Strathclyde

Discussion Papers
in Economics

University of Strathclyde Business School

The interaction of scale economies and energy quality

David Comerford

No. 22 – 9

Department of Economics
University of Strathclyde, Glasgow

The interaction of scale economies and energy quality

David Comerford*

Abstract

Ubiquitous in the natural resources literature is that resource scarcity is associated with high resource prices that incentivise the exploitation of marginal resources and the usage of alternatives. In this paper, this incentive is higher profitability in the extractive sector, rather than simply a higher price for its output. Since energy is an essential input to the economy, its supply affects the marginal products of other factors, and in general equilibrium its supply affects the costs that the extractive industry faces. Energy sector profitability is therefore ambiguously affected by the quality of available energy resources. There are conditions related to the returns to scale in the economy which can cause lower energy sector profitability with lower energy quality. This means that marginal resources may be abandoned as high quality resources are lost. An economy which exhibits constant, or weakly increasing, returns to scale can operate at any level of energy quality, since profitability rises with falling energy quality and we observe results consistent with the usual Hotellings Rule. However an economy which exhibits strongly increasing returns to scale cannot operate with only low quality energy resources and profitability may fall with falling energy quality. It is therefore possible that an energy quality shock disincentivises, rather than incentivises, the use of marginal resources and alternatives. Ultimately, a strongly increasing returns to scale economy may have no steady state equilibrium under a decentralised market allocation, despite such an allocation being technologically feasible.

Keywords: Non-Renewable Resources, Energy, General Equilibrium, Returns to Scale.
JEL Classification: Q30, Q43.

*Corresponding author: David Comerford, Department of Economics, University of Strathclyde, Glasgow, G4 0QU; Email: david.comerford@strath.ac.uk

1 Introduction

The purpose of this paper is to explore the economy's response to being faced with only lower quality energy resources, with a view to characterising the situations under which this is problematic. By "Problematic" in this context, I mean that we are unable to use all available resources; or that as resources become scarcer, we receive a price signal that makes the use of resources less efficient, and does not incentivise the use of alternatives. The scenario considered is that capital assets are used to supply energy to the economy, but the manufacture of capital assets can be an energy intensive process. If higher quality energy resources are no longer available, and the use of lower quality resources is to be expanded by applying more capital inputs to their exploitation, then the relative energy (output) to capital (input) price movement will have to be consistent with this expansion. A three good economy is therefore of the minimum complexity required to investigate this issue, and given the feedbacks between available energy resources and economy wide prices, a general equilibrium approach is appropriate.

In the theoretical natural resource economics literature, natural resource scarcity is accompanied by a rise in their price. This leads to fairly sanguine conclusions with regards to the exhaustion of non-renewable resources. Since Hotelling (1931), the defining characteristic of the optimal depletion of non-renewable resources is that resource prices should rise at a rate related to the rate that can be earned by extraction and investing the financial proceeds. This rising price ensures that resources that are initially unprofitable to exploit eventually become profitable, and that there are incentives both to economise on the use of the resources and to develop alternatives. This foundation to the literature has subsequently been built upon, e.g. Holland (2008) describes models of resource extraction that generate a peak in the extraction rate during the extraction period using a partial equilibrium approach (since interest rates and backstop prices do not depend upon energy used in the aggregate economy); Dasgupta and Heal (1974) extend The Hotelling's framework to a general equilibrium setting without changing the conclusion that non-renewable resource prices rise without bound as they become more scarce; and Aghion, Howitt, Brant-Collett, and Garcaia-Peanalosa (1998) describe a two sector general equilibrium model in which growth can be sustained despite declining availability of non-renewable natural resources, that are essential for production, through investment in intellectual capital. In all these cases, natural resource scarcity is accompanied by a rise in their price. Holland claims that price movements will be smoothly increasing because "oil is virtually costless to store in its natural reservoir ... even completely myopic firms without secure property rights would wait to produce from these [higher cost] deposits until the price were high enough to cover the extraction costs". This statement reveals, I believe, a possible shortcoming in this approach: yes, the resources can be left in the ground at zero cost, but there is no guarantee that the intermediate goods which are used to extract these resources will be reasonably priced in future. It may be the case that as resources become scarce, the price of intermediates rises faster than energy prices, and so lower quality energy resources can never be profitably exploited.

The model presented in this paper has several distinctive features. A muti-sector economy is necessary in order for endogenous energy sector input prices, and heterogenous energy resource quality is necessary since we need a marginal firm who decides to produce or to exit. However, the analysis reveals that economies of scale in the intermediate goods sector play a crucial role in determining how the economy responds to declining availability of energy resources: Constant returns, or a low level of increasing returns, are consistent with energy scarcity causing energy prices to rise faster than the prices of intermediates, and so for the economy to profitably expand into lower quality energy resources; However, if the degree of returns to scale in the intermediate goods sector is strong enough then the reduction in this sector's productivity, caused by the restriction in its factor inputs, boosts the price of intermediates by more than the price of energy. The supply of energy therefore contracts rather than expands at the margin, and eventually the economy can collapse.

Models with increasing returns are widespread in other fields e.g. economic geography models with agglomeration effects; business cycle models with increasing returns as a partial explanations for size of fluctuations; endogenous growth models; and new-trade models of intra-industry trade. The mechanism used in this paper and which is common to many papers in the literature is well described in Ventura (2005)[1] in which the cost of final goods production is falling in the number of inputs (increasing returns to scale), but the number of inputs depends on demand from producers of final goods. In Ventura (2005), this leads to a multiplicity of equilibrium locations chosen by industrial sectors. In this paper we see the same non-linear effect as the cost of energy production is falling in the productivity of intermediate inputs, but the productivity of intermediate input production depends on demand from energy producers (increasing returns to scale), which is partially determined by the quality of available energy resources. This interaction between possible increasing returns to scale and energy resource limitations has not previously been considered, and it is easy to imagine that it may be important. For example perhaps the ability and profitability of deep oil drilling is only possible because there is a full manufacturing supply chain that is predicated on the existence of automobile and aerospace industries. Perhaps if there was no cheap oil available, the contraction of the automobile and aerospace sectors could affect the manufacturing supply chain in such a way as to drastically increase costs/decrease productivity in the sector that manufactures equipment for deep oil drilling. This in turn may mean that, without cheap oil sustaining the automobile and aerospace sectors, the deep water drilling sector is unprofitable, and so it would not exist in cheap oil's absence.

