Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Energy intensity, growth and technical change

CAMA Working Paper 46/2018
September 2018

Akshay Shanker
Crawford School of Public Policy, ANU
Centre for Applied Macroeconomic Analysis, ANU

David Stern
Crawford School of Public Policy, ANU
Centre for Applied Macroeconomic Analysis, ANU

Abstract

World and U.S. energy intensities have declined over the past century, falling at an average rate of approximately 1.2–1.5 percent a year. The decline has persisted through periods of stagnating or even falling energy prices, suggesting the decline is driven in large part by autonomous factors, independent of price changes. In this paper, we use directed technical change theory to understand the autonomous decline in energy intensity and investigate whether the decline will continue. We show in an economy with no state-dependence, where existing knowledge does not make R&D more profitable, energy intensity continues to decline, albeit at a slower rate than output growth, due to energy-augmenting innovation. However, in an economy with extreme state-dependence, energy intensity eventually stops declining because labor-augmenting innovation crowds out energy-augmenting innovation. Our empirical analysis of energy intensity in 100 countries between 1970 and 2010 suggests a scenario without extreme state dependence where energy intensity continues to decline; in either case, energy intensity never declines faster than output grows, and so energy use always increases, as long as the extraction cost of energy stays constant.

Keywords

Energy, Directed Technological Change, Economic Growth

JEL Classification

O33, O41, Q43

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

.

ENERGY INTENSITY, GROWTH AND TECHNICAL CHANGE

Akshay Shanker[†] & David Stern [⋆] [‡]

Center for Applied Macroeconomic Analysis, Crawford School of Public Policy, Australian National University, 132 Lennox Crossing, ACT 2601, Australia

27 September 2018

World and U.S. energy intensities have declined over the past century, falling at an average rate of approximately 1.2–1.5 percent a year. The decline has persisted through periods of stagnating or even falling energy prices, suggesting the decline is driven in large part by autonomous factors, independent of price changes. In this paper, we use directed technical change theory to understand the autonomous decline in energy intensity and investigate whether the decline will continue. We show in an economy with no state-dependence, where existing knowledge does not make R&D more profitable, energy intensity continues to decline, albeit at a slower rate than output growth, due to energy-augmenting innovation. However, in an economy with extreme state-dependence, energy intensity eventually stops declining because labor-augmenting innovation crowds out energy-augmenting innovation. Our empirical analysis of energy intensity in 100 countries between 1970 and 2010 suggests a scenario without extreme state dependence where energy intensity continues to decline; in either case, energy intensity never declines faster than output grows, and so energy use always increases, as long as the extraction cost of energy stays constant.

Keywords: Energy, Directed Technological Change, Economic Growth

JEL Classification: O33, O41, Q43

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

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

[‡]We thank Hilde Bjørnland, Evan Comfort, Sofia Henriques, Astrid Kander, Per Krusell, Warwick McKibbin, and participants at the Scandinavian Energy Transition Workshops, 2018 Sustainable Resource Use and Economic Dynamics conference, the 2018 World Congress of Environmental and Resource Economists, the 2018 European Econometric Society meeting and seminars at the BI Norwegian Business School and Umeå University 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, the Research School of Asia and the Pacific at the Australian National University, the Center for Applied Macroeconomic Analysis and the Australian Research Council, Discovery Project (DP160100756) "Energy Efficiency Innovation, Diffusion and the Rebound Effect."

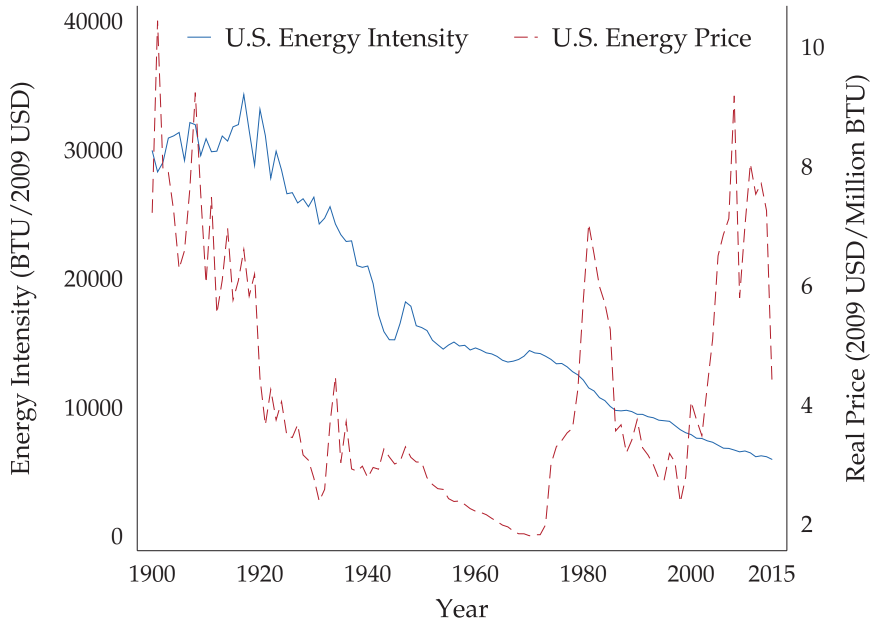

FIGURE 1.— U.S. energy intensity (including animal feed) and the real price of primary energy between 1900 and 2015. See the Appendix, Section C.1, for a detailed description of the data sources.

1. INTRODUCTION

Why does energy intensity, defined as joules of energy per unit of real output, fall as an economy grows? Why does energy intensity fall slower than the rate of output growth? Will energy intensity continue to fall in the absence of sustained energy price increases? In this paper, we use a model of endogenous growth with directed technical change to address these questions.

The past century has seen a persistent decline in energy intensity globally and in many individual countries. Figure 1 shows energy intensity for the United States (U.S.) and the real price of energy from 1900 to 2015. U.S. 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. Figure 1 also shows that the real price of energy fell between 1900 and 1950 and has fluctuated between 1950 and 2010. Whether or not there is an increasing trend in the price of energy in recent decades is unclear.[1]

[1]We conduct and discuss detailed tests on energy price trends in the Appendix, Section C.2. We find no trend in the price of individual fuels and energy carriers or the aggregate price (total cost of divided by BTU) including animal feed. However, we find a positive trend for the aggregate price

Moreover, even during periods of falling prices, particularly between 1980 and 2000, energy intensity continued to decline.

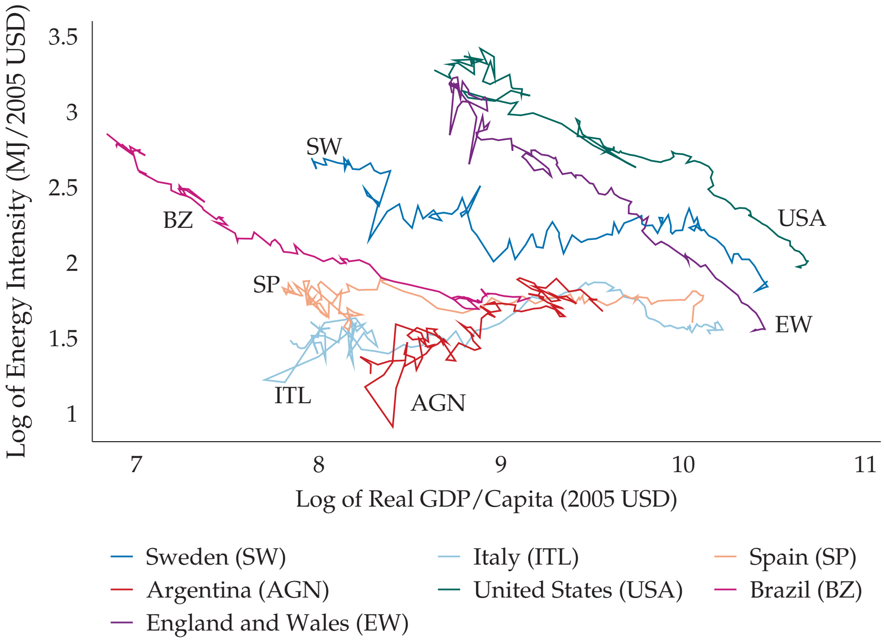

FIGURE 2.— Energy intensity and real GDP per capita for selected countries between 1900 and 2010.

The decline in energy intensity is also evident around the world. Figure 2 shows energy intensity and GDP for selected countries between 1900 and 2010 and Figure E.8 in the Appendix for 100 countries between 1971 and 2010 — per capita GDP growth is associated with a proportional decline in energy intensity. Moreover, as Figure 3 shows, there is a negative relationship between energy intensity and GDP per capita in the cross section of countries. Csereklyei et al. (2016) analyze data for 100 countries between 1970 and 2010 and find that the elasticity of energy intensity with respect to per capita GDP is -.3 in the cross-section and through time; the authors also find conditional and unconditional beta convergence and sigma convergence of energy intensity.

Because of the ambiguity over whether energy prices are trending upwards or not and because energy intensity continued to decline during periods of stagnating prices, energy economists have suspected that autonomous factors *tied to growth*,

excluding animal feed, possibly, in part driven by a positive correlation between cost shares and price movements. Animal feed was a significant part of the fuel mix in the U.S. at 1900, accounting for 20 percent of primary energy input. The share of animal feed fell to zero by 1960.

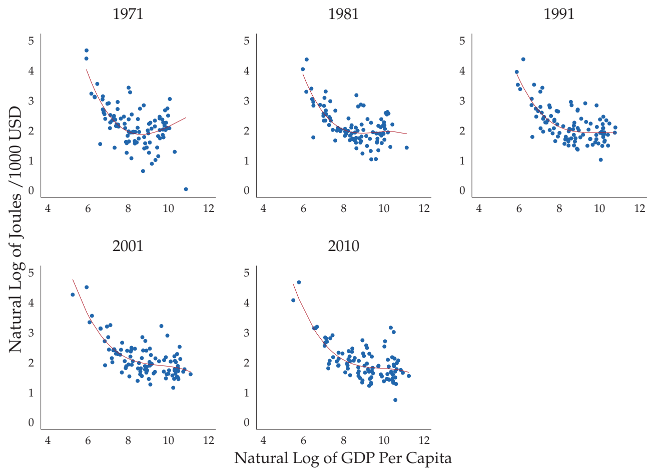

FIGURE 3.— Natural log of energy intensity and natural log of output for a cross section of countries. Csereklyei et al. (2016) describe the sources of the data. The regression lines are cubic polynomials, see the Appendix D.2 and also Section 3.

independent of price increases or energy scarcity, have played a role in the world-wide decline of energy intensity.[2] Indeed, climate policy models[3] 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 autonomous decline of energy intensity will continue and will be more rapid than in the past (Stern, 2017).

However, despite energy intensity being "central to the achievement of a range of policy goals, including energy security, economic growth and environmental sustainability" (IEA, 2017) and despite the well-established literature incorporating endogenous technical change in models estimating the costs of climate policy (Goulder and Schneider, 1999; Nordhaus et al., 2002; Jakeman et al., 2004; Popp, 2004), there has been no formal analytical study of what, in terms of economic incentives, could drive an autonomous decline in energy intensity and whether the autonomous, non-price, decline will continue. We undertake such an analysis in

[2]The climate policy modeling literature refers to the autonomous decline of energy intensity as Autonomous Energy Efficiency Improvement (AEEI). Though extensively debated, estimates of AEEI for the U.S. 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)). Newell et al. (1999), who use micro-level data on efficiency of air conditioners, find price increases induce technical change but "autonomous drivers of energy efficiency explain up to 62% of total changes in energy efficiency".

[3]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 an exogenous efficiency trend, for example, Popp (2004).

this paper.

We build a Schumpetarian endogenous growth model incorporating directed technical change (Acemoglu, 1998, 2002), where innovation can augment labor or energy. The quality improvement framework we use can be traced to Aghion and Howitt (1992), however, our treatment with a large number of firms most closely follows Acemoglu (2009), Chapter 14.[4] Following established estimates in the literature (see Stern and Kander (2012) and also Van der Werf (2008)), we assume that the elasticity of substitution between energy and labor services is less than one.

Throughout the paper, we analyze growth paths where the price of energy is constant. Our contention is not that fluctuating energy prices do not drive changes in energy intensity — Popp (2002) finds clear evidence that energy-augmenting innovation responds to prices — rather, by holding prices constant, we wish to isolate and understand the non-price component of energy intensity improvements and its long-run potential.

1.1. *Main findings*

Our first finding is that falling energy intensity without increasing real energy prices is *inconsistent* with balanced growth.[5] Rather, an autonomous decline of energy intensity implies structural change: a falling cost share of energy and a non-constant, but possibly converging growth rate of output. However, while the cost share falls, the quantity ratio of energy services to labor services increases — the rise in the labor services cost share is due to the increase in the wage rate. Using U.S. data on energy prices, wages, employment, energy use and GDP, we estimate the quantity ratio of labor services to energy services and find the share has indeed been falling suggesting the driver of falling energy intensity is energy-augmenting technical change, rather than structural change towards labor services. This suggests an answer to our first question: as Henriques and Kander (2010) found, energy intensity has been falling precisely because of energy-augmenting technological change, not structural change (in value terms) towards labor service sectors of the economy.

Since we can no longer use balanced growth paths (BGPs), we proceed to formalize asymptotic behavior. Motivated by evidence (Bloom et al., 2017) that innovation becomes harder as technologies advance, we begin with an economy without state dependence. Under no state dependence, the existing level of knowledge does not improve the profitability of research. In this case, we show asymptotic convergence to a growth path where energy intensity falls at a constant rate due to investment

[4]The structure of our modeling is also similar to Acemoglu and Cao (2015), except Acemoglu and Cao (2015) also incorporate radical innovations while we incorporate directed technical change.

[5]As defined by Hassler et al. (2016) — see Subsection 2.3.

in energy-augmenting technologies. Consistent with the data, energy intensity declines slower than output grows, implying energy use continues to increase. We also find that the asymptotic elasticity between energy intensity and output can be expressed as a function of the elasticity of substitution between energy and labor.

Along a growth path where real energy prices are constant, energy use increases, energy-augmenting technologies advance, and the price of energy services falls. The fall in the price of energy services reduces profitability and incentives for energy-augmenting research.[6] However, since the use of energy increases, the market size of energy services expands, improving the incentives to perform research that advances energy-augmenting technologies. In the scenario with no state dependence, the growing incentives from the expanding market size are enough to sustain energy-augmenting research.

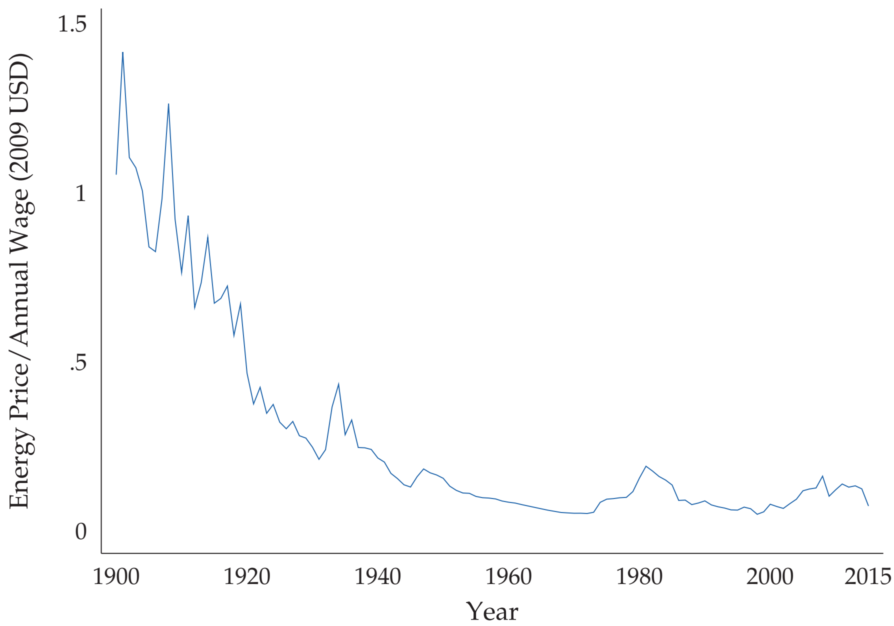

FIGURE 4.— Ratio of U.S. real energy price to real annual wage per employee.

