coefficient of -0.59 (std. err. 0.12).

3. NEO-SCHUMPETERIAN GROWTH MODEL WITH ENERGY AND DIRECTED TECHNICAL CHANGE

3.1. *The Model Environment*

Consider a continuous time economy where t, with $t \in [0, \infty)$, indexes time. Assume a representative consumer with linear utility, preferences over a time path for consumption are given by[5]

$$\int_{t=0}^{\infty} e^{-\rho t} C(t) \mathrm{d}t \tag{1}$$

[5]Linear utility simplifies analysis of the dynamics in the model. Aghion and Howitt (1992) makes the same assumption as does Acemoglu (2003) when analyzing dynamics.

where ρ is the discount factor. The resource constraint for the economy is

$$(2) \qquad Z(t) + C(t) + X(t) + \kappa(t)E(t) \leqslant Y(t)$$

where $Z(t)$ is the total level of R&D, $X(t)$ is total expenditure on machine varieties, $C(t)$ is consumption, $E(t)$ energy use, and $\kappa(t)$ is the extraction cost of energy.

At each t, we assume final goods $Y(t)$ are produced using two intermediate goods: an energy good $Y_E(t)$ and a labor good $Y_L(t)$

$$(3) \qquad Y(t) = \left(\gamma_E Y_E(t)^{\frac{\epsilon-1}{\epsilon}} + \gamma_L Y_L(t)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}$$

where ϵ is the elasticity of substitution and we assume $\epsilon < 1$. Competitive firms produce the intermediate goods using a continuum of machines indexed by i, where $i \in [0, 1]$. The intermediate production functions are

$$(4) \qquad Y_E(t) = (1-\alpha)^{-1}\, E(t)^{\alpha} \int_0^1 q_E(i,t)^{\alpha} x_E(i,t)^{1-\alpha} \mathrm{d}i$$

and

$$(5) \qquad Y_L(t) = (1-\alpha)^{-1}\, L(t)^{\alpha} \int_0^1 q_L(i,t)^{\alpha} x_L\,(i,t)^{1-\alpha}\, \mathrm{d}i$$

Let j index the sectors E and L. In the equations above, $x_j(i,t)$ is the quantity of machine type i in sector j. The term $q_j(i,t)$ denotes the quality of machine i at time t in sector j.

We assume labor supply is fixed with $L(t) = L$ and $L > 0$ for all t. The energy sector uses a fuel $E(t)$, which can be extracted at a fixed cost $\kappa(t)$. We use an exogenous extraction cost as our focus is on non-price induced technical change – we are interested in the response of technological change *given a path of prices*.[6].

The price of final output is normalized to one and all prices will be stated in terms of the final good. We assume that machines of all varieties have a constant production cost $(1-\alpha)$, and monopolists who own patents for the varieties make and sell machines to the intermediate producers.

[6]Some other researchers e.g. André and Smulders (2014) assume that the price follows a Hotelling rule. However, observed energy prices are inconsistent with a simple Hotelling rule which would imply an exponentially growing price (Hamilton, 2009)

Let $p_j^x(t,i|q)$ denote the price of machine i in sector j at time t with quality $q(i,t)$. And let $p_j(t)$ denote the price of the intermediate good $Y_j(t)$.

New entrants engage in research to improve machine varieties. We assume a quality ladder for each machine type as follows

$$q_j(i,t) = \lambda^{n_j(i,t)} q_j(i,0), \qquad j \in \{E, L\} \tag{6}$$

where $q_j(0,i)$ is the quality of machine i at time 0 and $n_j(t,i)$ equals the random number of incremental innovations on the machine variety up to time t. The arrival of a new innovation improves the machine quality by a factor λ. New entrants who have a successful innovation own a perpetual patent on the machine variety, however, once a new variety has been invented, the improved quality captures the market for the variety — Schumpeterian creative destruction. The following assumption [7] ensures the firm with the highest quality machine can charge the unconstrained monopoly price:

ASSUMPTION 1 $\quad \lambda \geqslant \left(\frac{1}{1-\alpha}\right)^{\frac{1-\alpha}{\alpha}}$

For each i, $q_j(i,t)$, is a random process. However, if $n_j(i,t) = n_j(t)$ for each t and j, a law of large numbers argument, the average machine quality

$$Q_j(t) := \int_0^1 q_j(i,t)\,\mathrm{d}i \tag{7}$$

will be deterministic,[8] determined by the innovation rate n_j as follows

$$\dot{Q}_j(t) = n_j(t) Q_j(t) \tag{8}$$

We now describe the innovation possibilities frontier (IPF) for two regimes, a regime without path dependence and a regime with path dependence. In the regime without path dependence, increasing RD effort is required to further improve machine quality as machine quality increases over time. Recent empirical evidence shows that increasing RD effort is needed to find new innovations Bloom et al. (2017).

---

[7]See Acemoglu (2009) sections 12.3.3 and 14.1.2 for discussion.

[8]More precisely, the average quality will equal $\mathbb{E}q(i,t)$ with probability one (Sun and Zhang, 2009). Acemoglu and Cao (2015) (Footnote 23) provide a concise derivation of Equation (8).

3.1.1. *Innovation Possibilities Frontier Without Path Dependence*

Final goods are used to undertake R&D which can be directed towards energy or labor machine improvements. A prospective entrant expending $z_j(i,t)$ units of R&D to improve the equality of machine i in sector j generates a flow rate of machine improvement equal to $\eta_j b_j(i,t)$, where η_j is an exogenous parameter and[9]

$$b_j(i,t) :\ = \frac{z_j(i,t)}{q_j(i,t)}$$

The denominator captures the idea that higher quality machines are more difficult to improve upon, cancelling out scale effects embedded in the quality ladder.

Total R&D expenditure will be

$$Z(t) = \int z_E(i,t)\,\mathrm{d}i + \int_0^1 z_L(i,t)\,\mathrm{d}i \tag{9}$$

3.1.2. *Innovation Possibilities Frontier With Path Dependence*

In the regime with path dependence, we assume a constant unit measure of scientists who undertake research directed towards energy or labor machine improvements. If a scientist works to improve the equality of machine i in sector j, the flow rate of machine improvement is η_j.

We assume prospective entrants can only choose to work in a sector j instead of choosing the specific machine variety i to improve. Once an entrant chooses j, they are randomly allocated to a machine variety i with no congestion, such that each variety i has at most one scientist allocated to it at t. Let s_j denote the measure of scientists working to improve machines in sector j. The flow rate of innovations in sector j will be $s_j \eta_j$. [10]

---

[9] Recall the flow rate tells us the probability of a machine improvement occurring during a small period of time Δt given R&D expenditure $z_j(i,t)$ is $\frac{\eta_j z_j(i,t)}{q_j(i,t)}\Delta t$. Formally,

$$\mathbb{P}\{n_j(t+\Delta t, i) - n_j(t,i) = 1\} = \frac{\eta_j z_j(i,t)}{q_j(i,t)}\Delta t + o(\Delta t)$$

[10] The random allocation of scientists across machine varieties ensures all varieties experience innovation in equilibrium. Acemoglu et al. (2012) make a similar assumption in discrete time. The assumption here would not be required if we assumed a product improvement environment, however, the Schumpeterian environment simplifies exposition of dynamics.

3.2. *Equilibrium Definitions and Characterization*

An allocation consists of time paths for consumption, aggregate spending on machines, energy and aggregate R&D expenditure $[C(t), X(t), E(t), Z(t)]_{t=0}^{\infty}$, R&D expenditure or scientists in the energy and labor sectors $[z_E(i,t), z_L(i,t)]_{i\in[0,1],t=0}^{\infty}$ or $[s_E(t), s_L(t)]_{t=0}^{\infty}$, prices and quantities of machines $\left[p_E^x(i,t\,|\,q), x_E(i,t)\right]_{i\in[0,1],t=0}^{\infty}$ and $\left[p_L^x(i,t\,|\,q), x_L(i,t)\right]_{i\in[0,1],t=0}^{\infty}$, value functions for monopolists for each machine in each sector $[v_L(i,t\,|\,q)]_{i\in[0,1],t=0}^{\infty}$ and $[v_E(i,t\,|\,q)]_{i\in[0,1],t=0}^{\infty}$ and interest and wage rates $[r(t), w(t)]_{t=0}^{\infty}$.

DEFINITION 3.1 An equilibrium is an allocation such that:

1. incumbent firms pick prices and machine quantities to maximize profits
2. final and intermediate goods producers maximize their profits
3. consumers maximize their utility
4. the final goods market clears
5. without path dependence, $s_E(t) = s_L(t) = 0$ and entrants choose R&D to maximize their net present value given the IPF in section 3.1.1
6. with path dependence, $Z_E(t) = Z_L(t) = 0$ and the allocation of scientists satisfies $s_E + s_L = 1$ and entrants choose a sector to innovative to maximize their net present value given the IPF in section 3.1.2.

We begin by documenting the familiar static conditions of the equilibrium, given technology.

3.2.1. *Static Equilibrium Conditions*

Profit maximization by final goods producers gives

$$\gamma_E Y(t)^{\frac{1}{\epsilon}} Y_E(t)^{-\frac{1}{\epsilon}} = p_E(t), \qquad \gamma_L Y(t)^{\frac{1}{\epsilon}} Y_L(t)^{-\frac{1}{\epsilon}} = p_L(t) \tag{10}$$

Profit maximization by the intermediate goods producers gives machine demands

$$x_E(i,t) = p_E(t)^{\frac{1}{\alpha}} p_E^x(i,t\,|\,q_E)^{\frac{1}{\alpha}} q_E(i,t) E(t) \tag{11}$$

$$x_L(i,t) = p_L(t)^{\frac{1}{\alpha}} p_L^x(i,t\,|\,q_L)^{\frac{1}{\alpha}} q_L(i,t) L \tag{12}$$

and first order conditions for energy and labor

$$\alpha(1-\alpha)^{-1}p_E(t)E(t)^{\alpha-1}\int q_E(i,t)^{\alpha}x_E(i,t)^{1-\alpha}\,\mathrm{d}i = \kappa(t) \tag{13}$$

$$\alpha(1-\alpha)^{-1}p_L(t)L^{\alpha-1}\int q_L(i,t)^{\alpha}x_L(i,t)^{1-\alpha}\,\mathrm{d}i = w_L(t) \tag{14}$$

Since the cost of machine varieties is $1-\alpha$ and by Assumption 1, monopolists who own the machine production technologies set $p_E^x = p_L^x = 1$, machine demands are

$$x_E(i,t) = p_E(t)^{\frac{1}{\alpha}}q_E(i,t)E(t), \qquad x_L(i,t) = p_L(t)^{\frac{1}{\alpha}}q_L(i,t)L \tag{15}$$

and monopolists' profits are

$$\pi_E(i,t\,|\,q) = \alpha p_E(t)^{\frac{1}{\alpha}}q_L(i,t)E(t), \qquad \pi_L(i,t\,|\,q) = \alpha p_L(t)^{\frac{1}{\alpha}}q_L(i,t)L \tag{16}$$

Note that profits are conditioned on the quality of machine type i in each sector, $q_j(i,t)$.

Define $Q_E\colon = \int_0^1 q_E(i,t)\,\mathrm{d}i$ and $Q_L\colon = \int_0^1 q_L(i,t)\,\mathrm{d}i$ as the average qualities of machines. Intermediate output, using (4) and (5) with (15) becomes

$$Y_E = (1-\alpha)^{-1}E(t)p_E(t)^{\frac{1-\alpha}{\alpha}}Q_E(t), \qquad Y_L = (1-\alpha)^{-1}Lp_L(t)^{\frac{1-\alpha}{\alpha}}Q_L(t) \tag{17}$$

Next, using the intermediate demand conditions, (10), we can write the ratio of prices as

$$\frac{p_E(t)}{p_L(t)} = \gamma\left(\frac{Y_E}{Y_L}\right)^{-\frac{1}{\epsilon}} = \gamma\left(\frac{E(t)}{L}\right)^{-\frac{1}{\epsilon}}\left(\frac{p_E(t)}{p_L(t)}\right)^{-\frac{1-\alpha}{\alpha\epsilon}}\left(\frac{Q_E(t)}{Q_L(t)}\right)^{-\frac{1}{\epsilon}}$$

and solving for the ratio of prices gives

$$\frac{p_E(t)}{p_L(t)} = \gamma^{\frac{\alpha\epsilon}{\sigma}}\left(\frac{E(t)Q_E(t)}{LQ_L(t)}\right)^{-\frac{\alpha}{\sigma}} \tag{18}$$

where $\sigma = 1+\alpha(\epsilon-1)$ and $\gamma = \frac{\gamma_E}{\gamma_L}$ is the elasticity of substitution between the factors of production. Using (17), we can also derive a ratio of intermediate goods

$$\frac{Y_E(t)}{Y_L(t)} = \left(\frac{E(t)Q_E(t)}{LQ_L(t)}\right)^{\frac{\epsilon\alpha}{\sigma}}\gamma^{\frac{\epsilon(1-\alpha)}{\sigma}} \tag{19}$$

Next, insert (10) into (17), use the expression for final output [11] and the ratio of intermediate goods above at (19) to derive

$$(20)\qquad Y(t) = (1-\alpha)^{-1}\left[\gamma_E^{\frac{\epsilon}{\sigma}}\left(E(t)Q_E(t)\right)^{\frac{\sigma-1}{\sigma}} + \gamma_L^{\frac{\epsilon}{\sigma}}\left(LQ_L(t)\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

We now derive a condition characterizing energy intensity. Use (13) and the definition of $Y_E(t)$ to arrive at

$$(21)\qquad \alpha\theta_E(t)\frac{Y(t)}{E(t)} = \kappa(t)$$

where

$$(22)\qquad \theta_E(t) := \frac{p_E(t)Y_E(t)}{Y(t)} = \left(1+\gamma^{-\frac{\epsilon}{\sigma}}\left(\frac{E(t)Q_E(t)}{Q_L(t)L}\right)^{\frac{1-\sigma}{\sigma}}\right)^{-1}$$

is the cost share of energy services in final production. The equality above comes from using (10) and (19). Use (22) along with (20) to write (21) as

$$(23)\qquad \gamma_E^{\frac{\epsilon}{\sigma(\sigma-1)}}\alpha(1-\alpha)^{\frac{\sigma-1}{\sigma}}\left(\frac{Y(t)}{E(t)}\right)^{\frac{1}{\sigma}} Q_E(t)^{\frac{\sigma-1}{\sigma}} = \kappa(t)$$

The above equation tells us that energy augmenting technical change is energy saving. That is, for a given level of output and real energy price, higher $Q_E(t)$ leads to lower energy use. Given a real energy price, we can decompose the effect of energy augmenting technologies into the change in energy use given a level of output and the change in energy use as output increases given a level of energy intensity:

$$(24)\qquad \frac{\partial E}{\partial Q_E} = \overbrace{\frac{\partial Y}{\partial Q_E}Q_E^{\sigma-1}}^{\text{Output effect}} + \overbrace{(\sigma-1)Y_E Q_E^{\sigma-2}}^{\text{Efficiency effect}}$$

In the derivative above, we have dropped the positive constant at the beginning of the LHS of (23) and the time index for simplicity. The efficiency effect is always negative, and the output is always positive. The following proposition tells us when increased energy efficiency results in a fall in energy use.