[1] Section 3.2 of Ventura (2005).

Given that a large intermediate sector lowers the cost of energy production, whilst a high level of energy production enables a large intermediate sector; the model here presents the possibility for development to go into reverse as this dependence, combined with the loss of energy resources, causes this productive equilibrium to be destabilised. This is suggestive of "The Big Push" story of economic development of Murphy, Schleifer, and Vishny (1989), in which coordinated industrial investment can make such investment profitable (when an investment at the margin would have been unprofitable) through the impact that the coordinated investment has on scale and efficiency. However, in Murphy, Schleifer, and Vishny (1989), endowments are constant and development is a coordination problem; whereas here I explicitly consider a decline (through resource depletion) in the endowment, such that the industrial equilibrium is destabilised[2] i.e. the issue here is the very existence of equilibrium rather than the need to coordinate on a better equilibrium.

The contribution that this article makes is to draw attention to the possibility that price movements in response to scarcity may not be favourable to bringing on substitutes or for using capital intensive but energy efficient alternatives, a possibility which is not considered in the existing literature. The paper is structured as follows: section 2 introduces the macroeconomy and section 3 presents some propositions which capture the mechanism underlying the basic argument presented. The energy sector in section 2 & 3 is completely abstract and has a very simple characterisation. Section 4 presents a specific '*Renewable Energy*' model which satisfies the characterisation of the energy sector given in section 2. Section 5 does likewise with a '*Fossil Fuel*' sector. Section 6 presents some illustrative results from the model, section 7 discusses the interaction between conclusions from this model and the incentives to innovate with policy implications, and section 8 concludes.

2 The macroeconomy

Households and the final goods sector in this model are standard. Final goods are produced using a constant returns, Cobb-Douglas technology, under perfect competition, using energy services,

[2] And I further do not have the "cottage" sector of Murphy, Schleifer, and Vishny (1989) to sustain output if there is no industrial sector.

intermediate goods, and labour as factor inputs. We can write:

$$
\begin{aligned}
Y_t &= a_t L_t^{1-\mu-\gamma} E_t^{\mu} Q_t^{\gamma} = C_t + \dot{K}_t + \delta K_t = w_t L_t + p_K(t) K_t + \Pi_t \\
\text{where} \quad & Y_t \equiv \text{output flow at time } t \\
& a_t \equiv \text{TFP at time } t \\
& L_t \equiv \text{Labour inputs at time } t \\
& E_t \equiv \text{Energy inputs at time } t \\
& Q_t \equiv \text{Intermediate good inputs at time } t \\
& C_t \equiv \text{Consumption at time } t \\
& K_t \equiv \text{Capital stock at time } t \\
& w_t \equiv \text{Wage rate at time } t \\
& p_K(t) \equiv \text{Rental rate of capital at time } t \\
& \Pi_t \equiv \text{Energy sector profit flow rate at time } t \\
& \mu \equiv \text{Income share paid to energy} \\
& \gamma \equiv \text{Income share paid to intermediates} \\
& \delta \equiv \text{Capital depreciation rate}
\end{aligned}
$$

For simplicity assume that TFP is constant and that population growth is zero, so we can write $Y_t = AE_t^{\mu} Q_t^{\gamma}$ (where $A \equiv a_t L_t^{1-\mu-\gamma} = const$). Households have CRRA preferences, and maximise lifetime utility taking the paths of wages, rental rates of capital and energy sector profits as given. Therefore the standard consumption Euler equation holds:

$$
\begin{aligned}
\frac{\dot{C}_t}{C_t} &= \frac{p_K(t) - \delta - \rho}{\epsilon} \\
\text{where} \quad & \rho \equiv \text{rate of time preference} \\
& \epsilon \equiv \text{coefficient of relative risk aversion}
\end{aligned}
$$

An interest rate, $p_K(t)$, less (greater) than $\rho + \delta$ implies that consumption is falling (rising), and consumption, C_t, greater (less) than $AE_t^{\mu} Q_t^{\gamma} - \delta K_t$ implies that capital stock is falling (rising).

Constant returns and perfect competition in the final goods sector give us the prices of energy and intermediate goods:

$$\begin{aligned}
p_E(t) &= \mu Y_t/E_t \\
p_Q(t) &= \gamma Y_t/Q_t \\
\text{define } z_t &\equiv \frac{p_E(t)}{p_Q(t)} \\
\Rightarrow Q_t &= \frac{\gamma}{\mu} z_t E_t \\
\text{so } p_Q(t) &= A\gamma^\gamma \mu^{1-\gamma} z_t^{\gamma-1} E_t^{\mu+\gamma-1}
\end{aligned}$$