Nonetheless, the rate of decline of energy intensity is slower than the rate of output growth because energy-augmenting technologies advance at a slower rate than

[6]To understand incentives to undertake energy-augmenting research, recall that profits of monopolistically competitive entrepreneurs are the product of the price a new invention can command and the number of units of the new product that can be sold, the market size. See Acemoglu (2009), ch. 15 for a detailed explanation of these market size and price effects.

labor-augmenting technologies. While the market size effect of increasing energy use is sufficient to maintain energy-augmenting innovation, the price effect of falling energy prices relative to wages (Figure 4) is stronger than the market size effect of greater per capita energy use, because the elasticity of substitution is less than one. The stronger price effect means there are greater incentives to undertake labor-augmenting research and in turn a faster rate of labor-augmenting innovation.

We also examine an economy with extreme state dependence.[7] We find that an economy with extreme state dependence has a BGP. While autonomous reductions in energy intensity can occur outside a BGP, in the absence of price changes, energy intensity eventually converges to a strictly positive constant. The reason is that state dependence compounds the stronger incentive to undertake labor-augmenting research arising from the price effect. The relative incentives to undertake labor-augmenting research continue to grow until they crowd out all energy-augmenting research, leading to an eventual end to the autonomous decline in energy intensity.

Informally observing the data on energy intensity trends suggests the extreme state dependence scenario where the decline in energy intensity eventually stops. For instance, across the cross-section of countries, energy intensity falls at a relatively low rate with respect to output for high income countries (see Figure 3 and discussion in Section D.2 of the Appendix) and many countries with low energy intensity do not experience a strong decline in energy intensity through time in Figure 2. However, the formal econometric analysis we conduct lends support to a scenario with no extreme state dependence and a continued decline in energy intensity. In particular, we use a polynomial regression to predict the rate of decline of energy intensity for given levels of output and energy intensity and find that there is a clear path along which energy intensity continues to decline at a rate consistent with the model without extreme state dependence.

1.2. *Related literature*

This paper 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 intensity improvements in generating possibilities of

[7] To clarify terminology, throughout the paper, we use the term extreme state dependence to refer to state dependence of the form where the elasticity of existing knowledge to producing new knowledge is one. When the elasticity is less than one, we have limited state dependence. While we do not explicitly analyze a model with incomplete state dependence, we present arguments suggesting the dynamics of an incomplete state dependence scenario will be similar to the scenario with no state dependence (Section 2.5.2).

growth where energy is scarce. André and Smulders (2014) use a directed technical change model to account for stylized facts of U.S. energy use. However, both authors do not explicitly analyze the role and potential for autonomous drivers of falling energy intensity to persist along a growth path. In particular, both papers also only study innovation under state dependence and use a Hotelling rule featuring an exponentially increasing energy price, which becomes the driver of long-run reductions in energy intensity.

Our finding that consumption of energy services increases along a growth path is similar to observations by Hart (2018) who builds a model in which consumption patterns shifts towards energy intensive goods. While ours is a supply side model, Hart (2018) also observes how energy efficiency of energy intensive services has been rising faster than the energy intensity of output: an implication of an increasing quantity share of energy services along a growth path.[8]

Our key finding that a model with extreme state dependence cannot feature a constant growth rate of energy-augmenting technical change along an asymptotic growth path is analogous to the main finding of Acemoglu (2003). Acemoglu (2003) shows how, under state dependence, long-run technical change can only be labor-augmenting rather than capital-augmenting for interest rates to remain constant. Our model under extreme state dependence is similar to Acemoglu's, except with capital replaced by energy. A specification with limited or no state dependence is ruled out by Acemoglu precisely because it would lead to a fall in capital intensity given constant interest rates — just as energy intensity falls with constant energy prices in the model presented here with no state dependence.

Finally, Stern and Kander (2012) econometrically fit an exogenous technical change growth model to Swedish data for the 19th and 20th Centuries finding that the rate of energy-augmenting technical progress slowed and the rate of labor-augmenting technical progress accelerated over the 19th and 20th Centuries in Sweden as energy intensity declined. Our model provides a theoretical explanation. Sweden initially had high energy intensity and has converged towards lower energy intensity countries as shown in Figure 2. Therefore, even under no state-dependence the rate of energy-augmenting technical change should have been high initially and fall as the rate of labor-augmenting technical change increased.

[8]See equation 41 in section (4).

2. SCHUMPETERIAN GROWTH MODEL OF ENERGY AND DIRECTED TECHNICAL CHANGE

2.1. *The model economy*

Consider a continuous time economy where t, with $t \in \mathbb{R}_+$, indexes time. Assume a risk neutral representative consumer, so preferences over a time path for consumption are:[9]

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

where ρ is the discount factor and $C(t)$ is consumption at time t. The consumer faces a standard dynamic optimization problem and chooses a path of consumption and assets [10], $(a(t))_{t=0}^{\infty}$, given a path of interest rates, $(r(t))_{t=0}^{\infty}$ and wage rates $(w_L(t))_{t=0}^{\infty}$ to maximize (1).[11] For any t, the resource constraint for the economy is:

$$Z(t) + C(t) + X(t) + \kappa(t)E(t) \leqslant Y(t) \tag{2}$$

where $Z(t)$ is the total level of R&D, $X(t)$ is total expenditure on machine varieties, $E(t)$ energy use and $\kappa(t)$ is the energy extraction cost, which is exogenous. We use an exogenous extraction cost as our focus is on autonomous endogenous technical change — we are interested in the response of technological change *given a path of prices*.[12]

At each t, final goods $Y(t)$ are produced using two intermediate goods: energy services $Y_E(t)$ and labor services $Y_L(t)$:

$$Y(t) := \left(\gamma_E Y_E(t)^{\frac{\epsilon-1}{\epsilon}} + \gamma_L Y_L(t)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}, \qquad \gamma_E, \gamma_L \in \mathbb{R}_+ \tag{3}$$

where ϵ is the elasticity of substitution and we assume $\epsilon \in (0, 1)$. Competitive firms produce intermediate goods using a continuum of machines indexed by i, where

[9] 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. Under a constant relative risk aversion (CRRA) utility function, the main results of our paper (Theorem 2.1, 1. and 4.) hold conditional on the assumption that any equilibrium path is an asymptotic balanced growth path (ABGP). However, we were unable to verify that every equilibrium path is an ABGP in the case of CRRA utility.

[10] Assets consist of equity in the innovating firms.

[11] The household maximization problem is standard and omitted, see for example, 8.2.2 in Acemoglu (2009); also note assets must satisfy 6) in Definition A.1

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

$i \in [0,1]$. The intermediate production functions are:

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

and

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

where $\alpha \in (0,1)$.

Let j index the sectors E and L. In the equations above, $q_j(i,t)$ denotes the highest quality of machine i at time t in sector j and $x_j(i,t \,|\, q_j(i,t))$ is the quantity of machine type i in sector j with quality $q_j(i,t)$. We assume labor supply is fixed with $L(t) = L$ and $L > 0$ for all t. The price of final output is normalized to one and all prices will be stated in terms of the final good price. We assume 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. Let $p_j^x(i,t|\, q)$ denote the price of machine i in sector j at time t with quality q. And let $p_j(t)$ denote the price of the intermediate goods $Y_j(t)$.

Turning to machine qualities, we assume a quality ladder for highest quality machines as follows:

(6) $$q_j(i,t) = \lambda^{n_j(i,t)} q_j(i,0), \qquad j \in \{E,L\}, i \in [0,1],\ t \in \mathbb{R}_+$$

where $q_j(i,0)$ is the quality of machine i at time 0 and $n_j(i,t)$ equals the random number of incremental innovations on the machine variety i up to time t. The arrival of a new innovation improves machine quality by a factor λ.

New entrants engage in research to improve machine varieties. 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 whole market for the variety — Schumpeterian creative destruction. The following assumption[13] ensures the firm with the highest quality machine can charge the unconstrained monopoly price:

Assumption 1 $\lambda \geqslant \left(\frac{1}{1-\alpha}\right)^{\frac{1-\alpha}{\alpha}}$.

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

[13]See Acemoglu (2009) Sections 12.3.3 and 14.1.2 for discussion.

R&D to improve the equality of machine i in sector j at time t generates a flow rate of machine improvement equal to $\eta_j b_j(i,t)$, where η_j is an exogenous parameter and:[14]

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

The denominator implies that higher quality machines are more difficult to improve upon, cancelling out state dependence embedded in the quality ladder.

Total R&D expenditure will be:

$$Z(t) := \int_0^1 z_E(i,t)\,\mathrm{d}i + \int_0^1 z_L(i,t)\,\mathrm{d}i \tag{7}$$

2.2. *Equilibrium characterization*

Objects in the model economy can be arranged in a tuple $\mathscr{E}$, where:

$$\mathscr{E} := \Big((\Omega,\mathscr{F},\mathbb{P}), q_E, q_L, p_E^x, p_L^x, p_E, p_L, x_L, x_E, z_E, z_L, C, X, E, Z, v_E, v_L, r, w, a\Big)$$

and

1. the tuple $(\Omega,\mathscr{F},\mathbb{P})$ is a probability space
2. the path of leading machine qualities, q_E and q_L, are $(\mathbb{R}_+^{[0,1]})^{\mathbb{R}_+}$ valued random variables defined on $(\Omega,\mathscr{F},\mathbb{P})$
3. the functions $p_E^x\colon [0,1]\times\mathbb{R}_+^2 \to \mathbb{R}_+$ and $p_L^x\colon [0,1]\times\mathbb{R}_+^2 \to \mathbb{R}_+$ are paths of monopolist prices
4. the functions $p_E\colon \mathbb{R}_+ \to \mathbb{R}_+$ and $p_L\colon \mathbb{R}_+ \to \mathbb{R}_+$ are intermediate goods prices
5. the functions $x_E\colon [0,1]\times\mathbb{R}_+^2 \to \mathbb{R}_+$ and $x_L\colon [0,1]\times\mathbb{R}_+^2 \to \mathbb{R}_+$ are machine quantities
6. the functions $z_E\colon [0,1]\times\mathbb{R}_+ \to \mathbb{R}_+$ and, $z_L\colon [0,1]\times\mathbb{R}_+ \to \mathbb{R}_+$ are paths for R&D
7. paths for consumption, total expenditure on machines, energy and total R&D that satisfy $C, X, E, Z \in \mathscr{C}^1(\mathbb{R}_+)$ where $\mathscr{C}^1(\mathbb{R}_+)$ is the space of all real valued continuously differentiable functions on $\mathbb{R}_+$
8. the functions $v_E\colon [0,1]\times\mathbb{R}_+^2 \to \mathbb{R}_+$ and $v_L\colon [0,1]\times\mathbb{R}_+^2 \to \mathbb{R}_+$ are paths for the value functions

---

[14]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,

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

9. the functions $r\colon \mathbb{R}_+ \to \mathbb{R}_+$, $w\colon \mathbb{R}_+ \to \mathbb{R}_+$ and $a\colon \mathbb{R}_+ \to \mathbb{R}_+$ are paths for the interest rate, wage rate, and assets.

Section A.1 of the Appendix gives a definition of equilibrium; we now turn to documenting the familiar static equilibrium conditions, given technology.

Profit maximization by final goods producers gives:

$$\gamma_E Y(t)^{\frac{1}{\epsilon}} Y_E(t)^{-\frac{1}{\epsilon}} = p_E(t), \qquad \gamma_L Y(t)^{\frac{1}{\epsilon}} Y_L(t)^{-\frac{1}{\epsilon}} = p_L(t) \tag{8}$$

At each t, profit maximization by the intermediate goods producers gives machine demands for the highest quality machine of type i:

$$x_E(i,t \,|\, q_E(i,t)) = p_E(t)^{\frac{1}{\alpha}} p^x_E(i,t \,|\, q_E(i,t))^{-\frac{1}{\alpha}} q_E(i,t) E(t) \tag{9}$$

$$x_L(i,t \,|\, q_L(i,t)) = p_L(t)^{\frac{1}{\alpha}} p^x_L(i,t \,|\, q_L(i,t))^{-\frac{1}{\alpha}} q_L(i,t) L \tag{10}$$

and first order conditions for energy and labor:

$$\alpha(1-\alpha)^{-1} p_E(t) E(t)^{\alpha-1} \int_0^1 q_E(i,t)^{\alpha} x_E\big(i,t \,|\, q_E(i,t)\big)^{1-\alpha} \,\mathrm{d}i = \kappa(t) \tag{11}$$

$$\alpha(1-\alpha)^{-1} p_L(t) L^{\alpha-1} \int_0^1 q_L(i,t)^{\alpha} x_L\big(i,t \,|\, q_L(i,t)\big)^{1-\alpha} \,\mathrm{d}i = w_L(t) \tag{12}$$

Since the cost of machine varieties is $1-\alpha$ and, by Assumption 1, monopolists who own the highest quality machine production technologies set $p^x_j(i,t \,|\, q_j(i,t)) = 1$, machine demands for the highest quality machine of type i are then:

$$x_E\big(i,t \,|\, q_E(i,t)\big) = p_E(t)^{\frac{1}{\alpha}} q_E(i,t) E(t), \quad x_L(i,t \,|\, q_L(i,t)) = p_L(t)^{\frac{1}{\alpha}} q_L(i,t) L \tag{13}$$

Profits for monopolists who own the leading edge machines are:

$$\pi_E\big(i,t \,|\, q_E(i,t)\big) = \alpha p_E(t)^{\frac{1}{\alpha}} q_E(i,t) E(t), \quad \pi_L(i,t \,|\, q_L(i,t)) = \alpha p_L(t)^{\frac{1}{\alpha}} q_L(i,t) L \tag{14}$$

Owners of lower quality machine production technologies receive zero profits.

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 leading machines. Intermediate output, using (4) and (5) with (13) becomes:

$$Y_E(t) = (1-\alpha)^{-1} E(t) p_E(t)^{\frac{1-\alpha}{\alpha}} Q_E(t), \quad Y_L(t) = (1-\alpha)^{-1} L p_L(t)^{\frac{1-\alpha}{\alpha}} Q_L(t) \tag{15}$$

Next, using the intermediate demand conditions, (8), we can write the ratio of prices as

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

where $\gamma = \frac{\gamma_E}{\gamma_L}$. And solving for the ratio of prices gives:

$$\text{(16)} \qquad \frac{p_E(t)}{p_L(t)} = \gamma^{\frac{\alpha\epsilon}{\sigma}} \left(\frac{E(t)Q_E(t)}{LQ_L(t)}\right)^{-\frac{\alpha}{\sigma}}$$

where $\sigma = 1 + \alpha(\epsilon - 1)$ is the elasticity of substitution between the factors of production. Using (15), we can also derive a ratio of intermediate goods:

$$\text{(17)} \qquad \frac{Y_E(t)}{Y_L(t)} = \left(\frac{E(t)Q_E(t)}{LQ_L(t)}\right)^{\frac{\epsilon\alpha}{\sigma}} \gamma^{\frac{\epsilon(1-\alpha)}{\sigma}}$$

Next, insert (8) into (15), use the expression for final output [15] and the ratio of intermediate goods above (17) to derive:

$$\text{(18)} \qquad Y(t) = (1-\alpha)^{-1} \left[\gamma_E^{\frac{\epsilon}{\sigma}} \left(E(t)Q_E(t)\right)^{\frac{\sigma-1}{\sigma}} + \gamma_L^{\frac{\epsilon}{\sigma}} \left(LQ_L(t)\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

We now derive a condition characterizing energy intensity. Use (11) and the definition of $Y_E(t)$ to arrive at:

$$\text{(19)} \qquad \alpha\theta_E(t)\frac{Y(t)}{E(t)} = \kappa(t)$$

where

$$\text{(20)} \qquad \theta_E(t) := \frac{p_E(t)Y_E(t)}{Y(t)} = \left(1 + \gamma^{-\frac{\epsilon}{\sigma}} \left(\frac{E(t)Q_E(t)}{Q_L(t)L}\right)^{\frac{1-\sigma}{\sigma}}\right)^{-1}$$

is the cost share of energy services in final production. The second equality results from using (8) and (17). Use (20) along with (18) to write (19) as:

$$\text{(21)} \qquad \gamma_E^{\frac{\epsilon}{\sigma(\sigma-1)}} \alpha(1-\alpha)^{\frac{\sigma-1}{\sigma}} \left(\frac{Y(t)}{E(t)}\right)^{\frac{1}{\sigma}} Q_E(t)^{\frac{\sigma-1}{\sigma}} = \kappa(t)$$

Equation (21) tells us that energy-augmenting technical change is energy saving.