[11]Note $Y(t) = \left(1+\gamma^{-1}\left(\frac{Y_L(t)}{Y_E(t)}\right)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}} Y_E(t)\gamma_E$

PROPOSITION 3.1 *If* $\theta_E < (1-\sigma)(1-\alpha)^{\frac{1-\sigma}{\sigma}}$*, then* $\frac{\partial E}{\partial Q_E} < 0$.

PROOF: Evaluate the derivative at (24) and simplify

$$\frac{\partial E}{\partial Q_E} = Q_E^{\sigma-2} Y \left(\gamma_E^{\frac{\epsilon}{\sigma}} Y^{\frac{1-\sigma}{\sigma}} (Q_E E)^{\frac{\sigma-1}{\sigma}} + \sigma - 1 \right) \tag{25}$$

Now note the expression for output at (20) and θ_E at (22) to arrive at the result.

Q.E.D.

When real energy prices are fixed, the result implies the rebound effect is less than 1 if the cost-share of energy is less than $(1-\sigma)(1-\alpha)^{\frac{1-\sigma}{\sigma}}$.

3.2.2. *Equilibrium Innovation*

Since consumers are risk neutral, the interest rate in the economy will be ρ. Let $v_j(i, t \mid \lambda q)$ denote the value of a successful innovation of machine i in sector j with quality q at time t. We have

$$v_j(i, t \mid \lambda q) = \mathbb{E}_t \int_{s=t}^{T} e^{-\rho s} \pi(i, s \mid q) \, \mathrm{d}s \tag{26}$$

where T is the random stopping time after which a new innovation replaces the incumbent. Free entry and exit implies

$$\eta_j \frac{v_j(i, t \mid \lambda q)}{q} = 1, \qquad j \in \{E, L\} \tag{27}$$

holds for all t and $i \in [0, 1]$. Assuming the value function is differentiable implies $\dot{v}_j(i, t \mid q) = 0$. The Hamilton Jacobi Bellman (HJB) equation for $v_j(t, i \mid q)$ is

$$\dot{v}_j(i, t \mid q) = \big(\rho + \eta_j b_j(i, t)\big) v_j(i, t \mid q) - \pi_j(i, t \mid q) \tag{28}$$

See details in Acemoglu (2009) Equation (14.13), Acemoglu and Cao (2015), Equation (14). Immediately giving, using (27) and (16),

$$b_E(t) + \rho = \eta_E \lambda \alpha p_E(t)^{\frac{1}{\alpha}} E(t), \qquad b_L(t) + \rho = \eta_L \lambda \alpha p_L(t)^{\frac{1}{\alpha}} L \tag{29}$$

Note the above equation implies $b_j(t)$ is no longer conditioned on the machine variety — the rate of innovations for all machines in a sector is equal. Because scale effects are cancelled out in the innovation function, the relative rate of innovation in both sectors is unaffected by path dependence, and

$$\frac{\rho+\eta_E b_E(t)}{\rho+\eta_L b_L(t)} = \frac{\eta_E}{\eta_L}\overbrace{\left(\frac{p_E(t)}{p_L(t)}\right)^{\frac{1}{\alpha}}}^{\text{Price effect}}\underbrace{\frac{E(t)}{L}}_{\text{Market size effect}} = \frac{\eta_E}{\eta_L}\left(\frac{Q_E(t)}{Q_L(t)}\right)^{-\frac{1}{\sigma}}\left(\frac{E(t)}{L(t)}\right)^{\frac{\sigma-1}{\sigma}} \tag{30}$$

As energy use grows, the response of the relative rates of innovation can be split into first order effects, how relative rates of innovation respond keeping $\frac{Q_E(t)}{Q_L(t)}$ fixed and 'second order effects', how relative rates of innovation respond after $\frac{Q_E(t)}{Q_L(t)}$ has adjusted in response to the first order effects. Now, keep the ratio $\frac{Q_E(t)}{Q_L(t)}$ fixed and consider the 'first order effects' of expanding energy use on relative research profitability. The market size effect increases research profitability for energy augmenting research, while an increase in energy use reduces the relative price of the energy intermediate, leading to a fall in energy research profitability. If $\sigma < 1$, then the price effect is stronger and the first order effects lead to a fall of profitability to energy research. Turning to the second order effects, to maintain the free entry/exit conditions, b_E must fall, and then $\frac{Q_E(t)}{Q_L(t)}$ falls, leading to a second order price effect running in the opposite direction to the first order price effect (recall again equation (16)). In the following sub-section, we show the second order effects from the falling $\frac{Q_E(t)}{Q_L(t)}$ ratio is enough to keep relative incentives to do research in both sectors in balance so innovation continues in both sectors.

Turning to the regime with path dependence, let

$$V_j(t) :\, = \int_0^1 v_j(i,t\,|\,q)\,\mathrm{d}i\,, \qquad j \in \{E, L\} \tag{31}$$

and let

$$\Pi_j(t) = \int_0^1 \pi_j(i,t\,|\,q)\,\mathrm{d}i\,, \qquad j \in \{E, L\} \tag{32}$$

Free entry and exit implies

$$V_E(t) \geqslant V_L(t), \qquad \text{if} \quad s_E > 0 \tag{33}$$

$$V_L(t) \geqslant V_E(t), \qquad \text{if} \quad s_L > 0 \tag{34}$$

The relative rates of innovation in the case with path dependence are more difficult to characterize using current profit ratios. To gain some intuition before moving to the formal result, suppose the relative rates of innovation depended only on the ratio of current expected profits

$$\begin{aligned} \frac{\rho + \eta_E s_E(t)}{\rho + \eta_L s_L(t)} &= \frac{\eta_E \Pi_E(t)}{\eta_L \Pi_L(t)} \\ &= \frac{\eta_E}{\eta_L} \overbrace{\frac{p_E(t)}{p_L(t)}^{\frac{1}{\alpha}}}^{\text{Price effect}} \underbrace{\frac{E(t)}{L}}_{\text{Market size effect}} \overbrace{\frac{Q_E(t)}{Q_L(t)}}^{\text{Scale effect}} = \frac{\eta_E}{\eta_L} \left(\frac{Q_E(t)}{Q_L(t)} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{E(t)}{L(t)} \right)^{\frac{\sigma-1}{\sigma}} \end{aligned} \tag{35}$$

As energy use grows, once again, because of the price effect, the first order effects lead to a fall in relative profitability to do energy research. However, there are now two second order effects; on the one hand, the subsequent fall in $\frac{Q_E(t)}{Q_L(t)}$ from higher profitability to do labor research pushes up the profitability for energy research through the second order price effect. On the other hand, as $\frac{Q_E(t)}{Q_L(t)}$ falls, the scale effect increases profitability to do labor research. The second order scale effect runs in *the same direction as the first order price effect*, pushing down relative profitability for energy research. Whether the relative profits for energy and labor augmenting research remain in a constant ratio or whether relative profits for energy keep falling depends on how fast energy use grows relative to technology. From (22), consider

$$\left(\frac{Q_E(t) E(t)}{Q_L(t) L} \right)^{\frac{\sigma-1}{\sigma}} = \frac{\theta_E(t)}{1 - \theta_E(t)} \gamma^{\frac{-\epsilon}{\sigma}} \tag{36}$$

Now use (21) and (23), to arrive at

$$\theta_E(t) = (\alpha(1-\alpha))^{\sigma-1} \kappa(t)^{1-\sigma} \gamma_E^{\frac{\epsilon}{\sigma-1}} Q_E(t)^{\sigma-1} \tag{37}$$

Thus, along a growth path, unlike in the regime without scale effects, if there is energy augmenting innovation, then the relative profitability of energy augmenting research must continue to fall. In the following sub-section, we show the continued fall in profitability means labor augmenting innovation must eventually crowd out all energy augmenting innovation.

3.3. *Equilibrium Growth Paths*

An equilibrium growth path is a balanced growth path if output and consumption both grow at a constant rate g, energy grows at a constant rate g_E, and Q_E and Q_L grow at constant rates g_{Q_E} and g_{Q_L} respectively. For any path $X(t)$, we will use $\hat{X}(t)$ to refer to the growth rate of $X(t)$ at time t, that is $\hat{X}(t) = \frac{\dot{X}(t)}{X(t)}$, and we use $X(t) \to Y(t)$ to mean $\lim_{t\to\infty} X(t) = Y(t)$ for a path $Y(t)$. [12]

PROPOSITION 3.2 *If along an equilibrium growth path, $\hat{Q}_E(t) > c$ for all t, where $c > 0$ and $\dot{\kappa}(t) = 0$, then $\hat{Y}(t)$, $\hat{E}(t)$, $\hat{Q}_L(t)$ and $\hat{Q}_E(t)$ cannot be constant.*

PROOF: By (20), $Y(t) = \tilde{F}(Q_E(t)E(t), Q_L(t)L)$ where $\tilde{F}$ is a homogeneous of degree one production function. Using Euler's Theorem, $g_{Q_E} + g_E = g_{Q_L}$ must hold for g, g_E, g_{Q_E} and g_{Q_L} to be constant (See Claim A.1 in the Appendix). However, $g_{Q_E} + g_E = g_{Q_L}$ yields a contradiction, since Equation (22) implies $\theta_E(t)$ is constant, which in turn implies $\hat{Q}_E(t) = 0$ and $E(t)/Y(t)$ is constant by (23) and (37).

Q.E.D.

Thus a BGP cannot feature non-price energy efficiency improvements. We study the possibility of non-price energy intensity improvements along *asymptotic* balanced growth paths (ABGP).[13]

DEFINITION 3.2 An equilibrium growth path is an asymptotic balanced growth path if $\hat{Y}(t) \to g$, $\hat{C}(t) \to g$, $\hat{E}(t) \to g_E$, $\hat{Q}_E(t) \to g_{Q_E}$, $\hat{Q}_L(t) \to g_{Q_L}$, $b_E(t) \to b_E^\star$, $b_L(t) \to b_L^\star$, $s_E(t) \to s_E^\star$ and $s_L(t) \to s_L^\star$, where all limits are real valued.

We now show an ABGP for the model with no path dependence can feature non-price energy efficiency improvements.

THEOREM 3.1 *If an equilibrium is an ABGP with no path dependence and $\dot{\kappa}(t) = 0$, then*

1. $\hat{Y}(t) \to \hat{Q}_L(t) \to g$ *with* $g > 0$
2. $\theta_E(t) \to 0$

[12] $X(t) \to Y(t)$ if and only if for every $\epsilon > 0$, there exists T such that for all $t > T$, $|X(t) - Y(t)| < \epsilon$.

[13] The definition of BGP and ABGP varies. Hassler et al. (2016) define a BGP in the same way as us. However,Acemoglu (2003) uses the term BGP to refer to what we define at ABGP. The distinction matters in our model since BGPs cannot feature non-price efficiency improvements while ABGPs can.

3. $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$
4. $\hat{E}(t) - \hat{Y}(t) \to \hat{Y}(t)\frac{(1-\theta_E(t))(\sigma-1)}{2-\theta_E(t)-\sigma}$

The proof of the theorem is in the Appendix. Part 1. of the above theorem says the growth of output converges to the growth rate of labor augmenting technologies, this happens because the cost share of labor converges close to 100% of output (part 2.). Part 3. tells us that in the limit, the growth rate of energy augmenting technologies converges to a constant rate that is slower than the rate of labor augmenting technologies, since we have assumed $\sigma < 1$. Part 4. tells us that energy efficiency declines, but at a slower rate than output growth.