There are two further sectors in this economy: an energy sector which uses intermediate goods to access energy resources and supply energy services; and an intermediate goods sector which rents the capital stock and produces intermediate goods. The properties that we place on the energy sector in this section and the next are that:

$$\begin{aligned}
E_t &= f(Q_E(t)) = E(z_t, s_t) \geq 0 \\
Q_E(t) &= Q_E(z_t, s_t) \geq 0 \\
\text{where} &\quad Q_E(t) \text{ is the quantity of intermediate goods used} \\
&\quad f \text{ is the energy from intermediates production function} \\
&\quad z_t \text{ is the energy to intermediates price ratio, taken as given} \\
&\quad s_t \text{ is an index, increasing in energy resource quality} \\
\text{s.t.} &\quad \frac{\partial E_t}{\partial z_t}, \frac{\partial Q_E(t)}{\partial z_t} > 0 \\
\exists z_{min}(s_t) \ \text{s.t.} &\quad E(z_{min}(s_t), s_t) = Q_E(z_{min}(s_t), s_t) = 0 \\
&\quad \frac{\mathrm{d}z_{min}(s_t)}{\mathrm{d}s_t} < 0 \text{ , \& } z_{min} \to \infty \text{ as } s \to 0 \\
\lim_{z\to\infty} E_t &= a(s_t) z_t^x \quad , \quad x \geq 0 \\
\lim_{z\to\infty} \left[\frac{Q_E(t)}{E_t}\right] &= b(s_t) z_t^y \quad , \quad 0 \leq y \leq 1 \\
\lim_{z\to z_{min}} E_t &= c(s_t) g(z_t) \quad , \quad g(z_{min}) = 0 \\
\lim_{z\to z_{min}} Q_E(t) &= d(s_t) g(z_t)
\end{aligned}$$

This is a fairly general specification for an energy sector, with output (and demand for inputs) increasing in the relative price of energy output to intermediates input, energy output increasing in the energy quality index, and with some simple regularity assumptions on the limiting behaviour at high and low price ratios. This general specification is sufficient to generate the phenomena described in this article, but two specific, not so general, details are required in order to observe these results. The first feature is price taking behaviour. This model assumes infinitesimal profit maximising firms who take prices as given. In these circumstances, the firms cannot internalise the effects of their supply decisions on economy wide prices. This is crucial and trivially important when elucidating a flaw in the allocation that may arise from a decentralised market: we would see nothing interesting if we analysed the social planner's solution (as was done in e.g. Dasgupta and Heal (1974) in their general equilibrium Hotellings model) or if we analysed the solution with a monopoly energy supplier.

The other important specific feature that this energy sector must exhibit is a minimum price for its output relative to its inputs at which it is willing to produce a positive quantity. The specific combination of price taking (which implies infinitesimal firms and decreasing returns to scale) and a minimum price at which these firms are willing to operate, calls for fixed costs in order to prevent productivity rising without bound as the scale of production goes to zero. This seems reasonable in context since it would seem that fixed costs are a realistic feature in the energy industry: the output of oil from the application of a very small quantity of deep water drilling equipment is not high, with decreasing returns to additional units of equipment, rather it is likely to be zero because these additional units of equipment are essential; likewise the output from a wind or solar farm that is disconnected from the grid is zero or very small, and the output from a wind turbine blade or silicon wafer in the absence of the rest of the components is definitely zero.

Finally it is important to note that whilst these features for the energy industry are necessary in order to see the phonomena described in this paper, the phenoma are not a necessary consequence of these features. It is not the case that describing a price taking energy industry with fixed costs is equivalent to assuming the results presented. We shall see that there is a large parameter region in which results are *standard*, and the interesting *new results* are a only exhibited in the presence of a sufficient degree of scale economies in the intermediate goods sector.

The intermediate goods sector is zero profit making, and rents capital stock to create intermediate goods. We write the intermediate goods production function as:

$$Q_t + Q_E(t) = \theta K_t^{\psi}, \quad \psi \geq 1 \tag{1}$$

where ψ determines the degree of returns to scale. θ is a normalisation parameter. Zero profits implies:

$$\begin{aligned} p_Q(t)(Q_t + Q_E(t)) &= p_K(t)K_t \\ \text{i.e. } p_K(t) &= \theta^{\frac{1}{\psi}} p_Q(t)(Q_t + Q_E(t))^{\frac{\psi-1}{\psi}} \end{aligned} \tag{2}$$

There is some evidence to suggest that manufacturing industries behave as if they are subject to increasing returns to scale. Hall (1989) explains the correlation of factor productivity with exogenous demand shocks using increasing returns and finds that increasing returns are particularly evident in the aggregate economy and in manufacturing sectors. Caballero and Lyons (1989) split the returns to scale evident in the aggregate economy into internal, firm level, constant or decreasing returns to scale, and positive external returns to scale. Their best estimate of the degree of scale economies in the US is that a sector which increases its inputs by 10% will see an increase in output of 8%, but if the whole economy increases its inputs by 10% then output will rise by 13%[3]. Basu and Fernald (1997) explain similar data as Caballero and Lyons (1989) as a reallocation effect towards more efficient firms rather than any real increasing returns at the micro-level, but agree that if we model the aggregate economy as a representative firm then increasing returns to scale are appropriate.

To simplify the analysis I consider only steady states of this economy. The consumption Euler equation implies that in steady state the rental rate of capital is a constant given by:

$$p_K^* = \rho + \delta \tag{3}$$

If rental rates are below this steady state rate, then households will be reducing their holdings of capital by saving at a rate that implies overall capital stocks are falling. Whilst if rental rates are above this steady state rate, then households will be increasing their holdings of capital by saving at a rate that implies overall capital stocks are rising. Equation (2) can be used to give a further condition on the rental rate of capital as a function of the energy to intermediates price ratio that can pin down the steady state of the whole economy, where we now drop the time subscripts to indicate that we consider only steady states of the economy:

$$p_K(z,s) = \theta^{\frac{1}{\psi}} A\gamma^{\gamma}\mu^{1-\gamma} z^{\gamma-1} E(z,s)^{\mu+\gamma-1} \left(\frac{\gamma}{\mu} z E(z,s) + Q_E(z,s)\right)^{\frac{\psi-1}{\psi}} \tag{4}$$

Clearly we always have $p_K(z,s) \geq 0$, $\forall z \in [z_{min}, \infty)$ and, by monotonicity of $E(z,s)$ and $Q_E(z,s)$ with respect to z, $p_K(z,s)$ is continuous over this set.