Now we turn to the dynamic aspects of equilibrium concerning technology choice.

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

Since consumers are risk neutral, the interest rate in the economy will be ρ for all t. 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:[16]

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

where $T_j(i, t)$ is the random stopping time after which a new innovation replaces the incumbent. Free entry and exit implies:[17]

$$\eta_j \frac{v_j(i, t \mid \lambda q)}{q} \leqslant 1, \qquad j \in \{E, L\}, \quad t \in \mathbb{R}_+ \tag{22}$$

holds for all t and $i \in [0, 1]$ and the inequality is an equality if $z_j(i, t) > 0$. Assuming the value function is differentiable implies $\dot{v}_j(i, t \mid q) = 0$. The Hamilton Jacobi Bellman (HJB) equation[18] for $v_j(t, i \mid q)$ is:

$$\dot{v}_j(i, t \mid q) = \big(\rho + \eta_j b_j(i, t)\big)\, v_j(i, t \mid q) - \pi_j(i, t \mid q) \tag{23}$$

Immediately giving, using (22) and (14):

$$\eta_E b_E(i, t) + \rho \geqslant \eta_E \lambda \alpha p_E(t)^{\frac{1}{\alpha}} E(t), \qquad \eta_L b_L(i, t) + \rho \geqslant \eta_L \lambda \alpha p_L(t)^{\frac{1}{\alpha}} L \tag{24}$$

where, once again, the inequality is an equality of $z_j(i, t) > 0$. Note the equation above implies $b_j(i, t)$ is no longer conditioned on the machine variety — there is a common rate of innovation for all machines in a sector. Let $b_j(t) := \int_0^1 b_j(i, t)\, \mathrm{d}i$ for $j \in \{E, L\}$.

Finally, for each i, $q_j(i, t)$ is a random process. However, at each t, the average machine quality $Q_j(t)$ for each sector j will be deterministic,[19] determined by the

---

[16]The random variable $T_j(i, t)$ has a distribution:

$$\mathbb{P}(T_j(i, t) \geqslant t + s) = e^{-\int_t^s \eta_j b_j(i, s_1)\, \mathrm{d}s_1}$$

[17]Recall that expenditure of $z_j(i, t)$ units of final good on R&D generates a flow rate of $\eta_j \frac{z_j(i,t)}{q_j(i,t)}$. Since the price of the final good is normalized to one, the value of spending one unit of the final good on research is $\frac{\eta_j v_j(i,t \mid \lambda q)}{q(i,t)} - 1$, which should not be strictly positive.

[18]See Acemoglu (2009), Equation (14.13).

[19]The average quality will equal $\mathbb{E}q(i, t)$ with probability one (Sun and Zhang, 2009), this is referred to as no aggregate uncertainty (NAU) and in this paper we view NAU as an assumption

innovation rate b_j as follows:[20]

$$\dot{Q}_j(t) = (\lambda - 1)\eta_j b_j(t) Q_j(t), \qquad j \in \{E, L\} \tag{25}$$

2.3. *Main theoretical results*

An equilibrium growth path is a balanced growth path if output and consumption both grow at a constant rate g, energy use 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, $\varphi(t)$, we will use $\hat{\varphi}(t)$ to refer to the growth rate of $\varphi(t)$ at time t, that is $\hat{\varphi}(t) = \frac{\dot{\varphi}(t)}{\varphi(t)}$, and we use $\varphi(t) \to \bar{\varphi}(t)$ to mean $\lim_{t\to\infty} \varphi(t) = \bar{\varphi}(t)$ for a path $\bar{\varphi}(t)$.[21]

PROPOSITION 2.1 *Let an economy $\mathscr{E}$ be an equilibrium. If $\hat{Q}_E(t) > \iota$ for all t, where $\iota > 0$, 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 (18), $Y(t) = \tilde{F}(Q_E(t)E(t), Q_L(t)L)$ 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 (20) implies $\theta_E(t)$ is constant, which in turn implies $\hat{Q}_E(t) = 0$ and $E(t)/Y(t)$ is constant by (21) and (37). *Q.E.D.*

Thus a BGP cannot feature autonomous energy intensity improvements. We study the possibility of autonomous energy intensity improvements along *asymptotic* balanced growth paths (ABGP).[22]

DEFINITION 2.1 An equilibrium economy $\mathscr{E}$ is an asymptotic balanced growth path (ABGP) if $\hat{Y}(t) \to g$, $\hat{C}(t) \to g$, $\hat{E}(t) \to g_E$, $\hat{Q}_E(t) \to g_{Q_E}$, $\hat{Q}_L(t) \to g_{Q_L}$, $b_E(t) \to b_E^\star$ and $b_L(t) \to b_L^\star$, where all limits are real valued.

ASSUMPTION 2 *The economy $\mathscr{E}$ satisfies:*

$$\lambda \eta_L \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} (1 - \theta_E(0))^{\frac{1}{1-\sigma}} L > \rho > (\lambda - 1)\eta_L \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} L \tag{26}$$

---

without providing a formal proof.

[20] Acemoglu and Cao (2015) (Footnote 23) provide a concise derivation of Equation (25).

[21] We say $\varphi(t) \to \bar{\varphi}(t)$ if and only if for every $\epsilon > 0$, there exists T such that for all $t > T$, $|\varphi(t) - \bar{\varphi}(t)| < \epsilon$.

[22] 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 autonomous intensity improvements while ABGPs can.

Note by (20), we have $\alpha p_L(t)^{\frac{1}{\alpha}} L = \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} (1-\theta_E(t))^{\frac{1}{1-\sigma}} L$ for each t. Thus, by the free entry and exit condition and (24), the first inequality of the assumption above ensures the growth rate of labor-augmenting technologies and output is positive. The second inequality ensures corporate assets do not grow faster than the discount rate, ensuring the transversality condition holds and (1) remains finite.

We now show that an ABGP features autonomous energy intensity improvements.

THEOREM 2.1 *If $\mathscr{E}$ is an ABGP and satisfies Assumptions 1 and 2 and $\dot{\kappa}(t) = 0$ for all t, then:*

1. $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$
2. $\theta_E(t) \to 0$
3. $\hat{Y}(t) \to \hat{Q}_L(t) \to g$ *with* $g = (\lambda - 1)(\lambda \eta_L \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} L - \rho)$
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. tells us the growth rate of energy-augmenting technologies converges to a rate lower than the growth rate of labor-augmenting technologies, since we have assumed $\sigma < 1$. Part 2. tells us the cost share of energy converges to zero. Part 3. tells us the growth of output converges to the growth rate of labor-augmenting technologies, because the cost share of labor converges to 100% of output (Part 2.).

Part 4. tells us energy intensity declines, but at a slower rate than output growth. Energy intensity declines because of energy-augmenting technical change (recall Equation (21)). And the rate of decline is slower than the rate of output growth since energy-augmenting technologies grow slower than labor-augmenting technologies (because of the price effect discussed at Equation (34)). On the other hand, labor services become an increasingly important contributor to output as the share of energy services falls to zero; thus output growth converges to the rate of growth of labor-augmenting technologies (see decomposition of output growth at Equation (44) in the Appendix).

The following result is an implication of Theorem 2.1, Part 2., and says that if initially there are no energy-augmenting technical advances (and energy intensity does not fall), then there exists some time T after which energy-augmenting technologies will begin to advance.

COROLLARY 2.1 *Let $\mathscr{E}$ be an ABGP. If $\dot{\kappa} = 0$ and:*

$$\lambda \eta_E \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(0)^{\frac{1}{1-\sigma}} E(0) - \rho < 0 \tag{27}$$

then there exists $T \geqslant 0$ such that for all $t > T$:

$$\eta_E b_E(t) = \eta_E \alpha \lambda \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{1}{1-\sigma}} E(t) - \rho > 0$$

Figure 5 shows ABGPs (in bold), assuming $\kappa(t)$ remains constant. To understand the direction of ABGPs intuitively, note output is growing and $b_E(t) > 0$, hence energy intensity must be falling and we verified in Theorem 2.1 that the elasticity of E/Y with respect to Y must converge to $\frac{\sigma-1}{2-\sigma}$.

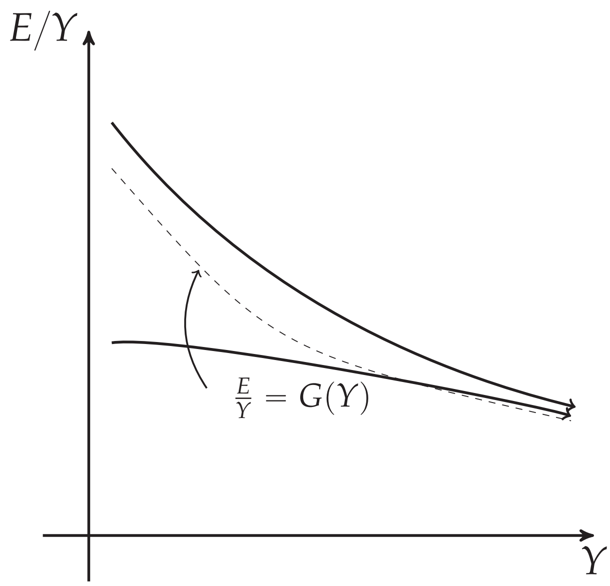

FIGURE 5.— ABGPs for an economy with no state dependence. ABGPs are shown in bold, converging to "stable path" (the dotted line).

Below, we show that, for a given level of output, energy intensity converges for all ABGPs (conditional convergence).

PROPOSITION 2.2 *Let $\dot{\kappa}(t) = 0$ for all t. If $\theta(0) + \sigma < 1$ and $\mathscr{E}$ is an ABGP, then:*

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

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

The Appendix, Section A.2.1, studies the transitional dynamics of ABGPs in detail, in particular why ABGPs starting below $G(Y)$ in Figure 5 must cross $G(Y)$ as they converge. The Appendix also shows that any equilibrium path is an ABGP.

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.

CLAIM 2.1 Let $\dot{\kappa}(t) = 0$ for all t. If for some t, $\tilde{Y}(t) = Y(t)$ 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. $\frac{\hat{\tilde{E}}(t) - \hat{\tilde{Y}}(t)}{\hat{\tilde{Y}}(t)} < \frac{\hat{E}(t) - \hat{Y}(t)}{\hat{Y}(t)}$

The Appendix, Section A.2.1, gives a proof of this claim.

2.4. *Extreme state dependence economy*

The main difference between the extreme state dependence and no state dependence economy detailed above are innovation incentives. Assume a constant unit measure of scientists who undertake research directed towards energy or labor machine improvements. If a scientist works to improve the quality 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 $s_j \eta_j$.[23]

To characterize incentives to innovate under state dependence, let:

$$V_j(t) := \int_0^1 v_j(i, t \mid q_j(i, t)) \, \mathrm{d}i \,, \qquad j \in \{E, L\}\,, t \in \mathbb{R}_+ \tag{28}$$

and let

$$\Pi_j(t) := \int_0^1 \pi_j(i, t \mid q_j(i, t)) \, \mathrm{d}i \,, \qquad j \in \{E, L\}\,, t \in \mathbb{R}_+ \tag{29}$$

[23]The random allocation of scientists across machine varieties ensures all varieties experience innovation in equilibrium. Acemoglu et al. (2012) make a similar assumption in a discrete time model.

The free entry and exit conditions under state dependence are:

$$\begin{aligned} V_E(t) &\geqslant V_L(t), \quad \text{if} \quad s_E > 0 \\ V_L(t) &\geqslant V_E(t), \quad \text{if} \quad s_L > 0 \end{aligned} \tag{30}$$

The definition of an economy and equilibrium are identical to the no state dependence case, with b_j replaced by s_j and the free entry and exit conditions specified by (30) instead of (22). We use $\mathscr{E}_S$ to denote an extreme state dependence economy.

The first state dependence result tells us that sustained autonomous energy intensity improvements are not possible.

THEOREM 2.2 *If $\mathscr{E}_S$ is an ABGP and $\dot{\kappa}(t) = 0$, then $\hat{Y}(t) \to \hat{Q}_L(t) \to g$ where $g = (\lambda - 1)\eta_L$, $\theta_E(t) \to \theta_E^\star$ where $\theta_E^\star > 0$ and $\hat{E}(t) - \hat{Y}(t) \to 0$.*

Moreover, if the energy extraction costs are low enough or energy-augmenting technologies are advanced enough, then there exists a BGP with no energy-augmenting research.

PROPOSITION 2.3 *Let $\mathscr{E}_S$ be an equilibrium. If $\dot{\kappa}(t) = 0$ for all t, then 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:*

$$\left(\frac{\kappa}{Q_E(0)}\right)^{1-\sigma} \leqslant \frac{(\alpha(1-\alpha))^{1-\sigma}\,\gamma^{\frac{\epsilon}{1-\sigma}}}{\frac{\rho+\eta_L}{\rho-\eta_L}+1}, \qquad t \geqslant 0 \tag{31}$$

Combining (37) and (19) with (31) implies:

$$\frac{E(0)}{Y(0)} \leqslant \frac{\alpha\kappa^{-1}\gamma_L^{\frac{\epsilon}{\sigma-1}}}{\frac{\rho+\eta_L}{\rho-\eta_L}+1} := \phi \tag{32}$$

must hold at time 0 for for a BGP to exist. Noting (36), if Q_E is constant along the BGP, then θ_E remains constant, and thus, by (19), $\frac{E}{Y}$ remains constant and satisfies (32) along the BGP.

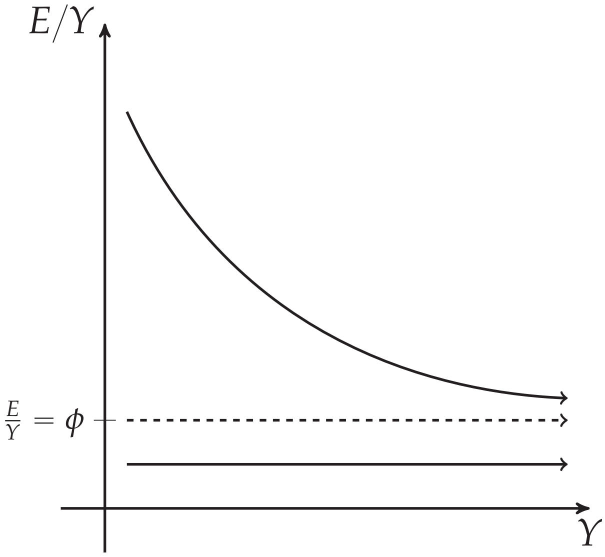

FIGURE 6.— ABGPs for an economy with extreme state dependence. ABGPs are shown in bold. The term ϕ is defined by (32).

Turning now to convergence under state dependence, consider Figure 6, which shows ABGPs for a state dependence economy with a constant real energy price. The dotted line is the BGP where $\frac{E}{Y} = \phi$, where ϕ is implied by Equation (32). By Proposition 2.3, any economy with energy intensity below ϕ has a unique BGP with no energy-augmenting technical change. However, again by Proposition 2.3, if energy intensity is strictly greater than ϕ, energy-augmenting technologies must advance. And by Theorem 2.2, Q_E and hence energy intensity must converge to a constant.

2.5. *Discussion of main results*

2.5.1. *Why energy-augmenting technologies advance more slowly than labor-augmenting technologies and stop advancing under extreme state dependence*

First, let us see why as an economy grows and energy-augmenting technologies advance, there are continued incentives to undertake energy-augmenting research in the no state dependence economy. Incentives to innovate (Equation (24)) are the product of the price effect and market size effect:

$$\eta_E b_E(t) = \eta_E \lambda \overbrace{\alpha p_E(t)^{\frac{1}{\alpha}}}^{\text{Price effect}} \underbrace{E(t)}_{\text{Market size effect}} - \rho \tag{33}$$

and, by Equation (20) and noting (61) in the Appendix, $\alpha p_E(t)^{\frac{1}{\alpha}} = \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{1}{1-\sigma}}$. Thus along an ABGP, if the price of energy, κ is constant and energy use expands, the price of energy services falls.[24] At the same time, the market size of E increases, sustaining incentives to innovate.