By contrast, under path dependence, sustained non-price energy intensity improvements are not possible.

THEOREM 3.2 *If an equilibrium path with path dependence is an ABGP and $\dot{\kappa}(t) = 0$, then $\hat{Y}(t) \to \hat{Q}_L(t) \to g$ where $g = \eta_L$, $\theta(t) \to \theta_E^\star$ where $\theta_E^\star > 0$ and $\hat{E}(t) - \hat{Y}(t) \to 0$.*

THEOREM 3.3 *In a path dependence regime, every equilibrium $\mathscr{E}$ is an ABGP. Moreover, if $\eta_L < \rho$, then there exists a unique equilibrium.*

Moreover, if the energy extraction costs are low enough or energy augmenting technologies are advanced enough, then there exists a BGP in the model with path dependence with no energy augmenting research. For the next result, let $\kappa(t) = \kappa$ for all t,

PROPOSITION 3.3 *In the path dependence regime, there exists a balanced growth path with $\eta_L s_L(t) = \eta_L$ and $s_E(t) = 0$ for all t if and only if*

$$\left(\frac{\kappa}{Q_E(0)}\right)^{1-\sigma} \leqslant \frac{(\alpha(1-\alpha))^{1-\sigma}\gamma^{\frac{\epsilon}{1-\sigma}}}{\gamma^{\frac{2\epsilon}{\sigma-1}}\frac{\rho+\eta_L}{\rho-\eta_L}+1}, \qquad t \geqslant 0 \tag{38}$$

Combining (36) with (38) implies

$$\frac{E(0)}{Y(0)} \leqslant \frac{\alpha\kappa^{-1}}{\gamma^{\frac{2\epsilon}{\sigma-1}}\frac{\rho+\eta_L}{\rho-\eta_L}+1} := \phi \tag{39}$$

must hold at time 0 for for a BGP to exist. Noting (36), if $Q_E(0)$ is constant along the BGP, then $\theta_E(t)$ remains constant, and thus, by (21), $\frac{E(t)}{Y(t)}$ remains constant and satisfies (39) along the BGP.

3.3.1. *Convergence*

Figure 3.3.1 shows ABGPs with constant prices (in bold) for the regime with no path dependence.

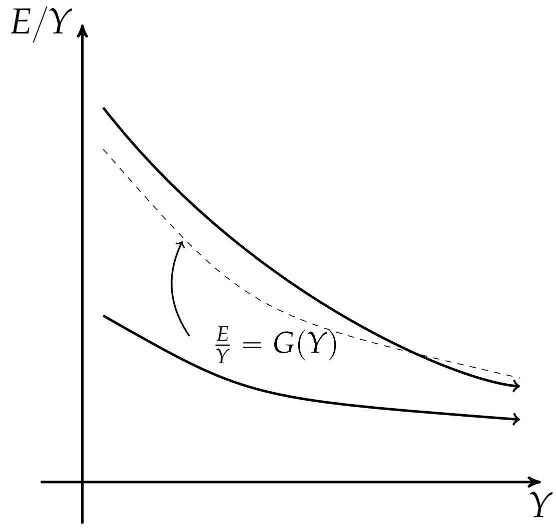

FIGURE 5.— ABGPs for a regime with no path dependence (in bold). All ABGPs feature declining energy intensity, but must converge to the dotted curve.

To verify the direction of the ABGPs, note output is growing and $b_E(t) > 0$, hence energy intensity must be falling. We verified in Theorem 3.1 that the elasticity of E/Y to Y must converge to $\frac{\sigma-1}{2-\sigma}$. We can also verify energy intensity for a given level of output converges for all ABGPs (conditional convergence):

PROPOSITION 3.4 *Any equilibrium with path dependence is an ABGP with*

$$\frac{E(t)}{Y(t)} \to G(Y(t))$$

where $G\colon \mathbb{R}_{++} \to \mathbb{R}_{++}$ *is a monotone decreasing function satisfying* $G(x) \to 0$.

For conditional convergence of energy to take place, ABGPs starting with higher energy intensity must experience a faster decline in energy intensity relative to output. For the next result, let $\{E, Y, Q_E, Q_L, \theta_E\}$ and $\{\tilde{E}, \tilde{Y}, \tilde{Q}_E, \tilde{Q}_L, \tilde{\theta}_E\}$ be two equilibrium paths with no path dependence.

CLAIM 3.1 If $\tilde{Y}(t) = Y(t)$ and $\tilde{E}(t) > E(t)$, then

1. $\frac{\hat{\tilde{Q}}_E(t)}{\hat{\tilde{Q}}_L(t)} > \frac{\hat{Q}_E(t)}{\hat{Q}_L(t)}$
2. $\hat{\tilde{Q}}_E(t) > \hat{Q}_E(t)$
3. $\frac{\hat{\tilde{E}}(t) - \hat{\tilde{Y}}(t)}{\hat{\tilde{Y}}(t)} < \frac{\hat{E}(t) - \hat{Y}(t)}{\hat{Y}(t)}$

Intuitively, energy intensity falls faster in economies with higher energy intensity because higher energy intensity raises the profitability to energy augmenting research through both the market size *and* price effects. Recall equation (21),

$$b_E(t) = \alpha p_E(t)^{\frac{1}{\alpha}} E(t) - \rho$$

For a given level of output, $E(t)$ is higher if energy intensity is higher. Moreover, by (48) in the Appendix, $p_E(t)$ is also higher when energy intensity is higher.

Next, consider figure 6, which shows ABGPs for a model with path dependence and a constant real energy price. The dotted line is the BGP where $\frac{E}{Y} = \phi$, where ϕ is implied by Equation (39). By Proposition 3.3, any economy with energy intensity below ϕ has a unique BGP with no energy augmenting technical change. However, again by proposition 3.3, if energy intensity is strictly greater than ϕ, there must be some energy augmenting technical change. And by Theorem 3.2, Q_E and hence energy intensity must converge to a constant.

4. DISCUSSION: NO PATH DEPENDENCE OR PATH DEPENDENCE?

We use Equation (23) to estimate $Q_E(t)$ using U.S. data between 1900 and 2015 and plot the estimate in figure 7. Considering Equation (13) for the wage rate, we also use the following equation

$$\gamma_L^{\frac{\epsilon}{\sigma(\sigma-1)}} \alpha (1-\alpha)^{\frac{\sigma-1}{\sigma}} \left(\frac{Y(t)}{L(t)} \right)^{\frac{1}{\sigma}} Q_L(t)^{\frac{\sigma-1}{\sigma}} = w_L(t)$$

to estimate $Q_L(t)$.[14] We then plot $\frac{Q_E(t)}{Q_L(t)}$ in figure 8.[15]

[14]We use annual wage per employee from the U.S. Employment Statistics and use total employment as $L(t)$.

[15]Note the parameter values in front of equation (23) only shifts estimates of $Q_E(t)$ without effecting its growth rate. As such, we set the parameter values to 1 for this exercise.

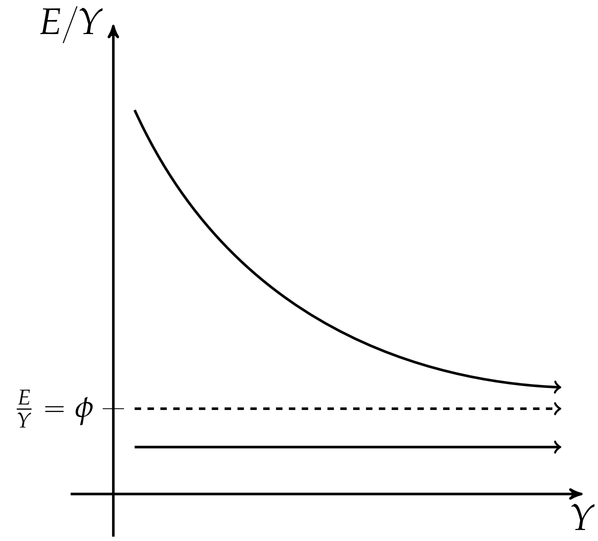

FIGURE 6.— ABGPs for a regime with path dependence. All ABGPs must converge to a BGP where energy intensity is constant.

Before 1950, the growth rate of of $Q_E(t)$ was positive, while prices declined, suggesting AEEI. During the same period, the bias of technical change, $\frac{Q_E(t)}{Q_L(t)}$ favoured energy augmenting technical change; energy augmenting technologies grew faster than labor augmenting technologies. In view of our analysis in section 3.3.1, the bias of research toward energy augmenting technical change suggests an initial condition for the U.S. where energy efficiency is sufficiently backward and $\frac{E(t)}{Y(t)}$ is high enough for $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} > 1$ (See 1. of Claim 3.1). After 1960, and before 1980, the bias of research favoured labor augmenting technologies, consistent with the long-run behaviour of $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)}$ predicted by the theory in this paper. However, the growth rate of $Q_E(t)$ declines to under zero by the early 1970s, before rising and falling again in a lagged response to prices.

Both the no path dependence and path dependence regime can be consistent with the path of $Q_E(t)$. If we suppose the U.S. economy is sufficiently energy inefficient in the early 1900s, then AEEI would have taken place under a path dependence regime (recall Proposition 3.3). The growth of $Q_E(t)$ does decline below zero just before the price shocks of the 1970s and the early 2000s, however, this does not imply a model with no non-price energy efficiency improvements along an ABGP. Both path dependence and non-path dependence regimes predict such a decline;

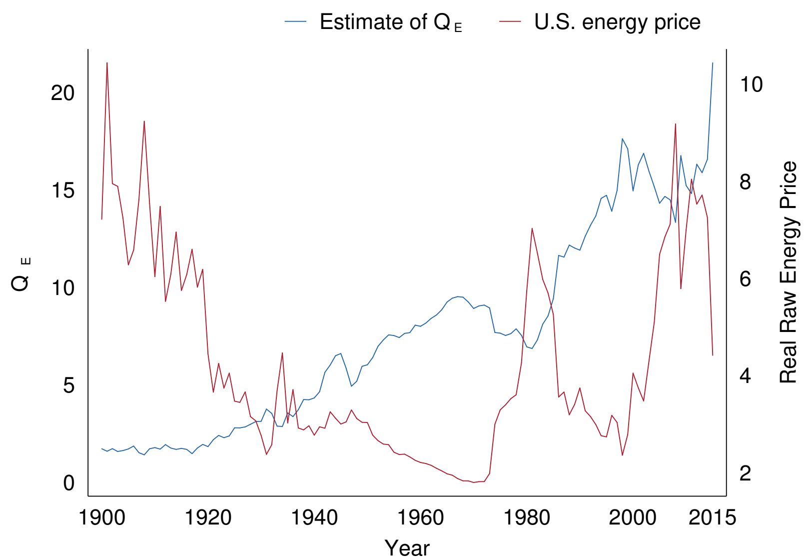

FIGURE 7.— Estimate of $Q_E(t)$ and the U.S. energy price, estimated assuming $\sigma = .3$

the decline $Q_E(t)$ is associated with a fall in energy prices. For a given level of $Q_E(t)$ and $Q_L(t)$, when prices fall, energy consumption increases, and recalling (30) and (35), energy augmenting technical change becomes relatively less profitable due to the stronger price effect.

Both the no path dependence and path dependence regime can also be consistent with the paths of energy intensity across countries shown by figures 2 and B.2; the figure shows many countries with low energy intensities that have close to constant paths of energy intensity. Under no path dependence, suppose that $b_E(t) + b_L(t) \leqslant \bar{B}(t)$ must hold in equilibrium for some value $\bar{B}(t)$. If $\bar{B}(t)$ were fixed or determined through the savings of consumers, then if $\bar{B}(t)$ is too low, the no arbitrage condition for energy may bind and only labor augmenting research will take place until either $\bar{B}(t)$ rises or the profitability for labor augmenting research falls sufficiently. Thus low growth countries with low energy intensities may experience no decline in energy intensity. On the other hand, under no path dependence, the countries showing constant energy intensity through time may be on their long-run BGP, to which all countries will converge to.