[3] i.e. in the notation of equation (1), this translates as $\psi \sim 1.3$ (ignoring the fact that capital services are only a subset of the whole economy, which is also subject to diseconomies of scale caused by declining energy resource quality). In the empirical trade exercise of Mohler and Seitz (2010), the elasticity of substitution in the CES import demand systems of European economies is found to lie in the range 3 - 5, which corresponds to a returns to scale parameter, $\psi \in (1.25, 1.5)$

3 Steady State Equilibrium

Given some energy quality s, the intersections of the equation (4) with the constant steady state value given by equation (3) defines steady state values, z^*. In this section I show that the parameter space for this economy can be divided into three: one parameter region that correspond to the common understanding of how prices respond to scarcity; one parameter region that can be dismissed as unrealistic, and one interesting *new region* that is the contribution of this paper.

Proposition 1. *Trivially, $K = 0$ is a steady state since, given zero capital stock, production and so investment is zero.*

Proposition 2. *For any given capital stock, $K > 0$, the market equilibrium exists and is unique.*

The equilibrium price ratio is related to capital stock by:

$$\begin{aligned} K &= \left(\frac{Q + Q_E}{\theta}\right)^{\frac{1}{\psi}} = \left(\frac{\frac{\gamma}{\mu} zE(z,s) + Q_E(z,s)}{\theta}\right)^{\frac{1}{\psi}} \\ \text{i.e. } K'(z) &> 0 \text{ , since } \frac{\partial E}{\partial z} \text{ , } \frac{\partial Q_E}{\partial z} > 0 \end{aligned}$$

$z \in (z_{min}(s), \infty)$ represents all possible price ratios that are associated with positive output from the energy sector. Clearly $K(z)$ is a bijection on $z > z_{min}(s)$ and so a given capital stock, K, will imply a particular price, z, by equation (4)[4]. There are therefore no problems of interpretation with a multiplicity of equilibria (though as we shall see, there may be multiple steady states). This monotone relationship between K and z, as well as the relationship already derived, Equation 4, between p_K and z, allows us to construct phase diagrams in (K, C) space based on the consumption Euler equation and the equation of motion for capital. First however, we need to characterise how Equation 4 behaves for different values of ψ.

Proposition 3. *$\exists \psi^{**} > 1$ such that $\psi > \psi^{**} \Rightarrow p_K(z,s) \rightarrow \infty$ as $z \rightarrow \infty$ and $\psi < \psi^{**} \Rightarrow p_K(z,s) \rightarrow 0$ as $z \rightarrow \infty$.*

[4]Note that this is a statement for all K, not just at the steady state K implied by the parameters of the model and by the energy quality index s. The renewables energy sector of section 4 has no energy sector dynamics and so this proposition holds whether or not the macroeconomy has settled into its steady state. However, the fossil fuel energy sector of section 5 has only had its price quantity relationships analysed assuming that it faces constant (i.e. steady state) prices. Therefore the signs of the partial derivatives of E and Q_E for this sector, although they are positive with respect to z as required, have strictly only been evaluated at steady state.

We can evaluate ψ^{**} by taking limits of equation (4):

$$\begin{aligned}
\text{Have } p_K(z,s) &= A\theta^{\frac{1}{\psi}}\gamma^{\gamma}\mu^{1-\gamma}z^{\gamma-1}E(z,s)^{\mu+\gamma-1}\left(\frac{\gamma}{\mu}zE(z,s)+Q_E(z,s)\right)^{\frac{\psi-1}{\psi}} \\
\text{so } \lim_{z\to\infty}\left[p_K(z,s)\right] &= const_1 \times \lim_{z\to\infty}\left[z^{\gamma-1}E(z,s)^{\mu+\gamma-1+\frac{\psi-1}{\psi}}\left(\frac{\gamma}{\mu}z+\frac{Q_E(t)}{E_t}\right)^{\frac{\psi-1}{\psi}}\right] \\
&= const_2 \times \lim_{z\to\infty}\left[z^{\gamma-1+x(\mu+\gamma-1+\frac{\psi-1}{\psi})+\frac{\psi-1}{\psi}}\right]
\end{aligned}$$

$$\begin{aligned}
\text{Therefore } p_K(z,s) \to \begin{matrix}\infty\\0\end{matrix} \text{ as } z\to\infty \quad &\iff \quad \gamma-1+x\left(\mu+\gamma-1+\frac{\psi-1}{\psi}\right)+\frac{\psi-1}{\psi} \gtrless 0 \\
&\iff \quad \psi \gtrless \psi^{**} = \frac{1}{\gamma+\frac{x}{1+x}\mu}
\end{aligned}$$

Proposition 4. *$\exists\psi^* \in (1,\psi^{**})$ such that $\psi > \psi^* \Rightarrow p_K(z,s) \to 0$ as $z \to z_{min}$ and $\psi < \psi^* \Rightarrow p_K(z,s) \to \infty$ as $z \to z_{min}$.*

We can evaluate ψ^* by taking limits of equation (4):

$$\begin{aligned}
\text{Have } p_K(z,s) &= A\theta^{\frac{1}{\psi}}\gamma^{\gamma}\mu^{1-\gamma}z^{\gamma-1}E(z,s)^{\mu+\gamma-1}\left(\frac{\gamma}{\mu}zE(z,s)+Q_E(z,s)\right)^{\frac{\psi-1}{\psi}} \\
\text{so } \lim_{z\to z_{min}}\left[p_K(z,s)\right] &= const_1 \times \lim_{z\to z_{min}}\left[g(z)^{\mu+\gamma-1+\frac{\psi-1}{\psi}}\right]
\end{aligned}$$

$$\begin{aligned}
\text{Therefore } p_K(z,s) \to \begin{matrix}\infty\\0\end{matrix} \text{ as } z\to z_{min} \quad &\iff \quad \mu+\gamma-1+\frac{\psi-1}{\psi} \lesseqgtr 0 \\
&\iff \quad \psi \lesseqgtr \psi^* = \frac{1}{\gamma+\mu} < \psi^{**}
\end{aligned}$$