However, the relative incentive to innovate favors labor-augmenting research. With innovation in both sectors, by (24),

$$\frac{\rho + \eta_E b_E(t)}{\rho + \eta_L b_L(t)} = \frac{\eta_E}{\eta_L} \overbrace{\left(\frac{p_E(t)}{p_L(t)}\right)^{\frac{1}{\alpha}}}^{\text{Price effect}} \underbrace{\frac{E(t)}{L}}_{\text{Market size effect}} = \frac{\eta_E}{\eta_L} \left(\frac{Q_E(t)}{Q_L(t)}\right)^{-\frac{1}{\sigma}} \left(\frac{E(t)}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{34}$$

As both technologies advance, energy use grows,[25] the market size effect increases profitability for energy-augmenting research and the price effect decreases profitability. If $\sigma < 1$, then the price effect is stronger leading to an overall fall in the profitability of energy-augmenting research.

The fall in the relative profitability of energy-augmenting research means more R&D resources are allocated to labor-augmenting research and the growth rate of labor-augmenting technologies will be faster. When there is no state dependence, the fall in the ratio $\frac{Q_E(t)}{Q_L(t)}$ is sufficient to eventually balance relative profits, so the left hand side of (34) converges to a constant and innovation occurs in both sectors.

The exact relative rates of innovation under a rational expectations equilibrium in the case with state dependence are more difficult to characterize using current profit ratios. Nonetheless, we now consider the current profit ratios to gain intuition about the result we proved in Section 2.4. Suppose that for innovation to occur

---

[24]Fouquet (2008) shows that the price of energy services fell dramatically over the last few centuries in Britain.

[25]In this statement, we presume labor-augmenting technologies grow at least as fast as energy-augmenting technologies. Then by the first order condition for energy use, Equation (21), energy use expands. However, if output grows sufficiently slower than energy-augmenting technologies advance, energy use can fall. In Appendix B, we give the exact conditions under which energy use may fall when labor-augmenting technologies do not advance.

in both sectors, we must have

$$\frac{\rho + \eta_E s_E(t)}{\rho + \eta_L s_L(t)} = \frac{\eta_E \Pi_E(t)}{\eta_L \Pi_L(t)}$$

(35)
$$= \frac{\eta_E}{\eta_L} \overbrace{\left(\frac{p_E(t)}{p_L(t)}\right)^{\frac{1}{\alpha}}}^{\text{Price effect}} \underbrace{\frac{E(t)}{L}}_{\text{Market size effect}} \overbrace{\frac{Q_E(t)}{Q_L(t)}}^{\text{Scale effect}} = \frac{\eta_E}{\eta_L} \left(\frac{Q_E(t)}{Q_L(t)}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{E(t)}{L}\right)^{\frac{\sigma-1}{\sigma}}$$

As energy use grows, once again, the price effect leads to a fall in relative profitability to do energy research. However, now, as $\frac{Q_E(t)}{Q_L(t)}$ falls, the scale effect further increases the profitability of labor-augmenting 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 the fall in the technology ratio. Use (20) to write:

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

Now use (19) and (21) to arrive at

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

Thus, if $Q_E(t) \to \infty$, then we must have $\frac{\rho+\eta_E s_E(t)}{\rho+\eta_L s_L(t)} \to 0$, which cannot hold since $s_L \leqslant s$. Continued energy-augmenting technological advancement cannot now occur along any ABGP since incentives to invest in labor-augmenting technologies increase at such high a rate due to state dependence increasingly favoring labor-augmenting technologies, that investment in energy-augmenting technologies must keep decreasing for the no arbitrage condition to hold.

2.5.2. *The possibility of limited state dependence and population growth*

Aghion et al. (2016) find evidence of state dependence among firms in the auto industry who undertake 'clean' versus 'dirty' innovation. However, state dependence can be of a limited/ non-extreme form where the elasticity of the production of new innovations with respect to existing knowledge is less than one. Indeed, Aghion et al. (2016) find that the elasticity of the creation of new knowledge with respect to the existing level of knowledge is approximately .3. We briefly discuss the implications of such a specification. Suppose an innovation possibilities frontier of the form:

(38) $$\dot{Q}_j(t) = (\lambda - 1)\eta_j s_j(t) Q_j(t)^{\psi}$$

where $\psi \in (0,1)$. Such a formulation requires that s_j grows through population growth for sustained per capita growth to be possible, similar to Jones (1995). Such a model would imply a relative current profit ratio of the form:

$$\begin{aligned}\frac{\rho + \eta_E s_E(t)}{\rho + \eta_L s_L(t)} &= \frac{\eta_E}{\eta_L}\frac{\Pi_E(t)}{\Pi_L(t)} \\ &= \frac{\eta_E}{\eta_L}\left(\frac{p_E(t)}{p_L(t)}\right)^{\frac{1}{\alpha}}\frac{E(t)}{L}\left(\frac{Q_E(t)}{Q_L(t)}\right)^{\psi} = \frac{\eta_E}{\eta_L}\left(\frac{Q_E(t)E(t)}{Q_L(t)L(t)}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{Q_E(t)}{Q_L(t)}\right)^{\psi-1}\end{aligned}$$

When $Q_E(t) \to \infty$, $\theta_E(t) \to 0$ and $\left(\frac{Q_E(t)E(t)}{Q_L(t)L(t)}\right)^{\frac{\sigma-1}{\sigma}} \to 0$. And since $\left(\frac{Q_E(t)}{Q_L(t)}\right)^{\psi-1} \to \infty$, an ABGP where the relative profitability of labor and energy-augmenting technologies remain constant is possible; thus, under limited state dependence, energy intensity can continue to decline as in the case with no state dependence. [26]

3. EXTREME STATE DEPENDENCE OR NOT?

In this section, we revisit the observed energy intensity trends and discuss whether they are more consistent with extreme state dependence or non-extreme state dependence (either no state dependence or limited state dependence). An initial observation of the data does not conclusively suggest one scenario or another. Both the no state dependence and complete state dependence economy can be consistent with paths of energy intensity across countries through time. First, both economies experience convergence of energy intensities across countries given the level of output — consistent with conditional convergence shown by Csereklyei et al. (2016) and the increasingly 'tighter' relationship between output per capita and energy intensity seen in Figure 2 and Figure E.8 (in the appendix). Second, both models can also explain why many countries experience no declines in energy intensity. Figure 2 shows that some countries, such as Italy and Spain, with low energy intensities have close to constant paths of energy intensity. Under limited or no state dependence, initially, energy intensity may be so low given output such that (27) holds and no energy-augmenting innovation occurs. On the other

[26] In a model with population growth where scientists engage in R&D, using a similar process to the proof for Theorem 2.1, we can show $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{2-\psi}{1-\psi-\sigma}$ along an ABGP if we assume relative profitability is given by (35). However, we were unable to show that a rational expectations equilibrium implied ABGPs in this case. Thus our study in this paper focused on the lab equipment model with no population growth.

hand, under extreme state dependence, countries showing constant energy intensity through time may be on their long-run BGP, to which all countries will converge to.

Now consider the cross-sectional distribution of energy intensity over per capita GDP at a given time, shown by Figure 3. We see that energy intensity decreases across per capita GDP at a slower rate as per capita output increases and does not fall at all across higher income countries. Both these facts are inconsistent with the cross-section of countries being distributed along the "stable path" of the no state dependence scenario — the dotted line in Figure 5 (see the Appendix, Section D.2, for details and an analysis of the cross-sectional data). However, a non-extreme state dependence scenario may still explain the data if the cross-section of countries are not distributed along the dotted line of Figure 5. The purpose of our econometric analysis, which we now turn to, is to understand whether this could be the case.

3.1. *Econometric analysis*

In our econometric model, we wish to understand whether, controlling for fluctuations in energy prices, Figure 5 or 6 better represents the relationship between energy intensity growth rate and the level of per capita output and energy intensity between 1970 and 2010. To motivate the model, consider the following relationship in the no state dependence model:

$$(39) \qquad \hat{E} - \hat{Y} = (\sigma - 1)\hat{Q}_E - \sigma\hat{\kappa} = \max\left\{\Xi \left(\frac{E}{Y}\right)^{\frac{2-\sigma}{1-\sigma}} \kappa^{\frac{1}{1-\sigma}} Y - \rho, 0\right\} - \sigma\hat{\kappa}$$

where Ξ is a constant. To derive the equation, take growth rates across Equation (21), note (25) and consider the free entry and exit condition at (24) and then use Equations (53) and (19). The equation above tells us that the growth rate of energy intensity is a non-linear function of per capita output, energy intensity, the rate of change of the energy price and the energy price as follows. Since we cannot derive such an explicit relationship in the model with state dependence and since the models here are only a rough approximation of reality, we proceed to approximate the non-linear relationship between the energy intensity growth rate, per capita output and energy intensity by estimating d degree multivariate polynomials as follows:

$$(40) \qquad \%\Delta\left(EY_t^i\right) = \sum_{k=0}^{d}\sum_{l=0}^{d}\sum_{m=0}^{d}\sum_{n=0}^{d} \alpha_{k,l,m,n} \ln(Y_t^i)^k \ln(EY_t^i)^l \left(\%\Delta P_t\right)^m \left(\%\Delta Y_t^i\right)^n + \epsilon_t^i$$

We let t index 5 year periods between 1970 and 2010. For each country i at time t, $\%\Delta\left(EY_t^i\right)$ denotes the average growth rate of energy intensity over the 5 year period, Y_t^i denotes the real GDP per capita at the beginning of the period, $\%\Delta P_t$ denotes the growth rate of the real U.S. energy price over the period and $\%\Delta Y_t^i$ denotes the growth rate of per capita output over the period. We have included the rate of output growth in the regression equation since shocks to energy intensity and the level of output are likely correlated with the output growth rate. The term ln refers to the natural log and we use logs of the independent variables in the regression equation so we can interpret the slopes of vectors in the predicted stream plot below as elasticities. The $\alpha_{k,l,m,n}$ terms are coefficients we estimate through ordinary least squares using data 100 countries between 1970 and 2010 (Csereklyei et al. (2016) describe the sources of the data).

We approximate polynomials up to degree 7 and find a polynomial of degree 4 scores lowest on the Akaike Information Criteria (AIC). Using the estimated degree 4 polynomial, we generate predicted values of ΔEY, setting the energy price growth to zero and the per capita output growth rate to 2% (the mean output growth rate of the sample). For values of energy intensity and per capita GDP observed in the data, the predicted values of ΔEY are then plotted as the slope of vectors in the first panel of Figure 7. Observations of countries over 5 year periods are also plotted. The vectors give the predicted direction of movement of a country with a given energy intensity and per capita GDP. The predicted vectors are consistent with no extreme state dependence; countries converge to a path which has an increasing rate of decline of energy intensity (the stable path is concave, as implied by Figure 5). For instance, countries with initially low energy intensity may have a close to zero or very low rate of decline of energy intensity as implied by Equation (39). And as Equation (39) also implies, once these countries are rich enough, they experience a higher rate of decline in energy intensity. Moreover, in the first stream-plot of Figure 7, the elasticity between the rate of decline of energy intensity and the rate of output growth converges for all growth paths, as implied by Theorem 2.1, part.4.

To test whether the decline in energy intensity towards the bottom right area of the scatter plot of Figure 7 is significantly below zero, we give a stream-plot of the 90% upper confidence interval of the predicted decline in energy intensity vectors in Figure 7. An upward sloping vector in the second panel implies the *expected* energy intensity decline for a country at a given level of GDP per capita and energy intensity is not significantly different to zero. However, for rich countries with low energy intensity, the 90% vectors slope downwards, allowing us to reject the null that their average rate of decline is zero.

Finally, the broader arguments in favor of a scenario without extreme state dependence are also substantial. For example, the scale effects generated by a model with

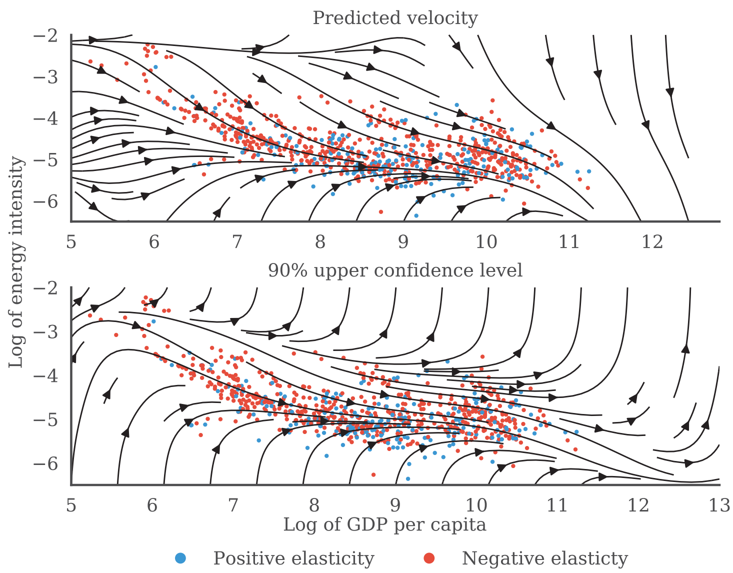

FIGURE 7.— Stream plot gives predicted value and upper confidence interval of decline of energy intensity assuming 2% per year growth of per capita GDP and no change in the real U.S. energy price. Scatter plot gives observed values of energy intensity and real GDP per capita for countries for 5 year periods between 1970 and 2010.

extreme state dependence and population growth lead to increasing per capita growth rates not observed in the data (see (Jones, 1995) and recent empirical evidence by Kruse-Andersen (2017) and Bloom et al. (2017)). Moreover, micro level evidence by Aghion et al. (2016) suggests that innovation in alternative energy motor vehicles, such as hybrid and electric cars, is characterized by limited state dependence.

4. STRUCTURAL CHANGE

A common view is that, rather than energy-augmenting technical change, structural change towards labor intensive service sectors has been responsible for driving down the energy intensity of output. Notably, Sue Wing (2008) shows most of the energy intensity improvement of output since 1950 in the U.S. has been due to shifts in the relative share of industries, rather than within sector energy efficiency improvements. It is also true that the cost share of service industries has been increasing for the past 50 years. By contrast, the directed technical change theory we present tells a story where the cost share of energy services falls but the *quantity share of energy services relative to labor services increases*; the driver of energy efficiency improvements then becomes energy-augmenting technical change (energy efficiency of energy intensive goods). To see this, recall (see discussion before equation (17)) the ratio of intermediate goods can be written as:

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

And then note equation part 2 of Theorem 2.1 implies $\theta(t) \to 0$, which by equation (36) implies the right hand side of the above equation converges to zero. In the theory presented here, the cost share of labor services does grow as the economy grows, but the structural change results from the increase in the value of labor services, not the quantity of labor services. Now decompose the growth rate of output as follows, into the growth rate of the energy intensive goods, minus the growth rate of labor intensive goods and minus the growth rate of energy-augmenting technological change:

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

We can see that if the quantity of energy intensive goods grows faster than the quantity of labor intensive goods, the left hand side of the above equation can only be negative if the growth rate of energy-augmenting technical change is positive.

Gentvilaite et al. (2015) and Kander et al. (2015) argue a similar view on the nature of structural change: "while deindustrialization appears to be a social fact, from the viewpoint of real output and the associated environmental burdens, the transition to the service economy is largely a price illusion"(Gentvilaite et al., 2015). A natural next step is to check whether the quantity ratio of labor services to energy services has been increasing or decreasing in the data; we perform this analysis by first estimating a time series of $Q_E(t)$ and $Q_L(t)$ for the U.S. data from 1900 (see Appendix D.1 for details). Once we have the time series of $Q_E(t)$ and $Q_L(t)$, we can directly calculate a time series for the ratio $\frac{Y_E(t)}{Y_L(t)}$ by using U.S. data on energy use and employment (see appendix C.1) and noting equation (17). We plot the results of the time series of $\frac{Y_E(t)}{Y_L(t)}$ in figure 8. The data show a clear decreasing trend for the quantity of labor to energy services up to 1970. After the 1970s, there is no increasing trend; the ratio increases, falls and then increases, plausibly due to a substitution away from energy intensive goods in response to price shocks. This view of the data lends support to the theory we presented in this paper – the quantity of energy intensive goods used has been increasing and energy-augmenting technical change plays the key role in observed energy efficiency improvements.