Perhaps one piece of evidence in favour of the path dependence scenario is the

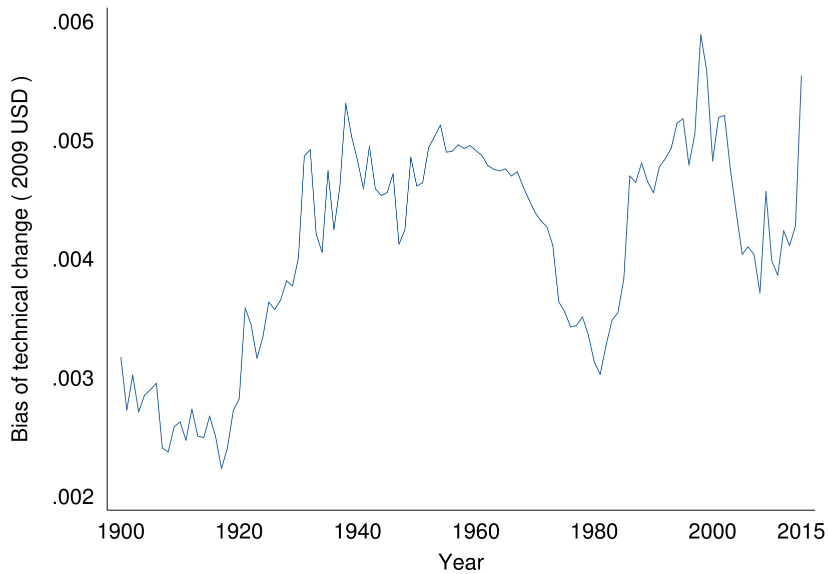

FIGURE 8.— The direction of technical change, estimated values of $\frac{Q_E(t)}{Q_L(t)}$ for the U.S., assuming $\sigma = .3$

cross-sectional elasticity of energy intensity to output, which falls close to between -20% to -6% (see table B.2 in the Appendix and Discussion). Consider the elasticity of energy intensity to output towards which ABGPs converge to, given by Theorem 3.1,

$$\hat{E} - \hat{Y} = \hat{Y}\frac{(\sigma - 1)(1 - \theta_E(t))}{2 - \theta_E(t) - \sigma}$$

With constant prices, the high elasticities seen in the cross-section of energy intensity to output can only be predicted along the ABGP in the no path dependence model if σ is higher than around .7 (see figure 9). However, such elasiticties of substitution are higher than the estimates in the literature, which range from close to zero (Hassler et al., 2016) to .5 (Stern et al., 2016). An elasticity of substitution higher than .7 is also not consistent with a smooth evolution of $Q_E(t)$. A high σ predicts a highly volatile estimate of $Q_E(t)$, rising by 350% in just one year (see section 4 by Hassler et al. (2016)); it is difficult to attribute $Q_E(t)$ in this case to technical change.

Finally, note the shape of the relationship between $\theta_E(t)$ and the elasticity of energy

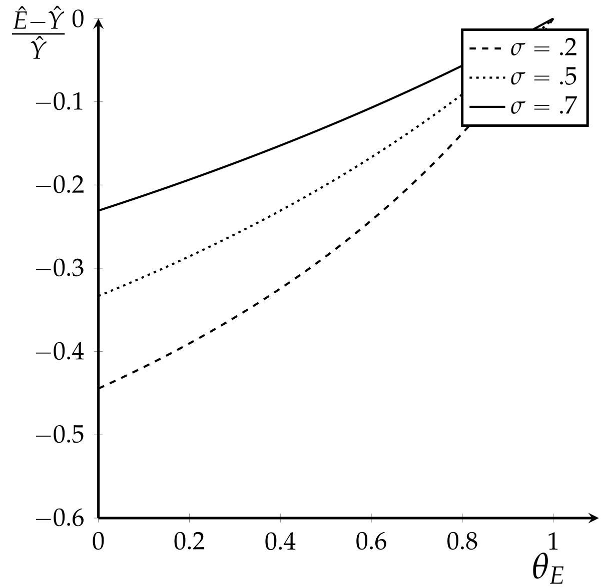

FIGURE 9.— Elasticity of Energy Intensity to Output along the 'stable path' to which all ABGPs in the model with no path dependence to.

intensity to out in figure 9: the elasticity is *lower* for lower levels of energy intensity. By contrast, our discussion of the in Section 2, in particular figures 2 and B.2 suggested the elasticity is lower in countries with higher energy intensity. Thus, for the no path dependence model to be consistent with the data, the cross section of countries cannot follow the dotted line 3.3.1; we require sufficiently many countries to be north of the dotted line in 3.3.1.

5. CONCLUSION

We developed an endogenous model of energy with directed technical change to understand the economic drivers and future potential for non-price energy efficiency improvements. Our analysis focused on understanding the trade-offs firms face between undertaking energy augmenting and non-energy augmenting research. Our key finding was that if energy prices do not continue to rise, then there is a relatively stronger incentive to undertake labor augmenting research, since the price of labor rises relative to the price of energy. Whether or not labor augmenting research completely crowds out energy augmenting research depends on whether or not innovation possibilities depend on scale effects; with scale effects, the relatively stronger incentive for labor augmenting research is compounded and autonomous energy efficiency improvements are eventually crowded out. On the other hand, in a model without scale effects, autonomous energy efficiency continue at a constant rate relative to output, albeit at a slower rate than labor augmenting technical change.

Despite the sharp theoretical predictions of the directed technical change model, we were unable to identify whether a model with scale effects or no scale effects more accurately accounts for past trends. This is because under both models, one, if energy intensity is high enough, autonomous energy efficiency improvements take place, and two, because fluctuating prices have led to notable fluctuations in the estimate of energy augmenting technology, which prevents us from determining whether energy efficiency is converging to a constant or not. Some mild evidence in favour suggestions the scale effects model may be more appropriate is the fact that the fall in energy intensity relative to output in the cross-section of high income countries is small, suggesting energy intensities in the cross-section converge to a constant. However, we also point out that the scale effects specification has received much debate in the endogenous growth literature since, with population growth, it implies increasing growth rates (see Jones (1995)).

APPENDIX A: PROOFS

A.1. *Characterizing Growth Paths*

For the following claim, let $\tilde{F}\colon \mathbb{R}^2_+ \to \mathbb{R}$ be homogeneous of degree one, increasing in both arguments and differentiable, let $Y(t) = \tilde{F}(X(t), Z(t))$ and let $\hat{Y}(t)$, $\hat{X}(t)$ and $\hat{Z}(t)$ denote the growth rates of $Y(t)$, $X(t)$ and $Z(t)$ respectively.

The following claim is used in the proof of Proposition 3.2.

CLAIM A.1 If $\hat{Y}$, $\hat{X}$ and $\hat{Z}$ are constant, then $\hat{X} = \hat{Z}$.

PROOF: Let g_Y, g_X and g_Z be the constant growth rates of $Y(t)$, $X(t)$ and $Z(t)$. Taking time derivatives of $Y(t)$ gives

$$\dot{Y}(t) = \tilde{F}_1(X(t), Z(t))\dot{X}(t) + \tilde{F}_2(X(t), Z(t))\dot{Z}(t)$$

and thus

$$\hat{Y}(t) = \frac{\tilde{F}_1(X(t), Z(t))X(t)}{\tilde{F}(X(t), Z(t))}\hat{X}(t) + \frac{\tilde{F}_2(X(t), Z(t))Z(t)}{\tilde{F}(X(t), Z(t))}\hat{Z}(t) :\ = (1 - \tilde{\theta}(t))\hat{X}(t) + \tilde{\theta}(t)\hat{Z}(t)$$

where, by Euler's Theorem (theorem 2.1 in Acemoglu (2009)),

$$(1 - \tilde{\theta}(t)) + \tilde{\theta}(t) = \frac{\tilde{F}_1\left(X(t), Z(t)\right)X(t) + \tilde{F}_2\left(X(t), Z(t)\right)Z(t)}{\tilde{F}\left(X(t), Z(t)\right)} = 1$$

Now suppose by contradiction that $g_X \neq g_Z$. First consider the case $g_X < g_Z$. We must have

$$g_Y = (1 - \tilde{\theta}(t))g_X + \tilde{\theta}(t)g_Z < (1 - \tilde{\theta}(t))g_Z + \tilde{\theta}(t)g_Z = g_Z \tag{40}$$

Since g_Y, g_X and g_Z are the growth rates of $Y(t)$, $X(t)$ and $Z(t)$,

$$Y(0)e^{tg_Y} = \tilde{F}\left(X(0)e^{tg_X}, Z(0)e^{tg_Z}\right)$$

dividing the LHS and RHS by e^{tg_Y} and by homogeneity of the function $\tilde{F}$, we have

$$\begin{aligned} Y(0) &= e^{t(g_Z - g_Y)}\tilde{F}\left(X(0)e^{t(g_X - g_Z)}, Z(0)\right) \\ &< \tilde{F}\left(X(0), Z(0)\right) = Y(0) \end{aligned}$$

where the inequality follows from our assumption $g_X < g_Z$, Equation (40) which says $g_Y < g_Z$, the fact that $\tilde{F}$ is increasing and observing $e^x < 1$ for any $x < 0$. However, $Y(0) < Y(0)$ yields a contradiction, implying $g_X \geqslant g_Z$. The case to rule out $g_X > g_Z$ is symmetric, establishing $g_X \ngtr g_Z$.

Q.E.D.

PROOF OF THEOREM 3.1: We first derive growth equations that hold in equilibrium. Take time derivatives of $\theta_E(t)$, defined by (22), to write

$$\hat{\theta}_E(t) = \frac{\sigma - 1}{\sigma}(1 - \theta(t))\left(\hat{Q}_E(t) + \hat{E}(t) - \hat{Q}_L(t)\right) \tag{41}$$

Similarly, take time derivatives of $Y(t)$, defined by (20), to write

$$\hat{Y}(t) = (1 - \theta_E(t))\,\hat{Q}_L(t) + \theta_E(t)\left(\hat{Q}_E(t) + \hat{E}(t)\right) \tag{42}$$

Now, note from (21), we have $\hat{\theta} = \hat{E} - \hat{Y}$. Use this expression in Equation (41) to write

$$\hat{E}(t) = \frac{\sigma - 1}{\sigma}(1 - \theta_E(t))(\hat{B}(t) + \hat{E}(t)) + \hat{Y}(t) \tag{43}$$

Next, combine (43) with (42) to write

$$\hat{E}(t) = \frac{\sigma}{1 - \theta_E(t)}\hat{Q}_E(t) - \hat{B}(t) \tag{44}$$

Let $\Gamma(t)$ define the ratio of relative profitability between energy and labor augmenting research as follows:

$$\Gamma(t) := \frac{\rho + \eta_E b_E(t)}{\rho + \eta_L b_L(t)}$$

Recall along an equilibrium innovation in both sectors, no arbitrage requires $\Gamma(t) = \left(\frac{Q_E(t)}{Q_L(t)}\right)^{-\frac{1}{\sigma}}\left(\frac{E(t)}{L(t)}\right)^{\frac{\sigma-1}{\sigma}}$. And thus,

$$\hat{\Gamma}(t) = \frac{1 - \sigma}{\sigma}\hat{E}(t) - \hat{B}(t)\left(\frac{1}{\sigma}\right) \tag{45}$$

Combine (44) with (45),

$$\begin{aligned}\hat{\Gamma}(t) &= \frac{\sigma-1}{\sigma}\hat{E} - \frac{\hat{B}(t)}{\sigma}\\ &= \frac{\sigma-1}{\sigma}\left(\frac{\sigma}{1-\theta(t)}\hat{Q}_E(t) - \hat{B}(t)\right) - \frac{\hat{B}(t)}{\sigma}\\ &= \frac{\sigma-1}{1-\theta_E(t)}\hat{Q}_E(t) - \hat{B}(t)\end{aligned}$$

Along an ABGP, $\hat{\Gamma}(t) \to 0$, and using the above, recalling $\hat{B}(t) = \hat{Q}_E(t) - \hat{Q}_L(t)$, deduce

$$\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{1-\theta(t)}{2-\theta(t)-\sigma} \tag{46}$$

Since $\hat{Q}_L(t)$ is positive and bounded below by a positive constant, we must have $\hat{Q}_E(t)$ is positive and bounded below by a positive constant. Accordingly, $\theta_E(t) \to 0$ by (37). By (42), $\hat{Y}(t) \to \hat{Q}_L(t)$. Now, noting (23), taking time derivatives, we arrive at

$$\hat{E}(t) - \hat{Y}(t) = (\sigma-1)\,\hat{Q}_E(t) \to \hat{Y}(t)\frac{(1-\theta(t))\,(\sigma-1)}{2-\theta(t)-\sigma} \to \hat{Y}(t)\frac{\sigma-1}{2-\sigma}$$

Q.E.D.