Whilst these propositions only strictly allow us to characterise $p_K(z,s)$ as $z \to z_{min}$ or $z \to \infty$, the function would have to be very strange to have many turning points. If we make a further regularity assumption (that will be true for the results presented in section 6) that there is at most one turning point, then we can describe the economy as a function of the degree of returns to scale in the intermediate goods sector:

- $\psi > \psi^{**}$ *Super Strong Increasing Returns to Scale* (SSIRS). Since $p_K \to 0$ as $z \to z_{min}$, $p_K \to \infty$ as $z \to \infty$, and given the assumption of a maximum of one turning point, then it must be the case that $p_K(z,s)$ is monotonically increasing in z. Therefore it will only cross the steady state value of p_K^* once (from below) at $z^* > z_{min}$. Given that $K(z)$ is a bijection, there is a single (unstable & repulsive) steady state $K^* > 0$. Equation (4) is graphed for this extreme case of SSIRS in figure 3.1. No stable productive economy (i.e. $Y^* > 0$) exists even with maximal energy resources availability. $K^* = 0$ is stable and attractive since as $K \to 0$ the rental rates paid to capital become insufficient for households to want to save enough to prevent the capital stock decaying away. Returns to scale are too strong for any this model to describe any sensible economic system and we do not consider this case further.

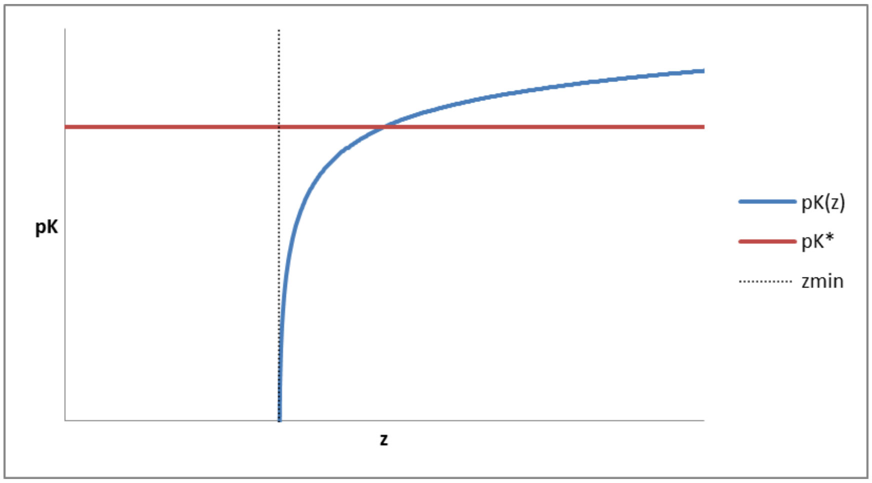

Figure 3.1: SSIRS: two steady states - a stable state at $K^* = 0$ and a higher unstable state.

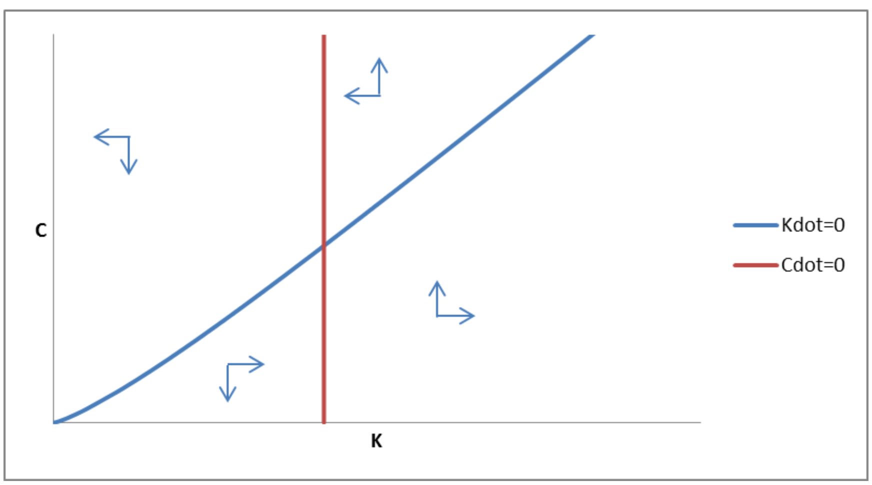

Figure 3.2: Phase diagram for SSIRS, showing the unstable steady state, $K^* > 0$, in (K, C) space

- $1 \leq \psi < \psi^*$ *Weakly Increasing (or Constant) Returns to Scale* (WIRS). Since $p_K \to \infty$ as $z \to z_{min}$ and $p_K \to 0$ as $z \to \infty$, and given the assumption of a maximum of one turning point, then it must be the case that $p_K(z, s)$ is monotonically decreasing in z. Therefore it will only cross the steady state value of p_K^* once (from above) at $z^* > z_{min}$. Given that $K(z)$ is a bijection, there is a single (stable & attractive) steady state $K^* > 0$. Equation (4) is graphed for WIRS in figure 3.3. $K^* = 0$ is unstable and repulsive since as $K \to 0$ the rental rates paid to capital become very large and households to want to save and accumulate capital. This economy accords with our intuitions: a productive economy ($Y^* > 0$) always exists and the response of the economy to a fall in energy quality is for the value of z_{min} to rise, the whole p_K curve to shift to the right, and the equilibrium energy price to intermediates price to rise. This relative energy price rise incentivises the full usage of energy resources by endogenously bringing previously unprofitable marginal resources into use.