How can we reconcile our analysis with studies, most notably Sue Wing (2008), arguing that structural change is responsible for improvements in energy inten-

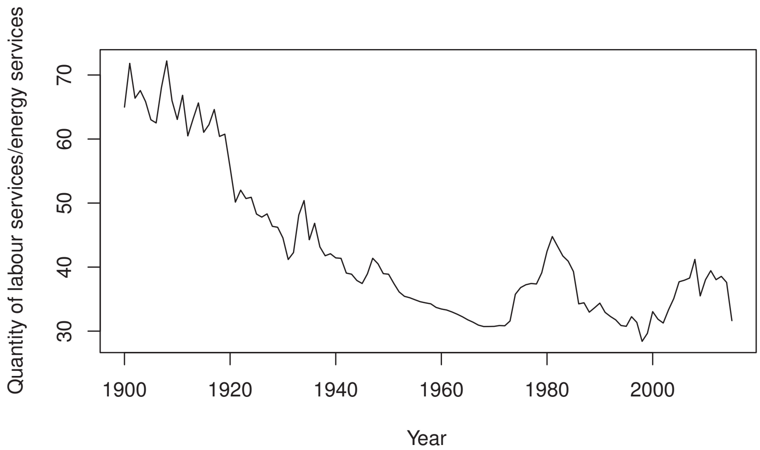

FIGURE 8.— Quantity share of labor to energy services in the U.S. We assume $\sigma = .5$ and $\alpha = .66$, differing values of σ below one do not alter the interpretation of the results.

sity? Sue Wing's analysis consists of 45 sectors and these sectors do not directly map to labor intensive or energy intensive sectors. At least in a supply side model such as ours, shifts in shares of industries may still be still driven by technological change. For example, supposing that gas utilities are more energy efficient than electric utilities, then a rise in the share of gas utilities can be seen as an improvement in a technology that augments gas utilities. Now, in an aggregate model such as ours, improvements in the aggregate energy-augmenting technology Q_E can be thought to incorporate structural change driven by improvements in individual sector technologies, technologies that augment industries as such gas utilities, for instance. (Another example is the the growth of information and communication technology sectors which can also be seen as energy saving, see Gentvilaite et al. (2015).) The key point is that Sue Wing's analysis does not tell us whether the type of structural change driving down aggregate energy intensity is a change towards labor services or a shift in the mix of energy sectors providing energy services.

5. CONCLUSION

This paper studied the relationship between technological change and a key factor input, energy, and asked why energy intensity has fallen at a rate *slower than output growth* without a commensurate upward trend in the real energy price. Our analysis argued that energy-augmenting innovation can occur along a growth path, but firms still face stronger incentives to augment labor rather than energy because the relative price of labor to energy has been increasing and labor is relatively scarce. We also argued that it was energy-augmenting innovation along a growth path that leads to a fall in energy intensity in the absence of increases in the energy price, not, as is popularly viewed, structural change towards labor intensive service sectors. And since output growth becomes dominated by the growth of labor-augmenting technologies, the stronger incentives for labor-augmenting research means energy intensity falls slower than the rate of output growth.

Whether or not energy-augmenting research continues along a growth path depends on whether or not there is extreme state dependence in innovation. Extreme state dependence (as in the analysis of Acemoglu (2003) applied to explain why capital intensity stays constant) compounds the stronger incentive for labor-augmenting research and labor-augmenting research eventually crowds out energy efficiency improvements. On the other hand, in a model without extreme state dependence, non-price energy efficiency improvements continue at a constant but slower rate than labor-augmenting innovation. Countries that initially have high energy intensity will initially see more rapid energy-augmenting technical change and over time labor-augmenting technical change will become increasingly relatively important. This explains and generalizes the findings of Stern and Kander (2012) for Sweden.

Our empirical analysis shows energy intensity continues to decline among high income countries with low energy intensity, suggesting a scenario without extreme path dependence. Nonetheless, with constant energy prices, energy intensity never declines faster than final output grows, implying energy use increases as long as the economy grows. Furthermore, we should not expect the decline in energy intensity to be more rapid than in the past unless the cost of extracting energy rises or policy changes innovation incentives. Autonomous improvements in energy efficiency are likely, therefore, to have a limited role in mitigating climate change.

APPENDIX A: TECHNICAL APPENDIX FOR SECTION 2

A.1. *Definition of equilibrium*

DEFINITION A.1 An economy $\mathscr{E}$ is an equilibriun if:

1. no aggregate uncertainty holds and for $j \in \{E, L\}$ and $i \in [0,1]$, the processes $(q_j(i,t))_{t=0}^{\infty}$ satisfy (6), where the arrival rate of innovations is identical across i in a sector $j \in \{E, L\}$: given by b_j for the no state dependence economy or s_j for the extreme state dependence economy
2. at each $t \in \mathbb{R}_+$ and $i \in [0,1]$ given $p_E(t)$ and $p_L(t)$, monopolists who own blueprints of machine type i with quality q pick prices, $p_E^x(i,t\,|\,q)$ and $p_L^x(i,t\,|\,q)$, to maximize profits
3. at each t and for each draw from the random variable q_E and q_L, given $p_E(t)$, $p_E^x(t)$, $p_L^x(t)$ and $p_L(t)$, intermediate goods producers choose $E(t)$, $L(t)$ and $x_E(i,t,q_E(i,t))$ and $x_L(i,t,q_L(i,t))$ to maximize profits
4. final goods markets producers maximize profits given $p_E(t)$ and $p_L(t)$ and the market for intermediate goods clears
5. given r and w, consumers choose a to maximize their inter-temporal utility
6. the corporate asset market clears, that is $a(t) = V_L(t) + V_E(t)$, where V_L and V_E are defined by (28)
7. given q_E and q_L, the value functions satisfy

$$v_j(i,t\,|\,q_j(i,t)) = \mathbb{E}_t \int_{s=t}^{T_j(i,t)} e^{-\rho s} \pi_j(i,s\,|\,q_j(i,t))\, \mathrm{d}s, \qquad j \in \{E, L\}, \mathbb{P} - a.e.$$

   where $\pi(i,s\,|\,q)$ is given in terms of the paths of prices, energy use and a quality q by (14)[27]
8. without state dependence, $s_E(t) = s_L(t) = 0$ and the free entry and exit condition

$$\eta_j \frac{v_j(i,t\,|\,q(i,t))}{q(i,t)} = 1, \qquad j \in \{E, L\}, \mathbb{P} - a.e.$$

   holds
9. with state dependence, $Z_E(t) = Z_L(t) = 0$ and the free entry and exit, (30), holds
10. the resource constraint (2) holds and $L(t) = L$ for all t.

[27]Recall expectations here are taken over T, where T is the random stopping time after which variety i in sector j experiences an innovation and the incumbent experiences zero profits thereafter. The distribution of T is pinned down entirely by the paths b_j and s_j for $j \in \{E, L\}$

A.2. *Proofs for section 2.3: main results for the no state dependence economy*

We begin with a proof for the claim that shows by autonomous energy efficiency improvements cannot occur along a BGP. 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 2.1.

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:

$$g_Y = (1 - \tilde{\theta}(t))g_X + \tilde{\theta}(t)g_Z < (1 - \tilde{\theta}(t))g_Z + \tilde{\theta}(t)g_Z = g_Z \tag{42}$$

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 (42) which says $g_Y < g_Z$, the fact that $\tilde{F}$ is increasing and observing $e^x < 1$ for any $x < 0$.

However, $Y(0) < Y(0)$ is a contradiction, implying $g_X \geqslant g_Z$. The case to rule out $g_X > g_Z$ is symmetric by replacing X with Z in the above steps, establishing $g_X \ngtr g_Z$.

Q.E.D.

Now we turn to proofs for the main results concerning the no state dependence economy. Some intermediate results concerning transitional dynamics of growth paths are given in a subsection below (subsection A.2.1). To prepare for the proof of the main result, we now derive growth equations that hold in equilibrium. Take time derivatives of $\theta_E(t)$, defined by (20), to write

$$\text{(43)} \qquad \hat{\theta}_E(t) = \frac{\sigma - 1}{\sigma}(1 - \theta_E(t))\left(\hat{Q}_E(t) + \hat{E}(t) - \hat{Q}_L(t)\right)$$

Similarly, take time derivatives of $Y(t)$, defined by (18), to write

$$\text{(44)} \qquad \hat{Y}(t) = (1 - \theta_E(t))\,\hat{Q}_L(t) + \theta_E(t)\left(\hat{Q}_E(t) + \hat{E}(t)\right)$$

Now, note from (19), we have $\hat{\theta} = \hat{E} - \hat{Y}$. Use this expression in Equation (43) to write

$$\text{(45)} \qquad \hat{E}(t) = \frac{\sigma - 1}{\sigma}(1 - \theta_E(t))(\hat{B}(t) + \hat{E}(t)) + \hat{Y}(t)$$

Next, combine (45) with (44) to write

$$\text{(46)} \qquad \hat{E}(t) = \frac{\sigma}{1 - \theta_E(t)}\hat{Q}_E(t) - \hat{B}(t)$$

Let $\Gamma\colon \mathbb{R}_+ \to \mathbb{R}_+$ define the profitability ratio between energy and labor-augmenting research:

$$\text{(47)} \qquad \Gamma(t) = \frac{\rho + \eta_E b_E(t)}{\rho + \eta_L b_L(t)}, \qquad t \in \mathbb{R}_+$$

If innovation occurs in both sectors at time t, no arbitrage requires

$$\text{(48)} \qquad \Gamma(t) = \left(\frac{Q_E(t)}{Q_L(t)}\right)^{-\frac{1}{\sigma}} \left(\frac{E(t)}{L(t)}\right)^{\frac{\sigma-1}{\sigma}}$$

And thus,

$$\hat{\Gamma}(t) = \frac{\sigma - 1}{\sigma}\hat{E}(t) - \hat{B}(t)\left(\frac{1}{\sigma}\right) \tag{49}$$

Now combine (46) with (49),

$$\begin{aligned}
\hat{\Gamma}(t) &= \frac{\sigma - 1}{\sigma}\hat{E} - \frac{\hat{B}(t)}{\sigma} \\
&= \frac{\sigma - 1}{\sigma}\left(\frac{\sigma}{1 - \theta_E(t)}\hat{Q}_E(t) - \hat{B}(t)\right) - \frac{\hat{B}(t)}{\sigma} \\
&= \frac{\sigma - 1}{1 - \theta_E(t)}\hat{Q}_E(t) - \hat{B}(t)
\end{aligned} \tag{50}$$

PROPOSITION A.1 *Let $\mathscr{E}$ be a no state dependence equilibrium and let assumptions 1-2 hold. If $b_E(0) = 0$, then there exists $t' > 0$ such that $b_E(t') > 0$.*

PROOF: Suppose by contradiction $b_E(t) = 0$ for all $t \in \mathbb{R}_+$. Then Q_E is constant, and by (37), θ_E is constant. By Assumption 2 and the free entry-exit condition (24), b_L is strictly positive implying $Q_L(t) \to \infty$ and by (46), $E(t) \to \infty$. Since, using (20), we can write

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

we must have $\alpha p_E(t)^{\frac{1}{\alpha}} E(t) \to \infty$. However, the free entry-exit conditions (24) are now violated because we assumed $b_E(t) = 0$ for all t, yielding a contraction.

Q.E.D.

PROOF OF THEOREM 2.1: First we verify there exists some t' such that for all $t > t'$, $b_E(t) > 0$ and $b_L(t) > 0$. By proposition A.1, there exists t' such that $b_E(t') > 0$ and $b_L(t') > 0$. Then by Claim A.2 in the Appendix, there exists $T > t'$ such that for all $t > T$

$$\frac{\eta_E b_E(t)}{\eta_L b_L(t)} \geqslant \frac{1 - \theta_E(t)}{2 - \theta_E(t) - \sigma}$$

By Assumption 2 and the free entry-exit conditions (24), b_L is bounded below by a strictly positive constant $\lambda \eta_L \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} (1 - \theta_E(0))^{\frac{1}{1-\sigma}} L - \rho$. Moreover, by (37), $\theta_E(t) \leqslant \theta(0)$ for all t, and, as such,

$$\eta_E b_E(t) \geqslant \frac{1 - \theta_E(t)}{2 - \theta_E(t) - \sigma}\eta_L b_L(t) \geqslant \frac{1 - \theta_E(0)}{2 - \sigma}(\lambda \eta_L \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} (1 - \theta_E(0))^{\frac{1}{1-\sigma}} L - \rho) > 0$$

Now using the free entry-exit conditions (24), the above equation implies (48) - (50) holds for all $t > T$.

Along an ABGP, since $\Gamma(t)$ converges to a positive constant, $\hat{\Gamma}(t) \to 0$. By (50), recalling $\hat{B}(t) = \hat{Q}_E(t) - \hat{Q}_L(t)$, we can establish part 1. of the theorem

$$\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma} \tag{51}$$

Since $b_E(t)$ is bounded below by a strictly positive constant for all $t > T$, we must have $\hat{Q}_E(t)$ is strictly positive for all $t > T$. Accordingly, $\theta_E(t) \to 0$ by (37), giving part 2 of the theorem. To establish part 3, write

$$b_L(t) = \lambda\eta_L\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}\left(1-\theta_E(t)\right)^{\frac{1}{1-\sigma}}L - \rho \to \lambda\eta_L\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}L - \rho \tag{52}$$

Thus $\hat{Q}_L(t) \to (\lambda-1)(\lambda\eta_L\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}L-\rho)$. Now note by (44), $\hat{Y}(t) \to \hat{Q}_L(t)$, establishing part 3.

Finally, for part 4, recall (21) and take taking time derivatives, to arrive at

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

Q.E.D.

PROOF OF PROPOSITION 2.2: Along an asymptotic balanced growth path $b_E(t) \to b_E^\star$ and $b_L(t) \to b_E^\star$, where by Theorem 2.1, $b_E^\star > 0$ and $b_L^\star > 0$.

Recall

$$\begin{aligned}
\eta_E\alpha p_E(t)^{\frac{1}{\alpha}}E(t) &= \eta_E\alpha\gamma_E^{\frac{\epsilon}{\sigma-1}}\theta_E(t)^{\frac{1}{1-\sigma}}E(t) \\
&= \eta_E\alpha\gamma_E^{\frac{\epsilon}{\sigma-1}}\theta_E(t)^{\frac{2-\sigma}{1-\sigma}}\frac{E(t)}{\theta_E(t)} \\
&= \eta_E\alpha^2\kappa(t)^{-1}\gamma_E^{\frac{\epsilon}{\sigma-1}}\theta_E(t)^{\frac{2-\sigma}{1-\sigma}}Y(t)
\end{aligned} \tag{53}$$

where the first equality is from Equation (63), the second equality follows from dividing through and multypying by $\theta_E(t)$ and the final equality follows Equation (19). Now, by Equation (51) in the proof of Theorem 2.1,

$$\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \to \frac{1-\theta_E(t)}{2-\sigma-\theta_E(t)} \tag{54}$$

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

$$\frac{\eta_E b_E(t)}{\eta_L b_L(t)} \to \frac{1-\theta_E(t)}{2-\sigma-\theta_E(t)}$$
$$\implies \frac{\eta_E \alpha^2 \kappa(t)^{-1} \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{2-\sigma}{1-\sigma}} Y(t) - \rho}{\alpha \eta_L \gamma_L^{\frac{\epsilon}{\sigma-1}} (1-\theta_E(t))^{\frac{1}{1-\sigma}} L - \rho} \to \frac{1-\theta_E(t)}{2-\sigma-\theta_E(t)}$$

The implication follows from the no arbitrage conditions, Equation (53) and (62). The above implies

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

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 (19), implies $\frac{E(t)}{Y(t)} \to G(Y(t))$.

Q.E.D.

A.2.1. *Transitional dynamics for the no state dependence economy*

Figure A.9 shows the dynamics for an equilibrium with no state dependence from two possible initial conditions. First, if $\frac{\eta_E b_E(0)}{\eta_L b_L(0)} \leqslant \frac{1-\theta_E(0)}{2-\theta_E(0)-\sigma}$, then Claim A.2 confirms $\frac{\eta_E b_E}{\eta_L b_L}$ rises till $\frac{\eta_E b_E(t)}{\eta_L b_L(t)} > \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$. Second, if $\frac{\eta_E b_E(t)}{\eta_L b_L(t)} > \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$, we are not able to verify exact conditions for the direction of movement of $\frac{\eta_E b_E(t)}{\eta_L b_L(t)}$. However, proposition (A.2) confirms any equilibrium is an asymptotic growth path, which, along with Theorem 2.1 confirms $\lim_{t\to\infty} \frac{\eta_E b_E(t)}{\eta_L b_L(t)} = \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$, implying the paths shown by figure A.9.