We now prepare some preliminary notation before turning to the proof of theorem 3.2. Noting $p_L(t) = \gamma_L\left(\frac{Y(t)}{Y_L(t)}\right)^{\frac{1}{\epsilon}}$, we have

$$\begin{aligned}p_L(t) = \gamma_L\left(\frac{Y(t)}{Y_L(t)}\right)^{\frac{1}{\epsilon}} &= \gamma_L\left(\gamma_L + \gamma_E\left(\frac{Y_E}{Y_L}\right)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{1}{\epsilon-1}}\\ &= \gamma_L^{\frac{\epsilon}{\epsilon-1}}\left(1 + \gamma^{\frac{\epsilon}{\sigma}}\left(\frac{Q_E(t)E(t)}{Q_L(t)L}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\epsilon-1}}\end{aligned}$$

where the third equality uses (19). Using (22), we can then write

$$\alpha p_L(t)^{\frac{1}{\alpha}} L = \alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}\left(1-\theta_E(t)\right)^{\frac{1}{1-\sigma}} L \tag{47}$$

similarly

$$\alpha p_E(t)^{\frac{1}{\alpha}} E(t) = \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{1}{1-\sigma}} E(t) \tag{48}$$

Now define

$$J_L(t) := \mathbb{E}_t \int_t^{T_L(t)} e^{-\rho s} \alpha p_L(s)^{\frac{1}{\alpha}} L \, \mathrm{d}s, \quad J_E(t) := \mathbb{E}_t \int_t^{T_E(t)} e^{-\rho s} \alpha p_E(s)^{\frac{1}{\alpha}} E(s) \, \mathrm{d}s$$

where $T_j(t)$ is the random stopping time after which an incumbent is replaced by a new entrant. Note the distribution of $T_j(t)$ does not depend on the individual machine i, since each machine experiences a common innovation, hence replacement rate $s_j \eta_j$. Recalling the definition of $V_L(t)$ from (31),

$$\begin{aligned} V_L(t) = \int_0^1 v_L\left(i, t \,|\, q\right) \mathrm{d}i &= \int_0^1 \mathbb{E}_t \int_{s=t}^{T_L(t)} e^{-\rho s} \pi_L(s) \, \mathrm{d}s \, \mathrm{d}i \\ &= \int_0^1 q(i,t) \mathbb{E}_t \int_{s=t}^{T_L(t)} e^{-\rho s} \alpha p_L(s)^{\frac{1}{\alpha}} L \, \mathrm{d}s \, \mathrm{d}i \\ &= Q_L(t) \mathbb{E}_t \int_{s=t}^{T_L(t)} e^{-\rho s} \alpha p_L(s)^{\frac{1}{\alpha}} L \, \mathrm{d}s \\ &= Q_L(t) J_L(t) \end{aligned}$$

The HJB equation for $J_L(t)$ will be

$$\dot{J}_L(t) = \left(\rho + \eta_L s_L(t)\right) J_L(t) - \alpha p_L(t)^{\frac{1}{\alpha}} L \tag{49}$$

and similarly, the HJB equation for $J_E(t)$ becomes

$$\dot{J}_E(t) = \left(\rho + \eta_E s_E(t)\right) J_E(t) - \alpha p_E(t)^{\frac{1}{\alpha}} E(t) \tag{50}$$

LEMMA A.1 *If an equilibrium with path dependence is an asymptotic balanced growth path with $s_E^\star > 0$ and $s_L^\star > 0$, then $\eta_L s_L^\star = \rho + 2\eta_E s_E^\star$.*

PROOF: By equation (37), since $Q_E(t) \to \infty$, $\theta_E(t) \to 0$. Now, if $\theta(t)_E \to 0$, using (47), we have

$$\alpha p_L(t)^{\frac{1}{\alpha}} L \to \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} L \tag{51}$$

which implies $J_L(t)$ converges to a constant. Note both the terms on the RHS of the HJB equation for $J_L(t)$, Equation (49), converge, implying $\dot{J}_L(t)$ converges. But since $J_L(t)$ converges to a constant, we must have $\dot{J}_L(t) \to 0$. Once again by the HJB condition, Equation (49),

$$J_L(t) \to \frac{\alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} L}{\rho + s_E^\star} \tag{52}$$

Also note

$$\hat{J}_L(t) = \frac{\dot{J}_L(t)}{J_L(t)} \tag{53}$$

The denominator converges a positive constant, while numerator converges to zero, implying $\hat{J}_L(t) \to 0$. Next, from the free entry and exit condition, since $s_E^\star > 0$ and $s_L^\star > 0$, there exists $\bar{T}$ such that

$$J_E(t) = \frac{Q_L(t) J_L(t)}{Q_E(t)}, \qquad t > \bar{T} \tag{54}$$

Taking time derivatives, we have

$$\hat{J}_L(t) - \hat{J}_E(t) = \hat{Q}_E(t) - \hat{Q}_L(t), \qquad t > \bar{T} \tag{55}$$

Rearranging gives

$$\hat{J}_E(t) = \hat{Q}_L(t) - \hat{Q}_E(t) + \hat{J}_L(t) \to \eta_L s_L^\star - \eta_E s_E^\star \tag{56}$$

where $\hat{J}_L(t) \to 0$ by the argument proceeding Equation (53). Now use the above with the HJB condition for $J_E(t)$, Equation (50),

$$\frac{\alpha p_E(t)^{\frac{1}{\alpha}} E(t)}{J_E(t)} = \rho + \eta_E s_E(t) - \hat{J}_E(t) \to \rho + 2\eta_E s_E^\star - \eta_L s_L^\star \tag{57}$$

Note the limit is a (possibly zero) constant. Moreover, by the HJB condition for $J_L(t)$, Equation (49),

$$\frac{\alpha p_L(t)^{\frac{1}{\alpha}} L}{J_L} = \rho + \eta_L s_L^\star - \hat{J}_L(t) \to \rho + \eta_L s_L^\star \tag{58}$$

since $\hat{J}_E(t) \to 0$. As such,

$$\begin{aligned}
\frac{\theta_E(t)}{1-\theta_E(t)} \gamma^{-\epsilon\sigma} &= \left(\frac{Q_E(t) E(t)}{Q_L(t) L} \right)^{\frac{\sigma-1}{\sigma}} \\
&= \frac{p_E^{\frac{1}{\alpha}} Q_E(t) E(t)}{p_L^{\frac{1}{\alpha}} Q_L(t) L} = \frac{p_E^{\frac{1}{\alpha}} J_L(t) E(t)}{p_L^{\frac{1}{\alpha}} J_E(t) L} \to \frac{\rho + 2\eta_E s_E^\star - \eta_L s_L^\star}{\rho + \eta_L s_L^\star}
\end{aligned}$$

where the first equality comes from (36), the second equality uses (18), the third uses the free entry and exit conditions (33) and convergence at the final step follows from (57) and (58) and noting the denominator converges to a strictly positive constant. Thus, if $\theta_E(t) \to 0$, then $\eta_L s_L^\star = \rho + 2\eta_E s_E^\star$.

Q.E.D.

CLAIM A.2 If an equilibrium with path dependence is an asymptotic balanced growth path with $s_E^\star > 0$ and $s_L^\star > 0$, then there exists $\bar{T}$ and M, with $M < \infty$ such that

$$\mathbb{E}_t \int_{s=t}^{T_E(t)} e^{\int_t^s \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}} \leqslant M < \infty, \qquad t > \bar{T}$$

PROOF: The distribution of $T_E(t)$ is

$$\mathbb{P}\left(T_E(t) \leqslant t + s\right) = 1 - e^{-\int_0^s \eta_E s_E(s_1 + t)\, \mathrm{d}s_1}$$

and the density of $T_E(t) - t$ is $\eta_E s_E(s) e^{-\int_0^s \eta_E s_E(s_1+t)\mathrm{d}s_1}$. Allowing us to write

$$\begin{aligned}\mathbb{E}_t \int_{s=t}^{T_E(t)} e^{\int_t^s \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s &= \mathbb{E}_t \int_0^{T_E(t)-t} e^{\int_0^s \hat{E}(\bar{s}+t) - \eta_E s_E(\bar{s}+t) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s \\ &= \eta_E s_E(s) \int_0^\infty \int_0^T e^{\int_0^s \hat{E}(\bar{s}+t) - \eta_E s_E(\bar{s}+t) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s \, e^{-\int_0^T \eta_E s_E(s_1+t)\mathrm{d}s_1} \, \mathrm{d}T\end{aligned}$$

Define

$$c := \lim_{t\to\infty} \{\hat{E}(t) - 2\eta_E s_E(t) - \rho\} \tag{59}$$

We now show $c < 0$. Taking growth rates across equation (23) gives $\hat{E}(t) - \hat{Y}(t) = (\sigma - 1)\hat{Q}_E(t)$, and thus $\hat{E}(t) - \hat{Y}(t) \to \eta_E s_E^\star(\sigma - 1)$, and $\hat{E}(t) \to \eta_E s_E^\star(\sigma - 1) + \eta_L s_L^\star$ since $\hat{Y}(t) \to \eta_L s_L^\star$ by (42). If $s_L^\star = 0$, then

$$c = \eta_E s_E^\star(\sigma - 1) + \eta_L s_L^\star - 2\eta_E s_E^\star - \rho = \eta_E s_E^\star(\sigma - 3) - \rho < 0 \tag{60}$$

On the other hand, if $s_L^\star > 0$, then by Lemma A.1, $\eta_L s_L^\star = \rho + 2\eta_E s_E^\star$

$$c = \eta_E s_E^\star(\sigma - 1) + \eta_L s_L^\star - \eta_E 2 s_E^\star - \rho = s_E^\star(\sigma - 1) < 0 \tag{61}$$

Next, note there exists $\epsilon > 0$ such that $c + \epsilon < 0$. Moreover, since $\hat{E}(t) - \eta_E s_E(t) - \rho \to \eta_E s_E^\star(\sigma - 2) + s_L^\star - \rho$, there exists $\bar{T}$ such that for all $t > \bar{T}$, we have

$$\hat{E}(t) - \eta_E s_E(t) - \rho < \eta_E s_E^\star(\sigma - 2) + \eta_L s_L^\star - \rho + \frac{\epsilon}{2} := c_1 \tag{62}$$

and

$$-\eta_E s_E(t) < -\eta_E s_E^\star + \frac{\epsilon}{2} := c_2 \tag{63}$$

The above two inequalities give

$$\begin{aligned} &\eta_E s_E(s) \int_0^\infty \int_0^T e^{\int_0^s \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s \, e^{-\int_0^T \eta_E s_E(s_1) \mathrm{d}s_1} \, \mathrm{d}T \\ &\leqslant \eta_E s_E(s) \int_0^\infty \int_0^T e^{\int_0^s c_1 \, \mathrm{d}\bar{s}} \, \mathrm{d}s \, e^{\int_0^T c_2 \mathrm{d}s_1} \, \mathrm{d}T = \int_0^\infty \frac{e^{T(c_1 + c_2)} - e^{Tc_2}}{c_1} \, \mathrm{d}T := M < \infty \end{aligned} \tag{64}$$

where the first inequality follows from monotonicity of the exponential function and noting (62) and (63). The second inequality comes from solving the inside integrals and the final inequality comes from noting $c_1 + c_2 = c + \epsilon < 0$.

Q.E.D.

PROOF OF THEOREM 3.2: If $s_E^\star > 0$, then there exists $\bar{T}$ such that for $t > \bar{T}$, the no arbitrage condition holds

$$\frac{J_E(t) Q_E(t)}{J_L(t) Q_L(t)} \geqslant 1, \qquad t > \bar{T} \tag{65}$$

to prove the theorem, we will show this condition cannot hold if $Q_E(t) \to \infty$ and $E(t)/Y(t) \to 0$. From the definition of $J_E(t)$ and $J_L(t)$, we have

$$\begin{aligned} \frac{J_E(t) Q_E(t)}{J_L(t) Q_L(t)} &= \frac{Q_E(t) \mathbb{E}_t \int_{s=t}^{T_E(t)} \pi_E(s)^{\frac{1}{\alpha}} E(s) e^{-\rho s} \, \mathrm{d}s}{Q_L(t) J_L(t)} \\ &= \kappa(t) (1 - \theta_E(t))^{1-\sigma} L \left(\frac{\theta_E(t)}{1 - \theta_E(t)} \right)^{\frac{1}{1-\sigma}} \\ &\quad \times \frac{E(t) Q_E(t)}{L Q_L(t)} \frac{\mathbb{E}_t \int_{s=t}^{T(t)} e^{\int_{\bar{s}=t}^s \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s}{J_L(t)} \\ &= \kappa(t) \left(1 - \theta_E(t)\right)^{1-\sigma} L \left(\frac{\theta_E(t)}{1 - \theta_E(t)} \right) \\ &\quad \times \frac{\mathbb{E}_t \int_{s=t}^{T(t)} e^{\int_{\bar{s}=t}^s \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s}{J_L(t)} \end{aligned} \tag{66}$$

where we have omitted constants in front of the RHS above for simplicity. The second equality above uses (48) and (37) to derive

$$\begin{aligned}\pi_E(s)^{\frac{1}{\alpha}}E(s) = \kappa(t)\theta_E(s)^{\frac{1}{1-\sigma}}E(s) &= Q_E(s)^{-1}E(s)\\ &= \kappa(t)Q_E(t)^{-1}E(t)e^{\int_{\bar{s}=t}^{s}\hat{E}(\bar{s})-\eta_E s_E(\bar{s})\,\mathrm{d}\bar{s}}\\ &= \kappa(t)\theta_E(t)^{\frac{1}{1-\sigma}}E(t)e^{\int_{\bar{s}=t}^{s}\hat{E}(\bar{s})-\eta_E s_E(\bar{s})\,\mathrm{d}\bar{s}}\end{aligned}$$

for $s > t$. Note once again, we have omitted constants in front of the RHS for simplicity. The third equality at (66) uses (36).

To complete the proof, by Claim A.2, $\mathbb{E}_t\int_{s=t}^{T}e^{\int_{\bar{s}=t}^{s}\hat{E}(\bar{s})-\eta_E s_E(\bar{s})-\rho\,\mathrm{d}\bar{s}}\,\mathrm{d}s < M$, where $M<\infty$. As such,

$$\frac{J_E(t)Q_E(t)}{J_L(t)Q_L(t)} \leqslant \kappa(t)\left(1-\theta_E(t)\right)^{\frac{1}{1-\sigma}}L\left(\frac{\theta_E(t)}{1-\theta_E(t)}\right)\frac{M}{J_L(t)}$$

Recall $\theta_E(t)\to 0$ by (37) if $Q_E(t)\to\infty$ and $J_L(t)$ converges to a constant, $\frac{J_E(t)Q_E(t)}{J_L(t)Q_L(t)}\to 0$. However, this contradicts (65).

Q.E.D.

PROOF OF PROPOSITION 3.3: We first prove the *if* statement of the proposition.