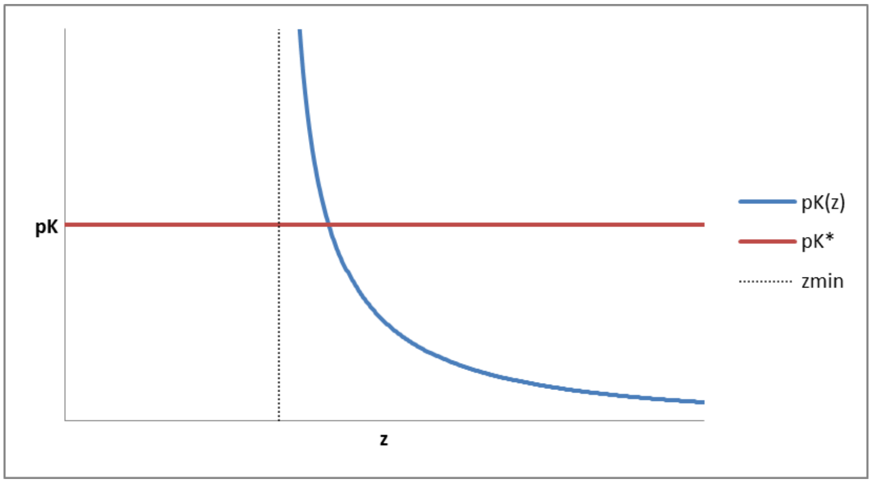

Figure 3.3: WIRS: two steady states - an unstable state at $K^* = 0$ and a stable state at $K^* > 0$.

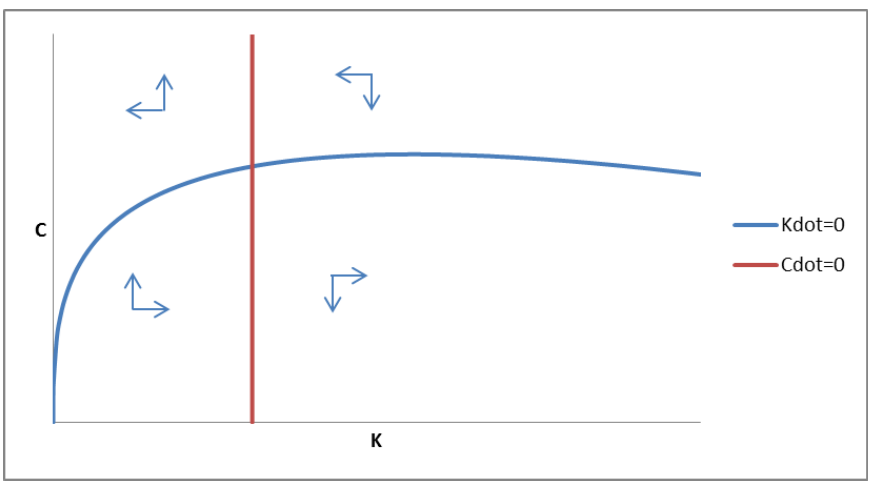

Figure 3.4: Phase diagram for WIRS, showing the stable steady state, $K^* > 0$, in (K, C) space

- $\psi^* < \psi < \psi^{**}$ *Strong Increasing Returns to Scale* (SIRS). Now have $p_K \to 0$ as $z \to z_{min}$ and $p_K \to 0$ as $z \to \infty$, with $p_K > 0$, $\forall z \in (z_{min}, \infty)$. Given the assumption of a maximum of one turning point, then there must indeed be a single turning point at $z^+(s)$ given by $\partial p_K(z^+(s), s)/\partial z = 0$. $K^* = 0$ is stable and attractive since as $K \to 0$ the rental rates paid to capital become insufficient for households to want to save enough to prevent the capital stock decaying away. But, so long as $p_K(z^+(s), s) > p_K*$ then there will be another stable steady state at $z^* > z^+ > z_{min}$ (corresponding to $K^* > 0$). Equation (4) is graphed for SIRS in figure 3.5, showing this stable productive steady state. At the productive equilibrium, this economy may look very similar to the WIRS economy, and may respond to a fall in energy quality in a similar way with relative price rises endogenously bringing previously unprofitable marginal resources into use. However, we cannot prove the

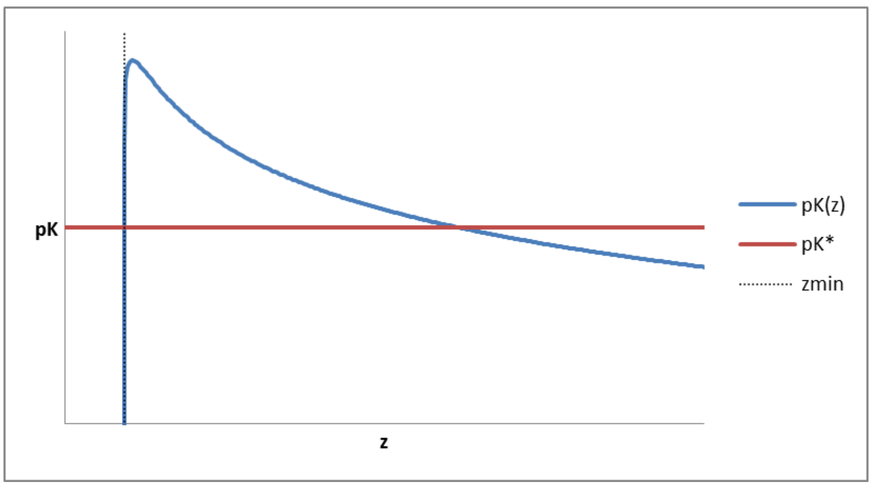

Figure 3.5: SIRS: may have three steady states: stable states at $K^* = 0$ and $K^* > 0$ with an unstable state between.