CLAIM A.2 Let $\mathscr{E}$ be an equilibrium with no state dependence. If assumptions 1-

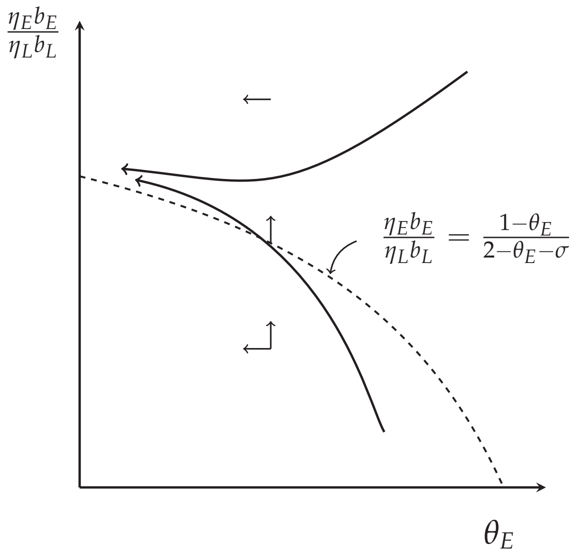

FIGURE A.9.— All ABGPs must converge to a BGP where energy intensity is zero.

2 hold and for some $\bar{t}$, $b_E(\bar{t}) > 0$, then there exists $T > \bar{t}$ such that for all $t > T$, $\frac{\eta_E b_E(t)}{\eta_L b_L(t)} \geqslant \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$.

PROOF: If $\frac{\eta_E b_E(\bar{t})}{\eta_L b_L(\bar{t})} \leqslant \frac{1-\theta_E(\bar{t})}{2-\theta_E(\bar{t})-\sigma}$, then $\frac{\hat{Q}_E(\bar{t})}{\hat{Q}_L(\bar{t})} \leqslant \frac{1-\theta_E(\bar{t})}{2-\theta_E(\bar{t})-\sigma}$. Recall the definition of Γ at equation (47) in the main text, and note (48) from the main text holds along the equilibrium path. Recalling Equation (50) from the main text and using $\frac{\hat{Q}_E(\bar{t})}{\hat{Q}_L(\bar{t})} \leqslant \frac{1-\theta_E(\bar{t})}{2-\theta_E(\bar{t})-\sigma}$, we have:

$$\hat{\Gamma}(\bar{t}) = \frac{\sigma-1}{1-\theta_E(\bar{t})}\hat{Q}_E(\bar{t}) - \hat{B}(\bar{t}) \geqslant 0 \tag{55}$$

Now take growth rates of Γ at $\bar{t}$ as defined at Equation (47) in the main text and use Equation (55) to arrive at:

$$\frac{\eta_E \dot{b}_E(\bar{t})}{\eta_L \dot{b}_L(\bar{t})} \geqslant \frac{\rho + \eta_E b_E(\bar{t})}{\rho + \eta_L b_L(\bar{t})} \tag{56}$$

which in turn implies:

$$\frac{\hat{b}_E(\bar{t})}{\hat{b}_L(\bar{t})} \geqslant \frac{\eta_L b_L(\bar{t})}{\eta_E b_E(\bar{t})}\frac{\rho + \eta_E b_E(\bar{t})}{\rho + \eta_L b_L(\bar{t})} > 1 \tag{57}$$

Thus if $\frac{\eta_E b_E(\bar{t})}{\eta_L b_L(\bar{t})} \leqslant \frac{1-\theta_E(\bar{t})}{2-\theta_E(\bar{t})-\sigma}$, then $\frac{\eta_E b_E(t)}{\eta_L b_L(t)}$ rises till $\frac{\eta_E b_E(T)}{\eta_L b_L(T)} > \frac{1-\theta_E(T)}{2-\theta_E(T)-\sigma}$ for some $T > \bar{t}$.

Moreover, if $\frac{\eta_E b_E(T)}{\eta_L b_L(T)} \geqslant \frac{1-\theta_E(T)}{2-\theta_E(T)-\sigma}$ for some T, then for any $t > T$, $\frac{\eta_E b_E(t)}{\eta_L b_L(t)} \geqslant \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$. To see why, suppose for some $t' > T$, we have $\frac{\eta_E b_E(t')}{\eta_L b_L(t')} < \frac{1-\theta_E(t')}{2-\theta_E(t')-\sigma}$, then there must be some $t'' \leqslant t'$ such that $\frac{\eta_E b_E(t'')}{\eta_L b_L(t'')} = \frac{1-\theta_E(t'')}{2-\theta_E(t'')-\sigma}$ and $\frac{\eta_E b_E(t'')}{\eta_L b_L(t'')}$ is decreasing. However, at t'', (55) must then hold, which in turn implies (56) holds and $\frac{\eta_E b_E(t'')}{\eta_L b_L(t'')}$ cannot be decreasing.

Q.E.D.

PROPOSITION A.2 *If $\mathscr{E}$ is a no state dependence equilibrium and $\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}(1-\theta_E(0))^{\frac{1}{1-\sigma}}L - \rho > 0$, then $\mathscr{E}$ is an ABGP.*

PROOF: To show $\mathscr{E}$ is an ABGP, we show $\lim_{t\to\infty} b_E(t) = b_E^\star$ and $\lim_{t\to\infty} b_L(t) = b_L^\star$.

Before commencing the proof, note that since Q_E is non-decreasing, by Equation (37), $\theta_E(t) \leqslant \theta_E(0)$ for all $t \in \mathbb{R}_+$. Moreover, use (24) and (62) to write equilibrium innovation in the labor sector as:

$$(58) \qquad \rho + \eta_L b_L(t) = \alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}(1-\theta_E(t))^{\frac{1}{1-\sigma}}L$$

implying $\eta_L b_L(t) \in [\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}(1-\theta_E(0))^{\frac{1}{1-\sigma}}L - \rho, \alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}L - \rho]$ for all t.

Next, by Claim A.2, there exists T such that for all $t > T$, $\frac{\eta_E b_E(t)}{\eta_L b_L(t)} \geqslant \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$. Thus, we have:

$$(59) \qquad b_E(t) \geqslant \frac{\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}(1-\theta_E(0))^{\frac{1}{1-\sigma}}L - \rho}{\eta_E}\frac{1-\theta(0)}{2-\sigma} := \underline{b}_E, \qquad t \geqslant T$$

Since $b_E(t) \geqslant \underline{b}_E > 0$ for all $t > T$, $\lim_{t\to\infty} Q_E(t) = \infty$ and by equation (37), $\lim_{t\to\infty}\theta_E(t) = 0$, which in turn implies by Equation (58) that $\lim_{t\to\infty}\eta_L b_L(t) = \alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}L - \rho$. Let $b_L^\star := \lim_{t\to\infty} b_L(t) = \alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}L - \rho$.

We now turn to show $\lim_{t\to\infty} b_E(t) = b_E^\star$ for some $b_E \in \mathbb{R}_{++}$. Recall the definition of $\Gamma(t)$ at equation (47), and note (48) holds along the equilibrium path. Recalling Equation (50) and using $\frac{\hat{Q}_E(t)}{\hat{Q}_L(t)} \geqslant \frac{1-\theta_E(t)}{2-\theta_E(t)-\sigma}$, we have:

$$\hat{\Gamma}(t) = \frac{\sigma-1}{1-\theta_E(t)}\hat{Q}_E(t) - \hat{B}(t) \leqslant 0, \qquad t \geqslant T$$

Since $\hat{\Gamma}(t) = \frac{\dot{\Gamma}(t)}{\Gamma(t)}$ and $\Gamma(t) \geqslant 0$, $\Gamma(t)$ must be decreasing for all $t > T$. However,

$$\Gamma(t) = \frac{\rho + b_E(t)}{\rho + b_L(t)} \geqslant \frac{\rho + \underline{b}_E}{\alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} L} > 0, \qquad t \geqslant T$$

implying $\lim_{t\to\infty} \Gamma(t) > 0$. Let $\bar{\Gamma} := \lim_{t\to\infty} \Gamma(t)$ and use the properties of limits of functions to conclude:

$$\lim_{t\to\infty} b_E(t) = \lim_{t\to\infty} \left(\Gamma(t)(b_L(t) + \rho) - \rho\right) = \left(\bar{\Gamma}(b_L^\star + \rho) - \rho\right) =: b_E^\star$$

where $b_E^\star \in \mathbb{R}_{++}$, thus completing the proof. *Q.E.D.*

PROOF OF CLAIM 2.1: We will use the growth equations and notation developed in the proof of Theorem 2.1. 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 (62) and (53). 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 (21), 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 (44) along with (46) to derive:

$$\hat{Y}(t) = \frac{1+\theta_E(t)(\sigma-1)}{1-\theta_E(t)}\hat{Q}_E(t) - \hat{B}(t) \tag{60}$$

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.*

A.3. *Proofs for state dependence economy, section 2.4*

We prepare some preliminary notation before turning to the proof of theorem 2.2. Noting $p_L(t) = \gamma_L \left(\frac{Y(t)}{Y_L(t)}\right)^{\frac{1}{\epsilon}}$, we have

$$\begin{aligned}
p_L(t) = \gamma_L \left(\frac{Y(t)}{Y_L(t)}\right)^{\frac{1}{\epsilon}} &= \gamma_L \left(\gamma_L + \gamma_E \left(\frac{Y_E}{Y_L}\right)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{1}{\epsilon-1}} \\
&= \gamma_L^{\frac{\epsilon}{\epsilon-1}} \left(1 + \gamma^{\frac{\epsilon}{\sigma}} \left(\frac{Q_E(t)E(t)}{Q_L(t)L}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\epsilon-1}}
\end{aligned} \tag{61}$$

where the third equality uses (17). Using (20), we can then write

$$\alpha p_L(t)^{\frac{1}{\alpha}} L = \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} \left(1-\theta_E(t)\right)^{\frac{1}{1-\sigma}} L \tag{62}$$

similarly

$$\alpha p_E(t)^{\frac{1}{\alpha}} E(t) = \alpha \gamma_E^{\frac{\epsilon}{\sigma-1}} \theta_E(t)^{\frac{1}{1-\sigma}} E(t) \tag{63}$$

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 (28),

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

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

$$\dot{J}_L(t) = \left(\rho + \eta_L s_L(t)\right) J_L(t) - \alpha p_L(t)^{\frac{1}{\alpha}} L \tag{64}$$

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

$$\dot{J}_E(t) = \left(\rho + \eta_E s_E(t)\right) J_E(t) - \alpha p_E(t)^{\frac{1}{\alpha}} E(t) \tag{65}$$

LEMMA A.1 *If an equilibrium with state 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 $\theta_E(t)_E \to 0$, using (62), we have

$$\alpha p_L(t)^{\frac{1}{\alpha}} L \to \alpha \gamma_L^{\frac{\epsilon}{\sigma-1}} L \tag{66}$$

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 (64), 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 (64),

$$J_L(t) \to \frac{\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}} L}{\rho + s_E^\star} \tag{67}$$

Also note

$$\hat{J}_L(t) = \frac{\dot{J}_L(t)}{J_L(t)} \tag{68}$$

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

$$J_E(t) = \frac{Q_L(t)J_L(t)}{Q_E(t)}, \qquad t > \bar{T} \tag{69}$$

Taking time derivatives, we have

$$\hat{J}_L(t) - \hat{J}_E(t) = \hat{Q}_E(t) - \hat{Q}_L(t), \qquad t > \bar{T} \tag{70}$$

Rearranging gives

$$\hat{J}_E(t) = \hat{Q}_L(t) - \hat{Q}_E(t) + \hat{J}_L(t) \to \eta_L s_L^\star - \eta_E s_E^\star \tag{71}$$

where $\hat{J}_L(t) \to 0$ by the argument proceeding Equation (68). Now use the above with the HJB condition for $J_E(t)$, Equation (65),

$$\frac{\alpha p_E(t)^{\frac{1}{\alpha}} E(t)}{J_E(t)} = \rho + \eta_E s_E(t) - \hat{J}_E(t) \to \rho + 2\eta_E s_E^\star - \eta_L s_L^\star \tag{72}$$

Note the limit is a (possibly zero) constant. Moreover, by the HJB condition for $J_L(t)$, Equation (64),

$$\frac{\alpha p_L(t)^{\frac{1}{\alpha}} L}{J_L} = \rho + \eta_L s_L^\star - \hat{J}_L(t) \to \rho + \eta_L s_L^\star \tag{73}$$

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

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

$$= \frac{p_E^{\frac{1}{\alpha}} Q_E(t)E(t)}{p_L^{\frac{1}{\alpha}} Q_L(t)L} = \frac{p_E^{\frac{1}{\alpha}} J_L(t)E(t)}{p_L^{\frac{1}{\alpha}} J_E(t)L} \to \frac{\rho + 2\eta_E s_E^\star - \eta_L s_L^\star}{\rho + \eta_L s_L^\star}$$

where the first equality comes from (36), the second equality uses (16), the third uses the free entry and exit conditions (30) and convergence at the final step follows from (72) and (73) 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.3 If an equilibrium with state 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

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

$$= \eta_E s_E(s) \int_0^\infty \int_0^T e^{\int_0^s \hat{E}(\bar{s}+t) - \eta_E s_E(\bar{s}+t) - \rho \,\mathrm{d}\bar{s}}\,\mathrm{d}s\, e^{-\int_0^T \eta_E s_E(s_1+t)\mathrm{d}s_1}\,\mathrm{d}T$$

Define

$$c := \lim_{t\to\infty}\{\hat{E}(t) - 2\eta_E s_E(t) - \rho\} \tag{74}$$

We now show $c < 0$. Taking growth rates across Equation (21) 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 (44). If $s_L^\star = 0$, then

$$c = \eta_E s_E^\star(\sigma - 1) + \eta_L s_L^\star - 2\eta_E s_E^\star - \rho = \eta_E s_E^\star(\sigma - 3) - \rho < 0 \tag{75}$$

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

$$c = \eta_E s_E^\star(\sigma - 1) + \eta_L s_L^\star - \eta_E 2 s_E^\star - \rho = s_E^\star(\sigma - 1) < 0 \tag{76}$$

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

$$\hat{E}(t) - \eta_E s_E(t) - \rho < \eta_E s_E^\star(\sigma - 2) + \eta_L s_L^\star - \rho + \frac{\epsilon}{2} := c_1 \tag{77}$$

and

$$-\eta_E s_E(t) < -\eta_E s_E^\star + \frac{\epsilon}{2} := c_2 \tag{78}$$

The above two inequalities give

$$\begin{aligned} &\eta_E s_E(s) \int_0^\infty \int_0^T e^{\int_0^s \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}} \, \mathrm{d}s \, e^{-\int_0^T \eta_E s_E(s_1) \mathrm{d}s_1} \, \mathrm{d}T \\ &\leqslant \eta_E s_E(s) \int_0^\infty \int_0^T e^{\int_0^s c_1 \, \mathrm{d}\bar{s}} \, \mathrm{d}s \, e^{\int_0^T c_2 \mathrm{d}s_1} \, \mathrm{d}T = \int_0^\infty \frac{e^{T(c_1 + c_2)} - e^{Tc_2}}{c_1} \, \mathrm{d}T := M < \infty \end{aligned} \tag{79}$$

where the first inequality follows from monotonicity of the exponential function and noting (77) and (78). 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 2.2: If $s_E^\star > 0$, then there exists $\bar{T}$ such that for $t > \bar{T}$, the no arbitrage condition holds

$$\frac{J_E(t) Q_E(t)}{J_L(t) Q_L(t)} \geqslant 1, \qquad t > \bar{T} \tag{80}$$

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

$$
\begin{aligned}
\frac{J_E(t)Q_E(t)}{J_L(t)Q_L(t)} &= \frac{Q_E(t)\mathbb{E}_t \int_{s=t}^{T_E(t)} \pi_E(s)^{\frac{1}{\alpha}} E(s) e^{-\rho s}\, \mathrm{d}s}{Q_L(t)J_L(t)} \\
&= \kappa(t)(1-\theta_E(t))^{1-\sigma} L \left(\frac{\theta_E(t)}{1-\theta_E(t)}\right)^{\frac{1}{1-\sigma}} \\
&\quad \times \frac{E(t)Q_E(t)}{LQ_L(t)} \frac{\mathbb{E}_t \int_{s=t}^{T(t)} e^{\int_{\bar{s}=t}^{s} \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}}\, \mathrm{d}s}{J_L(t)} \\
&= \kappa(t) \left(1-\theta_E(t)\right)^{1-\sigma} L \left(\frac{\theta_E(t)}{1-\theta_E(t)}\right) \\
&\quad \times \frac{\mathbb{E}_t \int_{s=t}^{T(t)} e^{\int_{\bar{s}=t}^{s} \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho \, \mathrm{d}\bar{s}}\, \mathrm{d}s}{J_L(t)}
\end{aligned}
\tag{81}
$$

where we have omitted constants in front of the RHS above for simplicity. The second equality above uses (63) 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 (81) uses (36).