Part 1: If

Suppose (38) holds, we will show an allocation satisfying 1.- 4. of Definition 3.1 with $\eta_L s_L(t) = \eta_L$ for all t is an equilibrium (satisfies 6. of Definition 3.1) and is a BGP. To show the allocation satisfies 6. of Definition 3.1, it is sufficient to confirm the free entry and exit condition holds along the allocation path, that is,

$$\frac{J_L(t)Q_L(t)}{J_E(t)Q_E(t)} \geqslant 1, \qquad \forall\, t\geqslant 0 \tag{67}$$

Since $s_E = 0$, $Q_E(t) = Q_E(0)$ for all t and we must have, using Equation (37),

$$\theta_E(t) = \theta_E^\star := \left(\alpha(1-\alpha)\right)^{\sigma-1}\kappa(t)^{1-\sigma}\gamma_E^{\frac{\epsilon}{\sigma-1}}Q_E(0)^{\sigma-1}, \qquad \forall\, t\geqslant 0 \tag{68}$$

By (47), $p_L(t)$ remains constant. As such,

$$J_L(t) = \frac{\alpha p_L(t)^{\frac{1}{\alpha}}L}{\rho+\eta_L} = \frac{\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}\left(1-\theta_E^\star\right)^{\frac{1}{1-\sigma}}L}{\rho+\eta_L}, \qquad \forall\, t\geqslant 0 \tag{69}$$

Since the replacement rate is zero for energy augmenting innovations,

$$
\begin{aligned}
J_E(t) &= \int_t^\infty \alpha p_E(t)^{\frac{1}{\alpha}} E(t) e^{-\rho s}\, \mathrm{d}s \\
&= \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E^{\star\,\frac{1}{1-\sigma}} \int_0^\infty E(t) e^{-\rho s}\, \mathrm{d}s \\
&= \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E^{\star\,\frac{1}{1-\sigma}} E(t) \int_t^\infty e^{(\eta_L-\rho)s}\, \mathrm{d}s \\
&= \frac{\alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E^{\star\,\frac{1}{1-\sigma}} E(t)}{\rho - \eta_L}
\end{aligned}
$$

The second equality uses (48). For the third equality, note when $\hat{Q}_E(t) = 0$, $\hat{E}(t) = \eta_L = \hat{Y}(t)$ for all t by (23) and (42). The third equality follows from solving the integral.

We can now write

$$
\begin{aligned}
\frac{J_L(t)Q_L(t)}{J_E(t)Q_E(t)} &= \gamma^{\frac{-\epsilon}{\sigma-1}} \left(\frac{1-\theta_E^\star}{\theta_E^\star} \right)^{\frac{1}{1-\sigma}} \left(\frac{LQ_L(t)}{E(t)Q_E(t)} \right) \frac{\rho - \eta_L}{\rho + \eta_L} \\
&= \gamma^{\frac{-2\epsilon}{\sigma-1}} \left(\frac{1-\theta_E^\star}{\theta_E^\star} \right) \frac{\rho - \eta_L}{\rho + \eta_L} \\
&= \frac{J_L(0)Q_L(0)}{J_E(0)Q_E(0)}
\end{aligned}
\tag{70}
$$

for all $t \geqslant 0$. The second equality follows from (36). The third equality follows from our observation that $\theta_E(t)$ is constant for all $t \geqslant 0$, given by (68).

Using algebra, and (68), we can verify

$$
\gamma^{\frac{-2\epsilon}{\sigma-1}} \left(\frac{1-\theta_E^\star}{\theta_E^\star} \right) \frac{\rho - \eta_L}{\rho + \eta_L} \geqslant 1 \iff \left(\frac{\kappa}{Q_E(0)} \right)^{1-\sigma} \leqslant \frac{(\alpha(1-\alpha))^{1-\sigma} \gamma^{\frac{\epsilon}{1-\sigma}}}{\gamma^{\frac{2\epsilon}{\sigma-1}} \frac{\rho+\eta_L}{\rho-\eta_L} + 1}
\tag{71}
$$

And thus, if (38) holds, then $\frac{J_L(0)Q_L(0)}{J_E(0)Q_E(0)} \geqslant 1$, which by (70), implies (67). To confirm the equilibrium allocation is a BGP, note by (23) and (42), $\hat{Y}(t) = \eta_L$ and $\hat{E}(t) = \eta_L$ for all t. Moreover, $s_L(t) = 1$, $s_E(t) = 0$, $\hat{Q}_L(t) = \eta_L$ and $\hat{Q}_E(t) = 0$ for all t.

Now we turn to the *only if* statement of the proposition.

Part 2: Only If

Let an equilibrium allocation be a BGP with $\eta_L s_L(t) = \eta_L$. We show (38) must hold. Since $s_E(t) = 0$ for all t, (70) must hold along the growth path by the argument between proceeding Equation (67) above. Because the path is an equilibrium path, the free entry and exit conditions (67) must hold, and since (70) holds,

$$\gamma^{\frac{-2\epsilon}{\sigma-1}} \left(\frac{1-\theta_E^\star}{\theta_E^\star} \right) \frac{\rho - \eta_L}{\rho + \eta_L} \geqslant 1 \tag{72}$$

which, by (71), implies (38) holds. *Q.E.D.*

A.2. *Convergence of Growth Paths*

PROOF OF PROPOSITION 3.4: Along an asymptotic balanced growth path $z_E(t) \to z_E^\star$ and $z_L(t) \to z_E^\star$, where by Theorem 3.1, $z_E^\star > 0$ and $z_L^\star > 0$.

Recall

$$\begin{aligned} \eta_E \alpha p_E(t)^{\frac{1}{\alpha}} E(t) &= \eta_E \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{1}{1-\sigma}} E(t) \\ &= \eta_E \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{2-\sigma}{1-\sigma}} \frac{E(t)}{\theta_E(t)} \\ &= \eta_E \alpha^2 \kappa(t)^{-1} \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{2-\sigma}{1-\sigma}} Y(t) \end{aligned} \tag{73}$$

where the first equality is from Equation (48), the second equality follows from dividing through and multypying by $\theta_E(t)$ and the final equality follows Equation (21). Now, by Equation (46) in the proof of Theorem 3.1,

$$\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{1-\theta_E(t)-\sigma}{2-\sigma-\theta_E(t)} \tag{74}$$

But since $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} = \frac{\eta_E z_E(t)}{\eta_L z_L(t)}$, we must have

$$\begin{aligned} \frac{\eta_E z_E(t)}{\eta_L z_L(t)} &\to \frac{1-\theta_E(t)-\sigma}{2-\sigma-\theta_E(t)} \\ \implies \frac{\eta_E \alpha \kappa(t)^{-1} \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{2-\sigma}{1-\sigma}} Y(t) - \rho}{\eta_L \gamma_L^{\frac{\epsilon}{\sigma-1}} (1-\theta_E(t))^{\frac{1}{1-\sigma}} L - \rho} &\to \frac{1-\theta_E(t)-\sigma}{2-\sigma-\theta_E(t)} \end{aligned}$$

The implication follows from the no arbitrage conditions, Equation (73) and (47). The above implies

$$Y(t) \to \frac{\frac{1-\theta_E(t)-\sigma}{2-\sigma-\theta_E(t)}\left(\alpha\eta_L\gamma_L^{\frac{\epsilon}{\sigma-1}}(1-\theta_E(t)^{\frac{1}{1-\sigma}}L-\rho\right)+\rho}{\alpha^2\gamma_E^{\frac{\epsilon}{1-\sigma}}\theta_E(t)^{\frac{2-\sigma}{1-\sigma}}} := H(\theta_E(t))$$

where $H\colon (0,\theta_E(0)) \to \mathbb{R}_+$. Note H is differentiable. By assumption, $\theta_E(0)+\sigma < 1$, which can be shown to imply H has a negative derivative and hence H is decreasing. Since H is decreasing, H is injective. Moreover, since $(0,\theta_E(0))$ is open, by the open mapping theorem, H has a continuous inverse, which we now denote as $G : \mathbb{R}_+ \to (0,\theta_E(0))$.

Finally, we show for any $\epsilon > 0$, there exists T such that for all $t > T$,$|G(Y(t)) - \theta_E(t)| < \epsilon$. There exists $\delta > 0$ such that $|G(x)-G(y)| < \epsilon$ for $|x-y| < \delta$. Moreover, there exists T such that for all $t > T$, we have $|Y(t)-H(\theta_E(t))| < \delta$, giving $|G(Y(t))-\theta_E(t)| < \epsilon$. This establishes that $G(Y(t)) \to \theta_E(t)$ and in view of Equation (21), implies $\frac{E(t)}{Y(t)} \to G(Y(t))$.

Q.E.D.

PROOF OF CLAIM 3.1: We will use the growth equations and notation developed in the proof of Theorem 3.1. First, write the no-arbitrage conditions as follows

$$\begin{aligned}
\frac{\eta_E\tilde{z}_E}{\eta_L\tilde{z}_L} &= \frac{\eta_E\alpha\kappa(t)^{-1}\gamma_E^{\frac{\epsilon}{\sigma-1}}\tilde{\theta}_E(t)^{\frac{2-\sigma}{1-\sigma}}\tilde{Y}(t)-\rho}{\eta_L\gamma_L^{\frac{\epsilon}{\sigma-1}}\left(1-\tilde{\theta}_E(t)\right)^{\frac{1}{1-\sigma}}L-\rho} \\
&> \frac{\eta_E\alpha\kappa(t)^{-1}\gamma_E^{\frac{\epsilon}{\sigma-1}}\theta_E(t)^{\frac{2-\sigma}{1-\sigma}}Y(t)-\rho}{\eta_L\gamma_L^{\frac{\epsilon}{\sigma-1}}\left(1-\theta_E(t)\right)^{\frac{1}{1-\sigma}}L-\rho} \\
&= \frac{\eta_E z_E}{\eta_L z_L}
\end{aligned}$$

where the equality uses (47) and (73). The inequality follows from noting the RHS is decreasing in θ_E holding and since $\tilde{Y}(t) = Y(t)$. This serves to prove part 1. of the claim.

To show part 2. of the claim, note

$$\begin{aligned}
\eta_E\tilde{z}_E &= \eta_E\alpha\kappa(t)^{-1}\gamma_E^{\frac{\epsilon}{\sigma-1}}\tilde{\theta}_E(t)^{\frac{2-\sigma}{1-\sigma}}\tilde{Y}(t)-\rho \\
&> \eta_E\alpha\kappa(t)^{-1}\gamma_E^{\frac{\epsilon}{\sigma-1}}\theta_E(t)^{\frac{2-\sigma}{1-\sigma}}-\rho = \eta_E z_E
\end{aligned}$$

As such $\hat{\hat{Q}}_E(t) > \hat{Q}_E(t)$. Next, taking time derivatives of Equation (23), we have

$$\begin{aligned}
\frac{\hat{\hat{E}}(t)}{\hat{\hat{Y}}(t)} &= (1-\sigma)\frac{\hat{\hat{Q}}(t)_E}{\hat{\hat{Y}}(t)} + 1 \\
&= \frac{(\sigma-1)\,\hat{\hat{Q}}(t)_E}{\frac{1+\hat{\theta}_E(\sigma-1)}{1-\hat{\theta}_E}\hat{\hat{Q}}(t)_E - \hat{\hat{B}}(t)} \\
&= \left(\frac{\frac{1+\hat{\theta}_E(\sigma-1)}{1-\hat{\theta}_E(t)}\hat{\hat{Q}}(t)_E - \hat{\hat{B}}(t)}{(\sigma-1)\,\hat{\hat{Q}}_E(t)}\right)^{-1} + 1 \\
&= \left(\frac{1+\tilde{\theta}(t)_E\,(\sigma-1)}{\left(1-\tilde{\theta}(t)_E\right)(\sigma-1)} + (1-\sigma)^{-1}\left(1-\frac{\tilde{z}_L(t)}{\tilde{z}_E(t)}\right)\right)^{-1} + 1 \\
&< \left(\frac{1+\theta_E(t)\,(\sigma-1)}{(1-\theta_E(t))\,(\sigma-1)} + (1-\sigma)^{-1}\left(1-\frac{z_L(t)}{z_E(t)}\right)\right)^{-1} + 1 \\
&= \frac{\hat{E}(t)}{\hat{Y}(t)}
\end{aligned}$$

where the second equality use (42) along with (44) to derive

$$\hat{Y}(t) = \frac{1+\theta_E(t)(\sigma-1)}{1-\theta_E(t)}\hat{Q}_E(t) - \hat{B}(t) \tag{75}$$

thus we have shown $\frac{\hat{\hat{E}}(t)}{\hat{\hat{Y}}(t)} < \frac{\hat{E}(t)}{\hat{Y}(t)}$, which directly implies part 3 of the claim.

Q.E.D.

APPENDIX B: FURTHER DATA ANALYSIS

B.1. *U.S. Data Sources*

B.1.1. *GDP, Implicit Price Deflator, Employment, and Wages*

This data is all sourced from the U.S. Bureau of Economic Analysis (BEA) National Income and Product Accounts (NIPA). GDP data is real GDP in chained 2009 dollars from 1929-2015. We use two measures of labor input: Full time equivalent employees (1929-2015) and hours worked from 1948-2015. Total wages are given by "compensation of employees" and deflated to 2009 using the GDP implicit price deflator. We then compute labor productivity and wage series per FTE worker and per hour for the two periods. We extended the GDP series to 1900 using the growth rates from estimates of the GNP from the Historical Statistics of the United States: Colonial Times to 1970 (HS70) Annual data on GNP is available from 1889. Page 236 in the latter source provides estimates of the labor share of national income for years prior to 1929. Page 224 has some estimates of the national income prior to 1929. Prior to 1919 these are for 5-year averages. For these earlier used we used these to obtain the ratio of national income to GNP for each 5-year period and then estimated annual national income by multiplying the ratio by annual GNP. We then multiplied average labor compensation shares in the national income to estimate labor compensation back to 1900. Historical Statistics of the United States: Millennial Edition gives employment from 1890 to 1990. We extend this forward to 2015 using the growth rate of total employment in the NIPA. We compute labor productivity and annual wage series per employee from 1900 on.