existence of a non-zero steady state for SIRS. All we know here are the limiting properties that $\lim_{z\to\infty} p_K(z, s) = \lim_{z\to z_{min}} p_K(z, s) = 0$, and that $z_{min} \to \infty$ as $s \to 0$. Therefore it is possible that the graph for Equation (4) for the SIRS economy looks like that shown in figure 3.7, i.e. with a globally stable $K^* = 0$ steady state. Indeed, we can *experimentally* construct a particular SIRS economy then with a particular $s = s_1$ such that the evaluated $p_K(z, s_1)$ function resembles figure 3.5. Then change the energy quality parameter to $s = s_2 < s_1$ such that the evaluated $p_K(z, s_2)$ function resembles figure 3.7, i.e. there is a point in the s parameter space at which the economy *collapses* as the $K^* > 0$ steady state ceases to exist. We can describe the transition from figure 3.5 to figure 3.7 as a collapse because there is a discontinuity in the steady state that the economy can reach. Once we lower the energy quality index, s, past a critical value, the steady state changes discontinuously from $K^* > 0$ to $K^* = 0$. This is unlike WIRS in which the economy exists at some positive level of

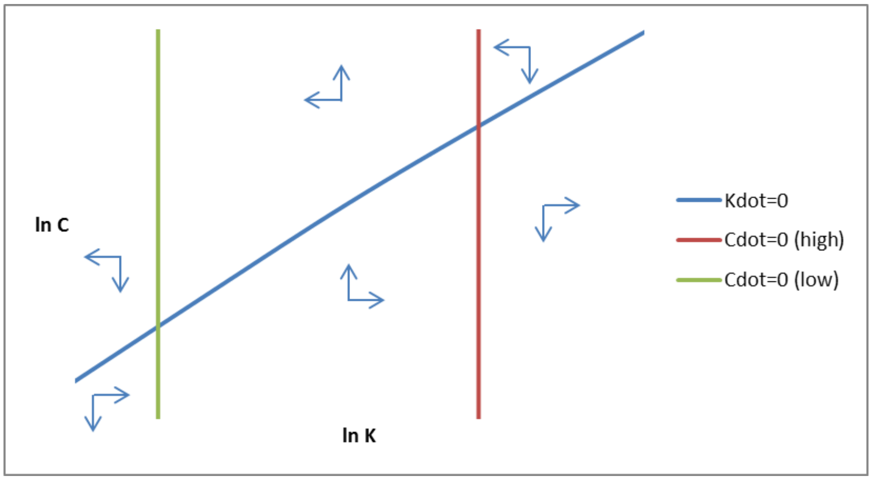

Figure 3.6: A possible phase diagram for SIRS, with 2 steady states, $K^* > 0$, the lower unstable and the higher stable, in (K, C) space. Log scale used because lower steady state close to $K = 0$.

production, irrespective of the severity of the resource restrictions, s, that are imposed.

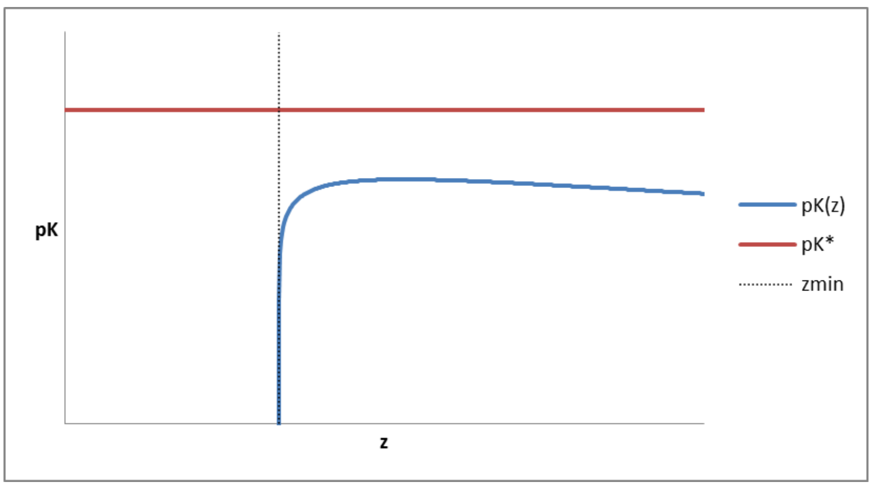

Figure 3.7: SIRS: may only have one stable steady state at $K^* = 0$.

The intuition for what is going on here is straightforward: exploiting energy resources requires intermediate goods as inputs and the scale of the energy sector will exogenously depend upon the quality of resources available, and endogenously upon the relative output to input price. An fall in energy quality is a supply shock to the energy sector which is felt throughout the whole economy. In the absence of significant scale economies in intermediate good production, this supply shock makes energy the scarce and hence expensive commodity, which mitigates the exogenous cause of the problem which was the decline in energy quality. If however scale economies are important in intermediate good production then as the economy contracts due to the effects of the exogenous decline in energy quality, productivity falls by a lot in the intermediate goods sector. This means