To complete the proof, by Claim A.3, $\mathbb{E}_t \int_{s=t}^{T} e^{\int_{\bar{s}=t}^{s} \hat{E}(\bar{s}) - \eta_E s_E(\bar{s}) - \rho\, \mathrm{d}\bar{s}}\, \mathrm{d}s < M$, 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 (80).

Q.E.D.

PROOF OF PROPOSITION 2.3: We first prove the *if* statement of the proposition.

Part 1: If

Suppose (31) holds, we will show an allocation satisfying 1.- 4. of Definition A.1 with $\eta_L s_L(t) = \eta_L$ for all t is an equilibrium (satisfies 6. of Definition A.1) and is a

BGP. To show the allocation satisfies 6. of Definition A.1, it is sufficient to confirm the free entry and exit condition holds along the allocation path, that is,

$$\frac{J_L(t)Q_L(t)}{J_E(t)Q_E(t)} \geqslant 1, \qquad \forall\, t \geqslant 0 \tag{82}$$

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

$$\theta_E(t) = \theta_E^{\star} := (\alpha(1-\alpha))^{\sigma-1}\,\kappa(t)^{1-\sigma}\gamma_E^{\frac{\epsilon}{\sigma-1}}Q_E(0)^{\sigma-1}, \qquad \forall\, t \geqslant 0 \tag{83}$$

By (62), $p_L(t)$ remains constant. As such,

$$J_L(t) = \frac{\alpha p_L(t)^{\frac{1}{\alpha}}L}{\rho+\eta_L} = \frac{\alpha\gamma_L^{\frac{\epsilon}{\sigma-1}}\left(1-\theta_E^{\star}\right)^{\frac{1}{1-\sigma}}L}{\rho+\eta_L}, \qquad \forall\, t \geqslant 0 \tag{84}$$

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 (63). For the third equality, note when $\hat{Q}_E(t) = 0$, $\hat{E}(t) = \eta_L = \hat{Y}(t)$ for all t by (21) and (44). The third equality follows from solving the integral.

We can now write

$$\begin{aligned}
\frac{J_L(t)Q_L(t)}{J_E(t)Q_E(t)} &= \gamma^{\frac{-\epsilon}{\sigma-1}}\left(\frac{1-\theta_E^{\star}}{\theta_E^{\star}}\right)^{\frac{1}{1-\sigma}}\left(\frac{LQ_L(t)}{E(t)Q_E(t)}\right)\frac{\rho-\eta_L}{\rho+\eta_L} \\
&= \gamma^{\frac{-2\epsilon}{\sigma-1}}\left(\frac{1-\theta_E^{\star}}{\theta_E^{\star}}\right)\frac{\rho-\eta_L}{\rho+\eta_L} \\
&= \frac{J_L(0)Q_L(0)}{J_E(0)Q_E(0)}
\end{aligned} \tag{85}$$

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 (83).

Using algebra, and (83), we can verify

$$(86) \qquad \gamma^{\frac{-2\epsilon}{\sigma-1}}\left(\frac{1-\theta_E^\star}{\theta_E^\star}\right)\frac{\rho-\eta_L}{\rho+\eta_L} \geqslant 1 \iff \left(\frac{\kappa}{Q_E(0)}\right)^{1-\sigma} \leqslant \frac{(\alpha(1-\alpha))^{1-\sigma}\,\gamma^{\frac{\epsilon}{1-\sigma}}}{\gamma^{\frac{2\epsilon}{\sigma-1}}\frac{\rho+\eta_L}{\rho-\eta_L}+1}$$

And thus, if (31) holds, then $\frac{J_L(0)Q_L(0)}{J_E(0)Q_E(0)} \geqslant 1$, which by (85), implies (82). To confirm the equilibrium allocation is a BGP, note by (21) and (44), $\hat{Y}(t) = \eta_L$ 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 (31) must hold. Since $s_E(t) = 0$ for all t, (85) must hold along the growth path by the argument between proceeding Equation (82) above. Because the path is an equilibrium path, the free entry and exit conditions (82) must hold, and since (85) holds,

$$(87) \qquad \gamma^{\frac{-2\epsilon}{\sigma-1}}\left(\frac{1-\theta_E^\star}{\theta_E^\star}\right)\frac{\rho-\eta_L}{\rho+\eta_L} \geqslant 1$$

which, by (86), implies (31) holds. *Q.E.D.*

APPENDIX B: COMPARATIVE STATICS OF EQUILIBRIUM ENERGY USE AND TECHNOLOGY

In this section, we study the effect of energy-augmenting technology on energy use. For a given level of output and real energy price, $\kappa(t)$, higher $Q_E(t)$ leads to lower energy use. Given a real energy price, we can decompose the effect of energy-augmenting technical change on energy use into the change in energy use given the level of output and the change in energy use as output increases given a level of energy intensity. Solve (21) for E and take the derivative with respect to Q_E to arrive at

$$(88) \qquad \frac{\partial E}{\partial Q_E} = \Phi\overbrace{\frac{\partial Y}{\partial Q_E}Q_E^{\sigma-1}}^{\text{Output effect}} + \Phi\overbrace{(\sigma-1)YQ_E^{\sigma-2}}^{\text{Efficiency effect}}$$

where $\Phi = \gamma_E^{\frac{\epsilon}{\sigma-1}}\alpha^\sigma(1-\alpha)^{\sigma-1}$ and we have dropped 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.

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

PROOF: Evaluate the derivative on the right hand side of (88) and simplify the equation

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

Now note the expression for output (18) and θ_E (20) 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 $(1-\sigma)(1-\alpha)^{\frac{1-\sigma}{\sigma}}$.[28]

APPENDIX C: DATA SOURCES AND DATA ANALYSIS ON FUEL PRICE TRENDS

C.1. *U.S. time series data sources*

C.1.1. *GDP, implicit price deflator, employment, and wages*

Data on GDP and wages from 1929 to 2015 are 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. 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. Total wages from 1929 to 2015 are given by compensation of employees and deflated to 2009 using the GDP implicit price deflator. 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 years prior to 1919 we use these averages 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 can then compute labor productivity and annual wage series per employee from 1900 on.

[28]In a similar exogenous technical change setting, the rebound is less than 1 if $\theta_E < (1-\sigma)$ (Saunders, 2015).

C.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 data were obtained from the HS70 and Appendix D of the EIAs Monthly Energy Review. These data also include animal feed, which was a large source of energy in 1900.

C.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.

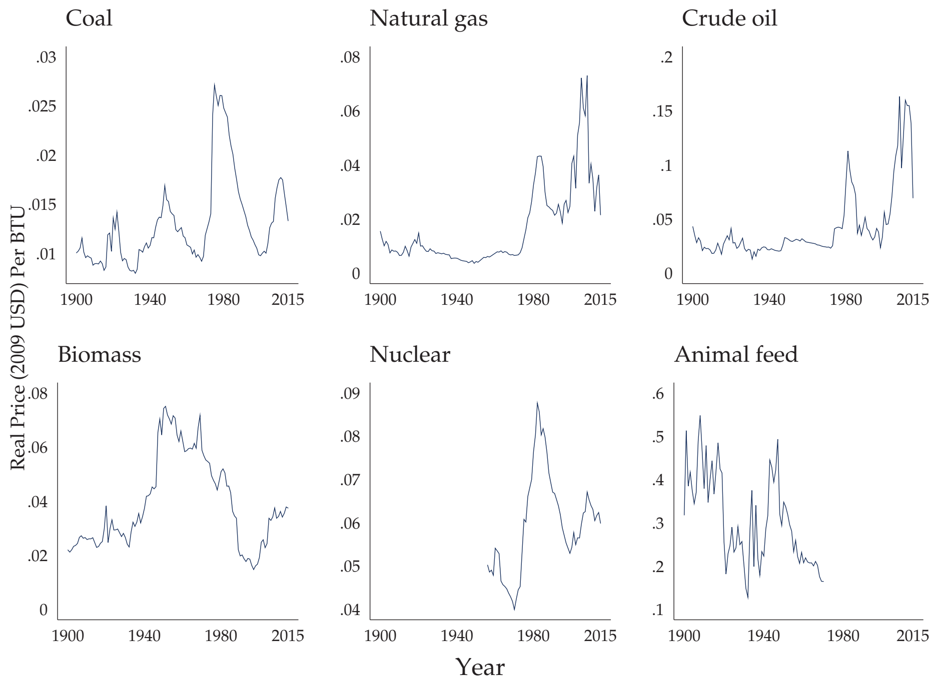

FIGURE C.1.— Real U.S. prices of energy carriers

C.2. *Trend tests on U.S. fuel prices*

Figure C.1 shows the real price in 2009 US Dollars per BTU for the key individual fuels. Even though each of these prices has increased absolutely over time, it is unclear whether these represent systematic trends or not. The aggregate energy price (Figure C.3) rises faster than most of the component prices because there is a positive correlation between price movements and cost shares.

Figures C.2 and C.3 show energy intensity and aggregate U.S. fuel price with and without animal feed. Animal feed was a significant part of the fuel mix prior to the 1930s, after which the two different price and energy intensity series converge. Note a clear positive trend in the aggregate price without animal feed. As noted above, there is a positive correlation between cost shares and changes in energy prices. In particular, oil price movements dominate the aggregate price in later decades.

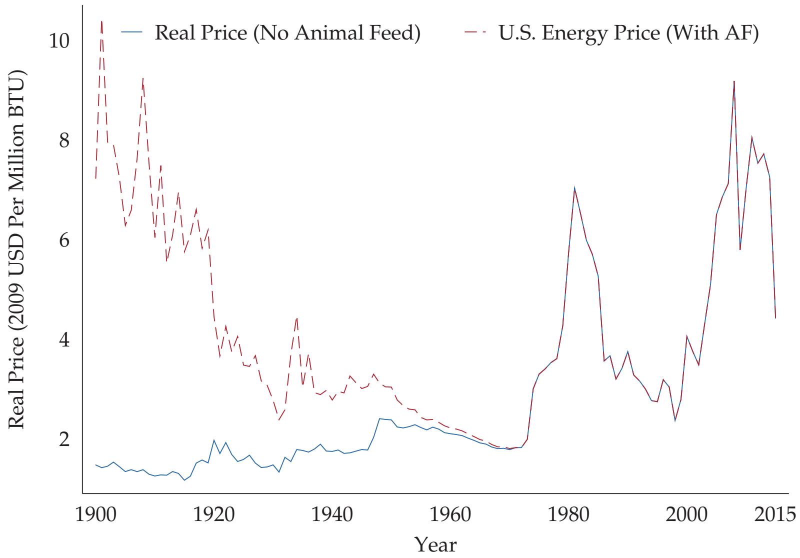

FIGURE C.3.— U.S. aggregate fuel price with and without animal feed (AF).

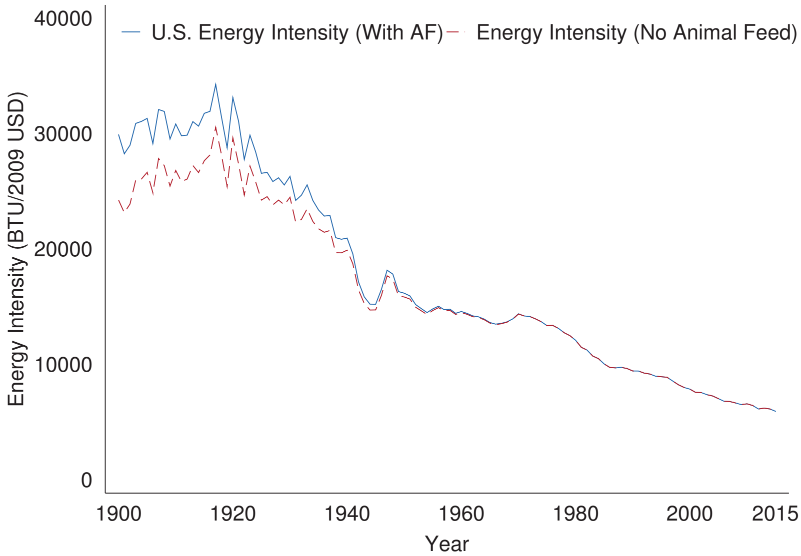

FIGURE C.2.— U.S. energy intensity with and without animal feed (AF).

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

$$y_{i,t} = \beta_{1,i} + \beta_{2,i} t_i + u_{i,t} \tag{90}$$

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:

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

where $\hat{\beta}_{2,i}$ is the estimate of the slope parameter and $\mathrm{se}\left(\hat{\beta}_{2,i}\right)$ its standard error, b is a parameter computed by Fomby and Vogelsang (2003), and

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

where RSS_1 is the sum of squared residuals from (90), 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

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

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

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

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 (Fomby and Vogelsang, 2003): $b = 2.466$, $t_{DAN} = 2.462$ at 1%; $b = 1.795$, $t_{DAN} = 2.052$ at 2.5%; $b = 1.322$, $t_{DAN} = 1.710$ 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 (90) 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 C.1

TREND TEST RESULTS FOR INDIVIDUAL PRICE SERIES — SIGNIFICANT TEST STATISTICS IN BOLD.

Log real price	N	β	$t_{DAN}01$	$t_{DAN}025$	$t_{DAN}05$	$t_{DAN}10$	J
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 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.

TABLE C.2

TREND TEST RESULTS FOR AGGREGATE PRICE SERIES — SIGNIFICANT TEST STATISTICS IN BOLD.

Log real price	β	$t_{DAN}01$	$t_{DAN}025$	$t_{DAN}05$	$t_{DAN}10$	J
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

We then tested for trends in the aggregate prices, the results are shown in figure C.2. 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.

APPENDIX D: ADDITIONAL DISCUSSION OF ENERGY INTENSITY DATA

D.1. *Evolution of the level of energy-augmenting technological change*

In this section, we revisit the observed trends of energy intensity, derive an estimate of energy-augmenting technical change and discuss how they can be explained under both the extreme and no state dependencve scenario.

First we solve Equation (21) for Q_E. In discrete time, adding a stationary — but probably serially correlated error term, we have:

$$Q_{E,t} + u_{E,t} = A_E \kappa_t^{\frac{\sigma}{\sigma-1}} \left(\frac{E_t}{Y_t}\right)^{\frac{1}{\sigma-1}} \tag{91}$$

where $A_E = \gamma_E^{-\frac{\epsilon}{(\sigma-1)^2}} \alpha^{-\frac{\sigma}{\sigma-1}} (1-\alpha)^{-1}$.[29] We compute the RHS of (91) using annual U.S. data from 1900 to 2015 (the data are described in the appendix); we assume $\sigma = .5$ apply the Hodrick-Prescott filter to obtain the estimate of $Q_{E,t}$ plotted in Figure D.4.[30] Before

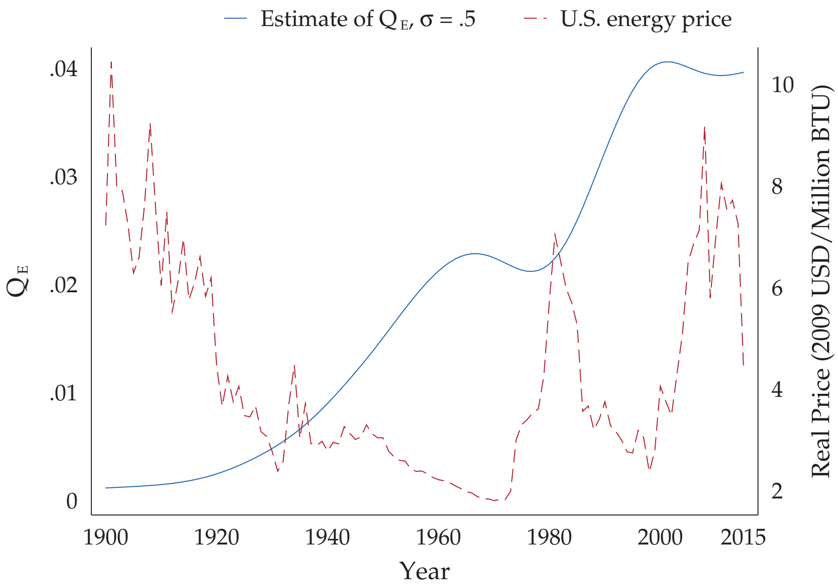

FIGURE D.4.— Smoothed estimate of energy technology stock ($Q_E(t)$) and the U.S. raw energy price. The estimates assume $\sigma = .5$.