B.1.2. *Primary Energy Consumption and Heat Rates*

We use data from the U.S. Energy Information Administration (EIA) website for 1949-2015 for consumption of the following energy carriers: Coal, natural gas, petroleum, nuclear electricity, hydroelectricity, geothermal energy, solar energy, wind energy, and biomass, in quadrillion BTU. The energy totals given for the non-combustible renewables and nuclear power are the equivalent quantity of fossil fuels that would be needed to generate the same amount of electricity. We used these heat rates, which are supplied by EIA to convert the price of electricity to a price per BTU of primary energy. Earlier energy quantity were obtained from the HS70 and Appendix D of the EIA's Monthly Energy Review. These data also include animal feed, which was a large source of energy in 1900.

B.1.3. *Energy Prices*

Using a combination of documents and databases on the EIA website we assembled fossil fuel production prices from the earliest available date to 2015. Oil wellhead prices are available in the online interactive data from 1859 to 2015. The natural gas wellhead price is available from 1922 but discontinued after 2012. For 2013-15 we use Henry Hub spot prices available on this page:

`http://www.eia.gov/dnav/ng/hist/rngwhhdA.htm`

Coal prices were available for 1949-2011 from the 2011 Annual Energy Review. For 2012-2015 we used the Annual Coal Reports. Biomass prices for 1970 to 2014 are available from this webpage:

`http://www.eia.gov/state/seds/data.cfm?incfile=/state/seds/sep_prices/total/pr_tot_US.html&sid=US`

Electricity prices are available from the EIA from 1960 to 2015. We use the industrial electricity price as a proxy for the wholesale price of electricity. Earlier energy prices were obtained from the HS70. For electricity prices we used the price series for "large users" and assumed the nominal price per million BTU was constant prior to 1917. For the price of animal feed we used the price per bushel of oats received by farmers from HS70. We assumed 32lbs of oats per bushel and 12MJ per kg of feed Pagan (1998). We use the price of lumber from HS70 to project back the cost of biomass energy for years before 1970. We use the growth rates of the price of oil to project natural gas prices back to 1900. In computing the total value of energy we multiply the electricity produced by nuclear and non-combustible renewables by the price of electricity. We then compute energy productivity and price series for raw BTU of primary energy and for quality-adjusted energy use.

B.2. *Energy Intensity and Output in the Cross-Section*

Details of cross-sectional data used in this section are in Stern et al. (2016). Figure B.2 gives a scatter plot of the natural log of energy intensity against the natural log of GDP per capita for five years between 1971 and 2010. The fitted regression line is a cubic polynomial; we compared the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC) for a linear model and cubic model, and found the cubic has the lowest score on both criteria. The polynomial regression shows the relationship between energy intensity is negative for lower level of energy intensity and output.

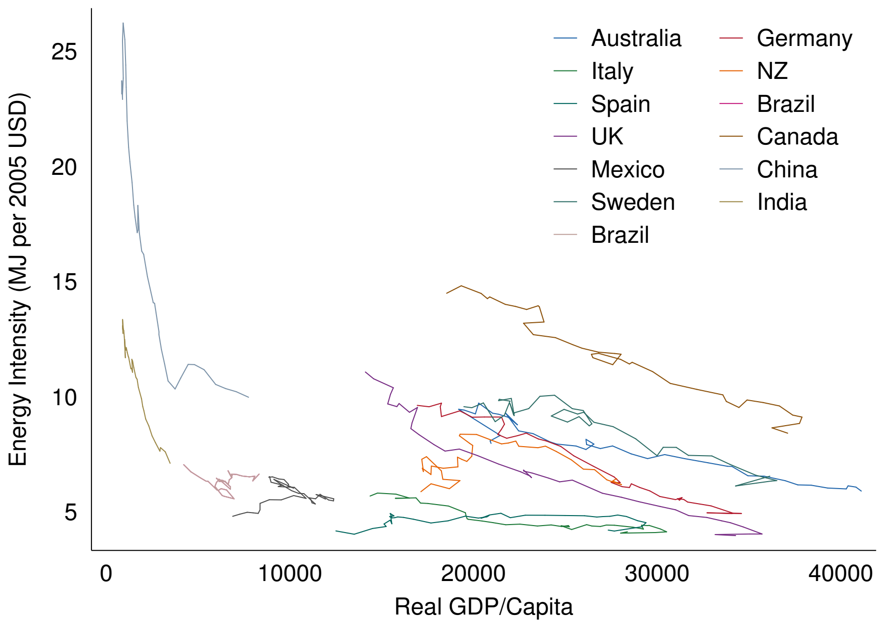

FIGURE B.1.— Energy intensity and real GDP/capita for selected countries between 1970 and 2010.

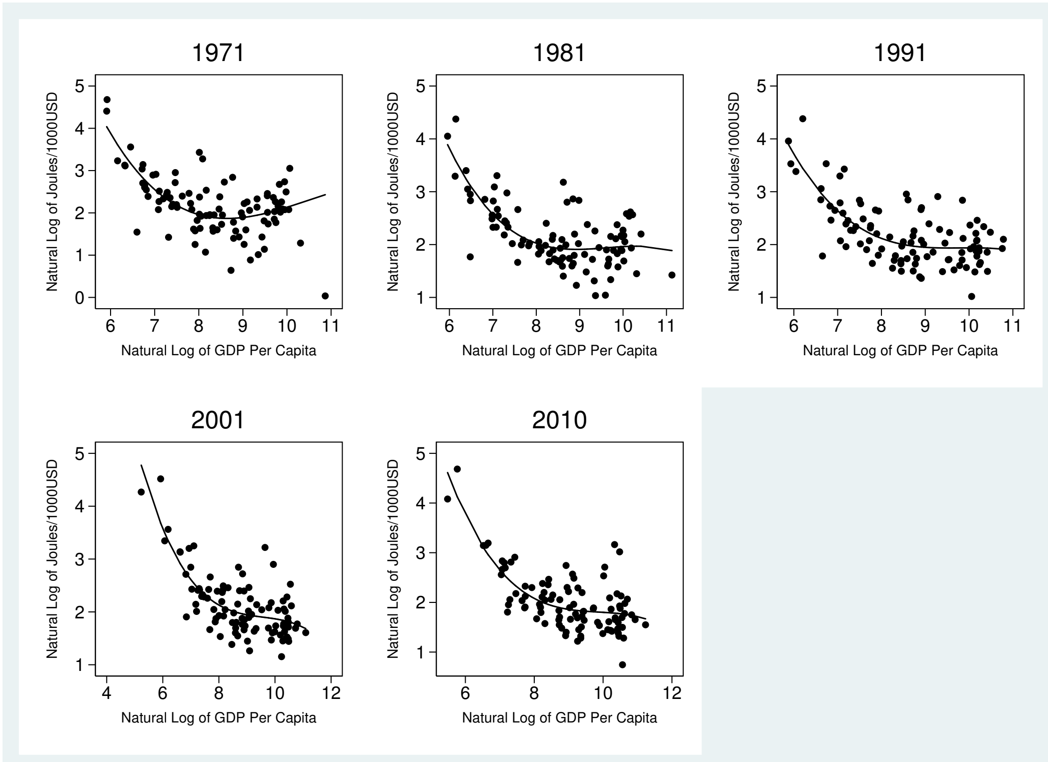

FIGURE B.2.— Caption

Using the polynomial coefficients from table B.2, we can estimate the elasticity of energy intensity to output in the cross-section. The elasticities are given in table B.2, and they reflect the analysis of the elasticity of energy intensity to output through time, from section 2. In particular, there is a negative relationship between energy intensity and output, however, for lower levels of output and energy intensity, the the fall in energy intensity associated with an increase in output is higher. For example, in 2010, for countries at a GDP per capita of 2,000 USD, a 1 percent increase in output is associated with a .5 percent fall in energy intensity, while for countries with a GDP per capita of 30,000, a 1 percent increase in output is associated with .06 percent fall in energy intensity.

TABLE B.1
REGRESSION TABLE

	1971	1981	1991	2001	2010
Ln(Y)	-11.28	-12.25**	-9.413*	-8.740**	-9.215**
	(-1.62)	(-2.81)	(-2.20)	(-3.18)	(-3.00)
$Ln(Y)^2$	1.111	1.280*	0.960	0.906**	0.931*
	(1.29)	(2.44)	(1.86)	(2.71)	(2.53)
$Ln(Y)^3$	-0.0354	-0.0443*	-0.0326	-0.0316*	-0.0315*
	(-1.00)	(-2.13)	(-1.58)	(-2.35)	(-2.17)
Constant	39.19*	40.84***	32.67**	30.22***	32.34***
	(2.13)	(3.41)	(2.81)	(4.09)	(3.84)
Observations	98	99	99	99	99

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE B.2
ELASTICITY OF ENERGY INTENSITY TO OUTPUT IN THE CROSS-SECTION

USD/ capita	1971	1981	1971	2001	2010
2,000	-0.371	-0.720	-0.470	-0.444	-0.522
10,000	0.405	-0.195	-0.026	-0.093	-0.082
20,000	0.574	-0.182	0.010	-0.093	-0.043
30,000	0.627	-0.233	-0.013	-0.135	-0.063

B.3. *U.S. Fuel Prices*

Figure B.3 shows the real price in 2009 US Dollars per BTU of each of the individual fuels. Even though each of these prices has increased absolutely over time, it is unclear whether these represent systematic trends or not. The energy price index rises faster than most of the component prices because there is a positive correlation between price movements and cost shares.

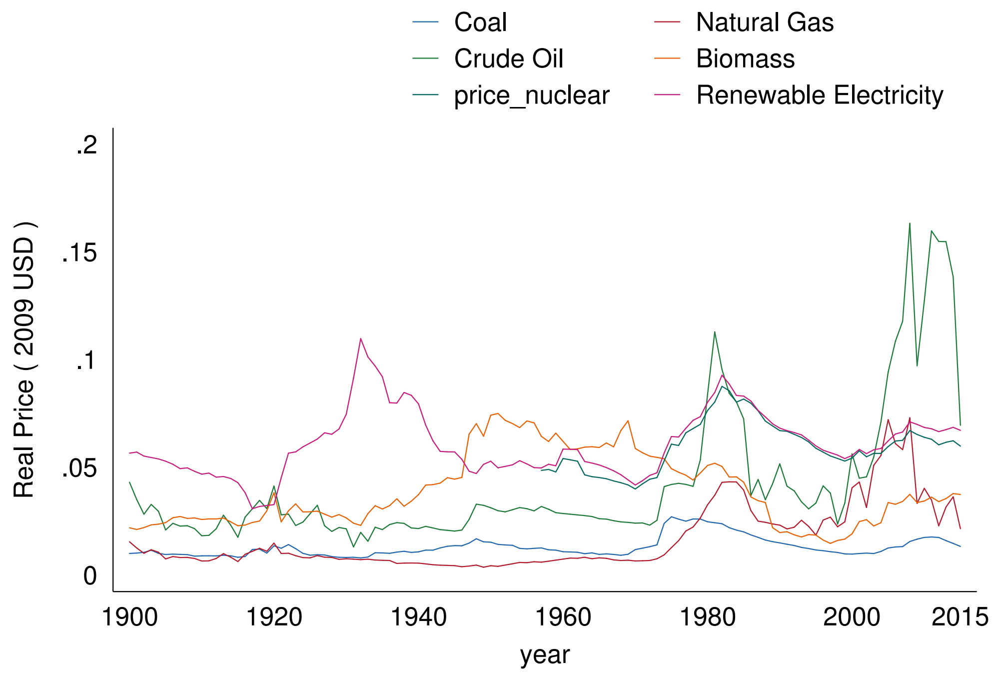

FIGURE B.3.— Caption

Figure B.4 and figure B.5 shows the energy intensity and aggregate U.S. fuel price with and without animal feed. Animal was a larger part of the fuel mix prior to the 1950s, after which both the price and energy intensity series converge. Note a clear positive trend in the price index without animal feed. (Why?)