1950, the growth rate of Q_E was positive, while prices declined, suggesting an autonomous decline of energy intensity. After 1960, and before 1980, the growth rate of Q_E declines to under zero by the early 1970s, before rising and falling again in a lagged response to prices.

Both the economy without state dependence and economy with state dependence can be consistent with the path of Q_E in the U.S. If we suppose the U.S. economy was sufficiently energy inefficient in the early 1900s, there could have been autonomous improvements in Q_E under a state dependence economy (recall Proposition 2.3). The growth of $Q_E(t)$ does decline below zero just before the price shocks of the 1970s and in the early 2000s, however, this does not imply a model with no autonomous energy efficiency improvements

[29]As A_E only changes the value of $Q_{E,t}$ without affecting its growth rate — and A_L similarly affects $Q_{L,t}$, we set $A_E = A_L = 1$ in the following.

[30]See figure E.7 for raw and smoothed estimates of Q_E for different values of σ. Hassler et al. (2016) assume there is no error in their equivalent of (91). But given our model is very simple, energy intensity likely adjusts slowly to its equilibrium value; and variables are measured with error — the addition of an error term seems more reasonable to us than not. However, values above .8, even with smoothing, give unreasonable jumps in line with the observation made by Hassler et al. (2016). As such, we maintain a reasonable value of σ to be between 0 and .5.

along an ABGP. Both complete state dependence and no state dependence predict such a decline, where the decline in the growth rate of $Q_E(t)$ is associated with a fall in energy prices. To see why, for a given level of $Q_E(t)$ and $Q_L(t)$, when prices fall, energy consumption increases, and recalling (34) and (35), energy-augmenting technical change becomes relatively less profitable due to the stronger price effect.

D.2. *Elasticity of energy intensity to output in the cross-section*

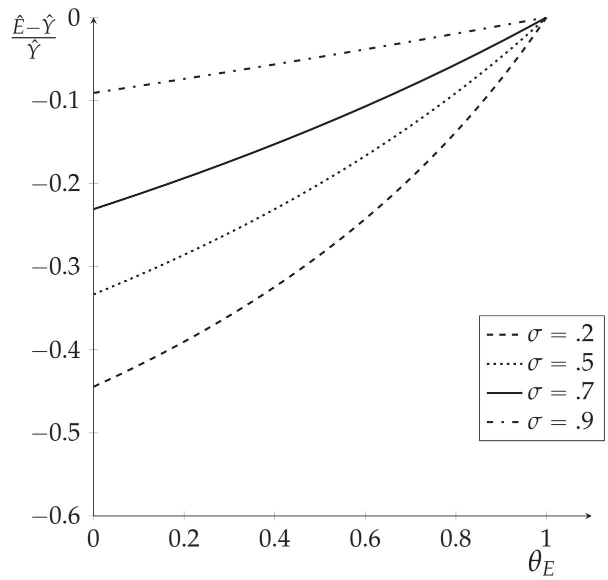

FIGURE D.5.— Elasticity of energy intensity to output along the "stable path".

On the surface, one pience of evidence in favor of the state dependence scenario is the low absolute elasticity of energy intensity to output through time and in the cross-section for countries with low energy intensity. Figure 3 from the introduction presented scatter plots of the natural log of energy intensity against the natural log of GDP per capita in PPP dollars at ten year intervals between 1971 and 2010. The fitted regression lines in the figure are cubic polynomials; 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 regression estimates are presented in Table D.3. Using these coefficients we can estimate the elasticity of energy intensity to output in the cross-section (Table D.4). There is a negative relationship between energy intensity and output, however, for higher levels of output, the fall in energy intensity associated with an increase in output is less. 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 .045 percent fall in energy intensity. At least in the cross section of countries, the elasticity of energy intensity to output is not constant and becomes close to zero among richer countries.

TABLE D.3

CUBIC POLYNOMIAL REGRESSION TABLE.

	1971	1981	1991	2001	2010
Ln(Y)	-11.28	-12.25**	-9.413*	-8.740**	-9.215**
	(6.947)	(4.368)	(4.274)	(2.744)	(3.073)
Ln(Y)^2	1.111	1.280*	0.960	0.906**	0.931*
	(0.865)	(0.525)	(0.517)	(0.335)	(0.368)
Ln(Y)^3	-0.0354	-0.0443*	-0.0326	-0.0316*	-0.0315*
	(0.0355)	(0.0208)	(0.0206)	(0.0134)	(0.0145)
Constant	39.19*	40.84***	32.67**	30.22***	32.34***
	(18.40)	(11.97)	(11.62)	(7.385)	(8.432)
Observations	98	99	99	99	99

Standard errors in parentheses

$*p < 0.05$, $**p < 0.01$, $***p < 0.001$

Consider the elasticity of energy intensity to output towards which ABGPs converge to, given by Theorem 2.1:

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

With constant prices, the low in absolute value elasticities seen in the cross-section of energy intensity to output can only be predicted along the ABGP in the no state dependence model if σ is approximately .9 (see Figure D.5). However, such elasticities of substitution are higher than the estimates in the literature, which range from close to zero (Hassler et al., 2016) to .68 (Stern and Kander, 2012). An elasticity of substitution of .9 is also not consistent with a reasonable evolution of $Q_E(t)$. As Figure E.7 in the appendix suggests, even with smoothing applied, a high σ implies Q_E rises and falls almost by 1000% between 1970 and 2015; it is difficult to attribute Q_E in this case to technical change.

Finally, note the shape of the relationship between $\theta_E(t)$ and the elasticity of energy intensity to output in Figure D.5: the elasticity is *more negative* for lower levels of energy intensity. Figure D.6 shows the elasticity of the average annual rate of energy intensity decline with respect to average annual output growth for the 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. Therefore, the elasticity is more negative in countries with higher energy intensity.

The evidence in this sub-section suggests the cross-section of countries cannot lie *cannot lie along the dotted line in Figure 5*. However, this does not mean the extreme state dependence

TABLE D.4

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

USD/ capita	1971	1981	1991	2001	2010
2,000	-0.527	-0.470	-0.470	-0.444	-0.522
10,000	0.176	0.055	-0.026	-0.093	-0.082
20,000	0.310	0.068	0.010	-0.093	-0.043
30,000	0.320	0.060	0.007	-0.1	-0.045

scenario represents the data. Dispersion of energy intensities between countries could be such that the non extreme state dependence scenario better represents the data but initial conditions give the impression that the cross sectional energy intensity converge to a constant as income rises. For instance, most countries could lie north of the dotted line of figure 5 and countries with higher incomes could have an initial value of energy intensity higher than the value implied by the stable path. Thus to understand clearly whether the extreme state dependence or no state dependence scenario is more realistic, we must assess whether there is sufficient decline in energy intensity among high income countries—this is the purpose of our analysis in section 3.

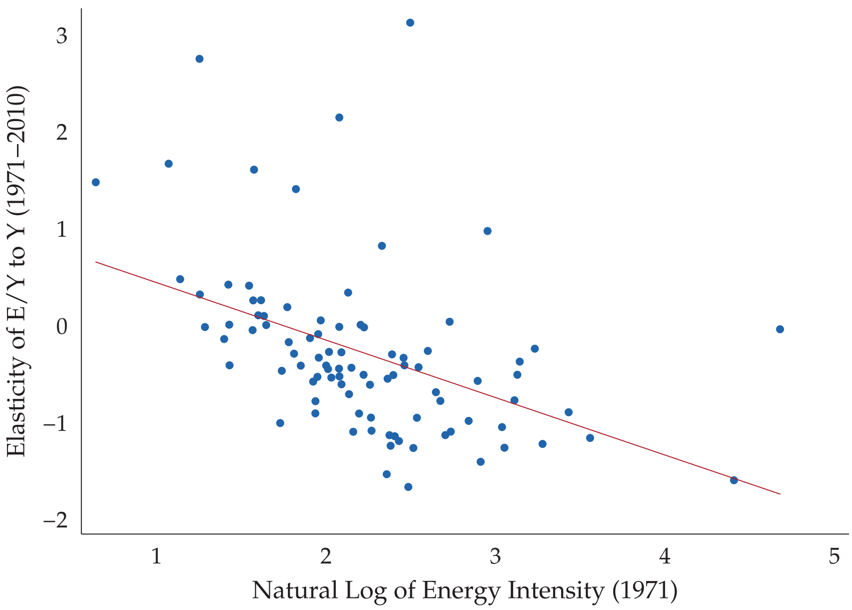

FIGURE D.6.— Elasticity of average rate of decline in energy intensity to average rate of output growth between 1971 and 2010. The line of best fit has a slope coefficient of -0.59 (std. err. 0.12).

APPENDIX E: ADDITIONAL FIGURES

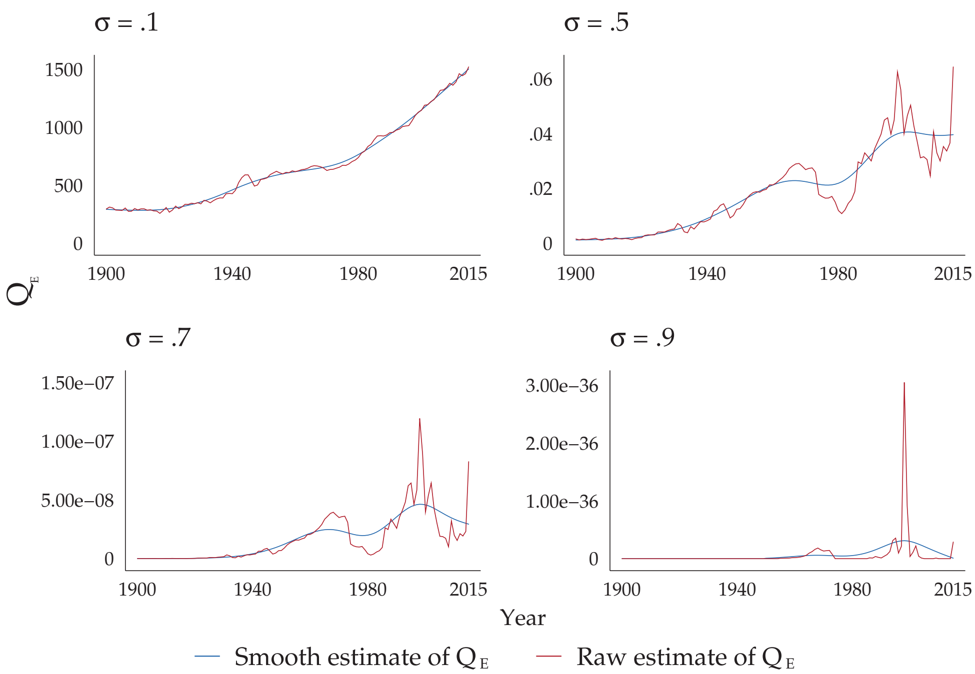

FIGURE E.7.— Smoothed and raw estimate of Q_E for different values of σ.

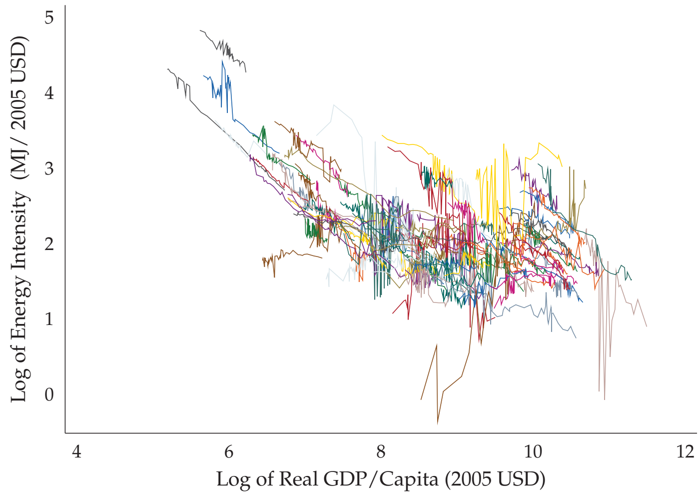

FIGURE E.8.— Natural log of energy intensity and natural log of output for a cross section of countries. Csereklyei et al. (2016) describe the sources of the data.

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. Labor-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., Hemous, D., feb 2012. The Environment and Directed Technical Change. American Economic Review 102 (1), 131–166.

Acemoglu, D., Cao, D., 2015. Innovation by Entrants and Incumbents. Journal of Economic Theory 157, 255–294.

Aghion, P., Hémous, D., Martin, R., Reenen, J. V., Dechezleprêtre, A., 2016. Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry. Journal of Political Economy 124 (1), 1–51.

Aghion, P. P., Howitt, P., 1992. A Model of Growth Through Creative Destruction. Econometrica 60 (2), 323–351.

André, F. J., 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 Reenen, J., Webb, M., 2017. Are Ideas Getting Harder to Find? NBER Working Paper No. 23782.
URL `http://www.nber.org/papers/w23782`

Bunzel, H., 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.

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

EIA, 2017. Annual Energy Outlook. Tech. rep., U.S. Energy Information Admistration.
URL https://www.eia.gov/outlooks/aeo/pdf/0383(2017).pdf

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

Fouquet, R., 2008. Heat, Power and Light: Revolutions in Energy Services. Edward Elgar, Cheltenham.

Gentvilaite, R., Kander, A., Warde, P., 2015. The role of energy quality in shaping long-term energy intensity in Europe. Energies 8 (1), 133–153.

Goulder, L. H., 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), 179–206.

Hart, R., 2018. Rebound, directed technological change, and aggregate demand for energy. Journal of Environmental Economics and Management 89, 218 – 234.
URL http://www.sciencedirect.com/science/article/pii/S0095069618301876

Hassler, J., Krusell, P., Olovsson, C., 2016. Directed Technical Change as a Response to Natural-Resource Scarcity. Working Paper.
URL hassler-j.iies.su.se/PAPERS/estc.pdf

Henriques, S. T., Kander, A., 2010. The modest environmental relief resulting from the transition to a service economy. Ecological Economics 70 (2), 271–282.

IEA, 2016. World Energy Outlook 2016, 1–13.
URL https://www.oecd-ilibrary.org/energy/world-energy-outlook-2016_weo-2016-en

IEA, 2017. Market Report Series: Energy Efficiency. Tech. rep., International Energy Agency.
URL https://www.iea.org/publications/freepublications/publication/market-report-series-energy-efficiency-2017-.html

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

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

Kander, A., Paolo, M., Warde, P., 2015. Power to the People: Energy in Europe over the Last Five Centurie. Princeton University Press, New Jersey.

Kruse-Andersen, P., 2017. Testing R&D-Based Endogenous Growth Models. Working Paper, 1–44.
URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2947528

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

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

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

Pagan, J., 1998. Carbohydrates in Equine Nutrition. In: Pagam, J. (Ed.), Advances in Equine Nutrition. Nottingham University Press, Nottingham.

Popp, D., 2002. Energy Prices and Induced Innovation. American Economic Review 1 (91), 67–75.

Popp, D., jul 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., Jaffe, A. B., 2010. Energy, The Environment, and Technological Change. Handbook of the Economics of Innovation 2 (1), 873–937.

Saunders, H. D., 2015. Recent Evidence for Large Rebound: Elucidating the Drivers and their Implications for Climate Change Models. The Quarterly Journal of the IAEE's Energy Economics Education Foundation 36 (1), 23–48.

Stern, D. I., 2004. Economic Growth and Energy. In: Cleveland, C. J. (Ed.), Encyclopedia of Energy. Vol. 2. Academic Press, San Diego, CA, pp. 35–51.

Stern, D. I., 2017. How Accurate are Energy Intensity Projections? Climatic Change 143 (3-4), 537–545.

Stern, D. I., Kander, A., 2012. The Role of Energy in the Industrial Revolution and Modern Economic Growth. The Energy Journal, International Association for Energy Economics 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., 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., 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.

Webster, M., Paltsev, S., 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., Perry, A. M., 1990. Low-Cost Strategies for Coping with CO2 Emission Limits (A Critique of CO2 Emission Limits: an Economic Cost Analysis for the USA by Alan Manne and Richard Richels). The Quarterly Journal of the IAEE's Energy Economics Education Foundation 11 (4), 35–59.