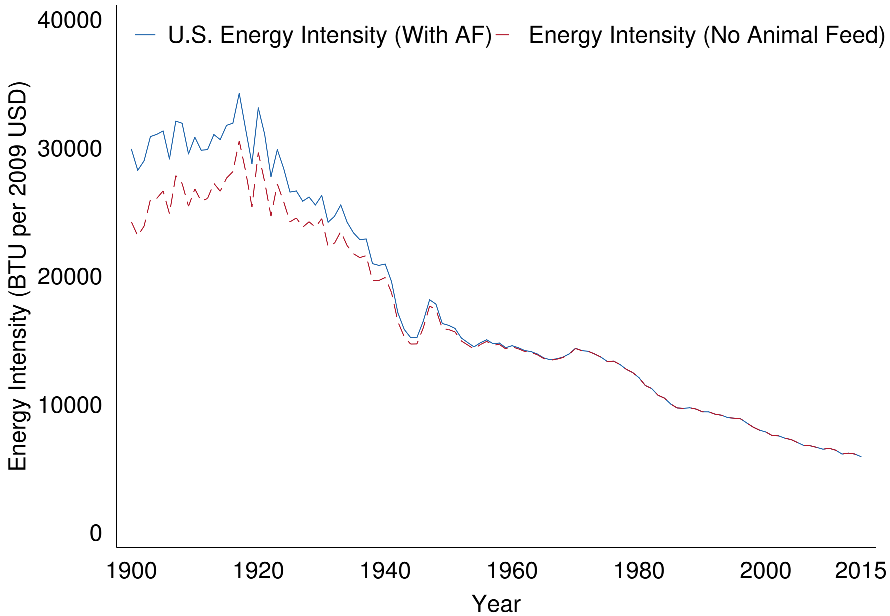

FIGURE B.4.— U.S. energy intensity with and without animal feed (AF)

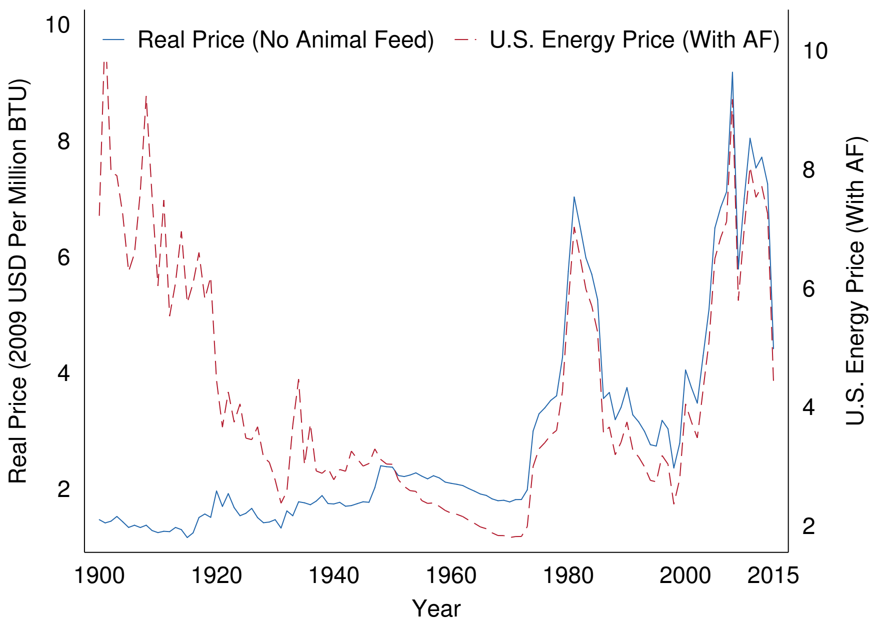

FIGURE B.5.— U.S. aggregate fuel price with and without animal feed (AF)

To test for trends, we use the t_{DAN} Fomby and Vogelsang (2003) or Dan-J Bunzel and Vogelsang (2005) test based on a modified t-test on the slope parameter of the simple linear trend regression model:

$$y_{i,t} = \beta_{1,i} + \beta_{2,i} t_i + u_{i,t} \tag{76}$$

where t is a linear time trend, i indicates the location or sample period of the data, u, is a stochastic process that may or may not be stationary and and β_1 and β_2 are regression parameters to be estimated. Then the trend test statistic is given by:

$$t_{DAN} = \frac{\hat{\beta}_{2,i}}{\text{se}\left(\hat{\beta}_{2,i}\right)} e^{-bJ}$$

where $\hat{\beta}_{2,i}$ is the estimate of the slope parameter and se $\left(\hat{\beta}_{2,i}\right)$ its standard error, b is a parameter derived by Bunzel and Vogelsang (2005) and

$$JSS = \frac{RSS_1 - RSS_4}{RSS_4}$$

where RSS_1 is the sum of squared residuals from (76), and RSS_4 is the sum of squared residuals from the following regression:

$$y_t = \sum_{i=0}^{9} \gamma_i t^i + v_t$$

The standard error $\mathrm{se}\left(\hat{\beta}_{2,i}\right)$ is computed as follows

$$\mathrm{se}\left(\hat{\beta}_{2,i}\right) = \sqrt{\hat{\sigma}\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}}$$

with $\bar{t} = T^{-1}\sum_{t=1}^{T} t$, where T is the sample size and

$$\hat{\sigma}^2 = \hat{\gamma}_0 + 2\sum_{j=1}^{T-1} \frac{\sin(j\pi/M)}{j\pi/M}\gamma_j$$

where $\hat{\gamma}_j = T^{-1}\sum_{t=j+1}^{T} \hat{u}_t\hat{u}_{t-j}$ is a function of the estimated residuals $\hat{u}$ and $M = \max\{0.02T, 2\}$.

The recommended value of b and the critical values of t_{DAN} for a two-tailed test are as follows: $b = 2.466$, $t_DAN = 2.462$ at 1%; $b = 1.795$, $t_DAN = 2.052$ at 2.5%; $b = 1.322$, $tDAN = 1.710$ at 5% and $b = 0.965$, $t_DAN = 1.329$ at the 10% significance level. Further values can be derived from the formulae in Bunzel and Vogelsang (2005). J can also be used in a left-tailed test of the null hypothesis that the errors in (76) contain a unit root autoregressive process or random walk. The null hypothesis is rejected for small values of the statistic. The critical values are 0.488 at 1%, 0.678 at 2.5%, and 0.908 at the 5% significance levels.

TABLE B.3

TREND TEST RESULT FOR INDIVIDUAL PRICE SERIES

Log real price	N	β	$t_{DAN}01$	$t_{DAN}025$	$t_{DAN}05$	$t_{DAN}10$	jt
Coal	116	0.00454	0.05304	0.17444	0.40376	0.76072	1.77433
Natural Gas	116	0.01603	3.62E-04	0.00528	0.03497	0.1456	3.99518
Crude Oil	116	0.01168	0.06151	0.22571	0.56434	1.12699	1.93742
Biomass	116	0.00126	1.80E-10	7.36E-08	5.10E-06	1.25E-04	8.96129
Nuclear	60	0.0049	4.43E-24	1.67E-17	7.24E-13	2.29E-09	22.56978
Renewable	116	0.00231	0.00888	0.04083	0.11966	0.26939	2.27309
Animal Feed	71	-0.02972	-0.10661	-0.43794	-1.18566	-2.51437	2.10566

There is no significant positive trend in any of the series except Animal Feed, which has a significant negative trend at the 10% level. All have a unit root, so the series are not stationary, but apart from animal feed they do not have a significant drift either.

We then tested for trends in the aggregate price indices with the following results

TABLE B.4
TREND TEST RESULT FOR AGGREGATE PRICE SERIES

Log real price	β	$t_{DAN}01$	$t_{DAN}025$	$t_{DAN}05$	$t_{DAN}10$	jt
Raw Energy Price with AF	-2.00E-03	-6.64E-07	-3.25E-05	-5.04E-04	-0.004	5.79794
Raw Energy Price without AF	0.01375	0.64543	1.45564	2.58247	3.98066	1.21205

For the data with animal feed there is no trend. For the series without animal feed there is a positive trend at the 5% significance level for the raw series.

REFERENCES

Acemoglu, D. (1998). Why Do New Technologies Complement Skills ? Directed Technical Change and Wage Inequality. *The Quarterly Journal of Economics*, 113(4):1055–1089.

Acemoglu, D. (2002). Directed Technical Change. *The Review of Economic Studies*, 69(4):781–809.

Acemoglu, D. (2003). Labour and Capital Augmenting Technical Change. *Journal of the European Economic Association*, 1(1):1–37.

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton University Press, Princeton, New Jersey.

Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012). The Environment and Directed Technical Change. *American Economic Review*, 102(1):131–166.

Acemoglu, D. and Cao, D. (2015). Innovation by entrants and incumbents. *Journal of Economic Theory*, 157:255–294.

Acemoglu, D., Gancia, G., and Zilibotti, F. (2015). Offshoring and Directed Technical Change. *American Economic Journal: Macroeconomics*, 7(3):84–122.

Aghion, P., Hemous, D., Martin, R., Reenen, J. V., and Dechezleprêtre, A. (2016). Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry. *Journal of Political Economy*, 124(1):1–52.

Aghion, P. and Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60(2).

André, F. J. and Smulders, S. (2014). Fueling Growth When Oil Peaks: Directed Technological Change and the Limits to Efficiency. *European Economic Review*, 69:18–39.

Bloom, N., Jones, C. I., Van Reen, J., and Webb, M. (2017). Are Ideas Getting Harder to Find?

Bunzel, H. and Vogelsang, T. J. (2005). Powerful Trend Function Tests That Are Robust to Strong Serial Correlation, with an Application to the Prebisch: Singer Hypothesis. *Journal of Business & Economic Statistics*, 23(4):381–394.

EIA (2017). Annual Energy Outlook. Technical report, U.S. Energy Information Admistration.

Fomby, T. and Vogelsang, T. (2003). Test of Common Deterministic Trend Slopes Applied to Quarterly Global Temperature Data. In Fomby, T. and Hill, R., editors, *Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later, Advances in Econometrics 17*. Elsevier.

Goulder, L. H. and Schneider, S. H. (1999). Induced Technological Change and The Attractiveness of CO2 Abatement Policies. *Resource and Energy Economics*, 21(3-4):211–253.

Hamilton, J. D. (2009). Understanding crude oil prices. *The Energy Journal*, 30(2):179–207.

Hassler, J., Krusell, P., and Olovsson, C. (2016). Directed Technical Change as a Response to Natural-Resource Scarcity.

IEA (2016). World Energy Outlook 2016.

Jakeman, G., Hanslow, K., Hinchy, M., Fisher, B. S., and Woffenden, K. (2004). Induced Innovations and Climate Change Policy. *Energy Economics*, 26(6):937–960.

Jones, C. (1995). R&D-Based Models of Economic Growth. *The Journal of Political Economy*, 103(4):759–784.

Kaufmann, R. K. (2004). The Mechanisms for Autonomous Energy Efficiency Increases : A Cointegration Analysis of the US Energy / GDP Ratio. *The Energy Journal*, 25(1):63–86.

Löschel, A. (2002). Technological Change in Economic Models of Environmental Policy: A Survey. *Ecological Economics*, 43(2-3):105–126.

Mckibbin, W. J., Pearce, D., and Stegman, A. (2004). Long Run Projection For Climate Change Scenarios. *Working Papers in International Economics*, (104).

Newell, R. G., Jaffe, A. B., and Stavins, R. N. (1999). The Induced Innovation Hypothesis and Energy-Saving Technological Change. *Quarterly Journal of Economics*, 114(3):941–975.

Nordhaus, W. (2002). Modeling Induced Innovation in Climate Change Policy. In Grubler, A., Nakićenović, N., and Nordhaus, W., editors, *Technological Change and the Environment*, pages 259–290. Resources for the Future Press.

Pagan, J. (1998). Carbohydrates in Equine Nutrition. In Pagan, J., editor, *Advances in Equine Nutrition*. Nottingham University Press, Nottingham.

Popp, D. (2004). ENTICE: endogenous technological change in the DICE model of global warming. *Journal of Environmental Economics and Management*, 48(1):742–768.

Popp, D., Newell, R. G., and Jaffe, A. B. (2010). *Energy, The Environment, and Technological Change*, volume 2. Elsevier B.V.

Stern, D., Csereklyei, Z., and Rubio- Varas, M. (2016). Energy and Economic Growth : The Stylized Facts Energy & The Economy. *Energy Economics*, 37(2).

Stern, D. I. (2004). Economic Growth and Energy. In Cleveland, C. J., editor, *Encyclopedia of Energy*, volume 2, pages 35–51. Academic Press, San Diego, CA.

Stern, D. I. and Kander, A. (2012). The Role of Energy in the Industrial Revolution and Modern Economic Growth. *The Energy journal*, 33(3):125–152.

Sue Wing, I. (2008). Explaining The Declining Energy Intensity of the U.S. Economy. *Resource and Energy Economics*, 30(1):21–49.

Sue Wing, I. and Eckaus, R. S. (2007). The Implications of the Historical Decline in US Energy Intensity for Long-Run CO2 Emission Projections. *Energy Policy*, 35(11):5267–5286.

Sun, Y. and Zhang, Y. (2009). Individual Risk and Lebesgue Extension Without Aggregate Uncertainty. *Journal of Economic Theory*, 144(1):432–443.

Van der Werf, E. (2008). Production Functions for Climate Policy Modeling: An Empirical Analysis. *Energy Economics*, 30(6):2964–2979.

Warr, B., Ayres, R., Eisenmenger, N., Krausmann, F., and Schandl, H. (2010). Energy use and economic development: A Comparative Analysis of Useful Work Supply in Austria, Japan, the United Kingdom and the US during 100 Years of Economic Growth. *Ecological Economics*, 69(10):1904–1917.

Webster, M., Paltsev, S., and Reilly, J. (2008). Autonomous efficiency improvement or income elasticity of energy demand: Does it matter? *Energy Economics*, 30(6):2785–2798.

Williams, R. H., Lave, L. B., and Perry, A. M. (1990). Low-Cost Strategies for Coping with CO Emission Limits (A Critique of CO Emission Limits: an Economic Cost Analysis for the USA by Alan Manne and Richard Richels). *The Energy Journal*, 11(4):35–59.