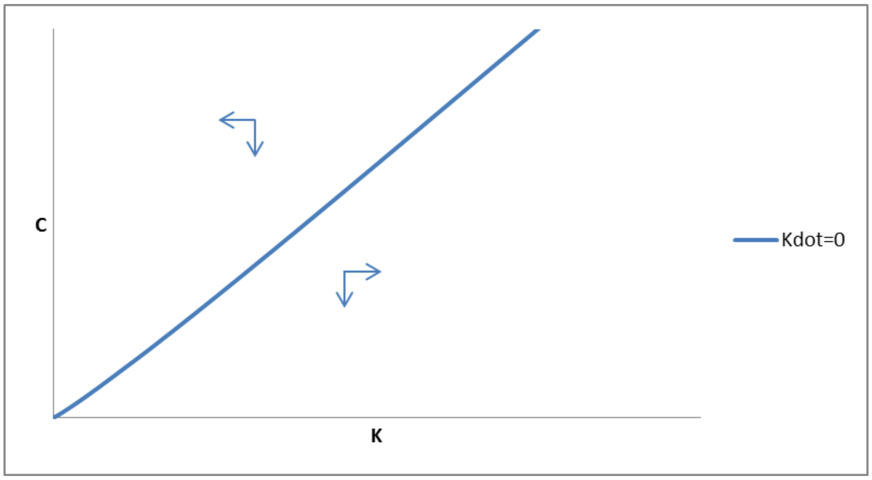

Figure 3.8: A possible phase diagram for SIRS. No steady states with $K > 0$ exists.

that intermediates become relatively scarce and expensive. The price effects move in the opposite direction needed to mitigate the exogenous cause of the problem, this exacerbates the problem by restricting energy production further. These effects multiply and eventually there is no energy price and intermediate goods price which can simultaneously produce positive output from the energy sector and allow the factors supplying the final goods sector to be paid their marginal products, whilst paying capital at the steady state interest rate. In this circumstance, even in the absence of any further declines in energy quality, the interest rate will be below the required rate of return and the economy will run down its capital stock towards the zero capital stock, zero production steady state.

4 A renewable energy sector

A model of some types of renewable energy is perfect for generating an energy sector whose behaviour is consistent with the assumptions made in section 2 and which can be studied *in steady state*. If the resources that are exploited are always there, i.e. next period's resources are not impacted by usage in this period, then there is no trade off across time that this sector needs to make. It will take prices as given now, and make an optimal choice now; the future does not matter. This describes resources like wind or solar resources available at a given site, and not timber or other biomass which needs to be managed with a more lifecycle view. Therefore in the subsequent discussion, read 'resources' perhaps as location specific wind speed or solar flux. It is appropriate to model such resources as being subject to decreasing returns to scale since there is a limit to the energy we can extract from a single location, no matter how much capital we deploy at that location. As previously discussed, it is natural that there is some fixed costs in for operating in any particular location: inada conditions leading to super-productive but miniscule factor inputs are unrealisitic. For any given set of prices, the best resources will be more profitable to exploit than more marginal resources. As the energy to intermediates price rises, more intermediates will be used exploiting a given resource, and marginal resources that previously were not exploited will now be brought into use.

Energy resources are owned by households who auction the right to exploit these resources to a continuum, $[e,\infty)$, of potential energy firms. The households therefore extract all the surplus and own the profit stream that the firms produce. The resulting energy market is competitive (i.e. price taking) with a continuum of differentiated firms, $j \in [e,\infty)$, each producing homogenous output, E_j using intermediate goods Q_j in a decreasing returns to scale production function that also exhibits costs indexed by j i.e. "high j" firms are exploiting poorer quality energy resources than "low j" firms and so, for a given quantity of inputs, Q_j, they produce a lower quantity, E_j of outputs. The production function is:

$$E_j = Q_j^{\beta} - j \ , \quad \beta \in (0,1)$$

Firms maximise profits, $\pi_j = p_E E_j - p_Q Q_j = p_E(E_j - (1/z)Q_j)$, taking prices as given. This gives:

$$Q_j = (\beta z)^{\frac{1}{1-\beta}}$$

This is independent of j i.e. all energy firms use the same quantity of inputs. Therefore profits and energy output are both decreasing in j. j is endogenously defined on $[e,r]$ where $e = 1/s > 0$ is the exogenous parameter representing the highest quality energy resources available, whilst r is an endogenous variable that is defined by $\pi_r = 0$ i.e. there is free entry in the energy sector and firms continue to enter, making positive profits, until the marginal firm makes zero profits. This gives:

$$r = Q_j^{\beta}(1-\beta) = (\beta z)^{\frac{\beta}{1-\beta}}(1-\beta)$$

The total inputs and outputs from the energy sector are calculated by summing over the firms from e to r i.e.

$$\begin{aligned}
E &= \int_e^r E_j \mathrm{d}j = \frac{1}{2}\left(1-\beta^2\right)(\beta z)^{\frac{2\beta}{1-\beta}} - e(\beta z)^{\frac{\beta}{1-\beta}} + \frac{1}{2}e^2 \\
Q_E &= \int_e^r Q_j \mathrm{d}j = (1-\beta)(\beta z)^{\frac{1+\beta}{1-\beta}} - e(\beta z)^{\frac{1}{1-\beta}} \\
\Pi &= \int_e^r \pi_j \mathrm{d}j = p_E(E - \frac{1}{z}Q_E)
\end{aligned}$$

The energy sector uses intermediate goods, and its output responds endogenously to the relationship between the output energy price and the input intermediates price. High quality resources are those which require low inputs per unit of energy produced whereas low quality resources require higher inputs per unit of energy produced. There is no limit imposed upon energy availability, however these unlimited resources will be of increasingly poor quality. If it is optimal to exploit a particular resource, then it is optimal to exploit every resource of higher quality, and so the available high quality resources are always exploited. Exploitation of lower quality resources is an increasing function of the energy to intermediates price ratio.

Appendix 1 shows that this renewables energy sector satisfies the properties of the generic energy sector specified in section 2 and hence that the aggregate economy should have the steady state behaviour of section 3. Section 6 uses this renewables energy sector to generate illustrative results for the economy under regimes of constant, weakly increasing, and strongly increasing returns to scale in the intermediate goods sector.

5 A fossil fuel energy sector

In a Hotelling model of non-renewable resource extraction, the owners of the resources face a trade off between extracting and supplying these resources to market, and leaving the resources in the ground and seeing their price rise. Optimal extraction equates the value of these options, and the basic result is that as a finite resource is extracted, its price should rise to compensate those owners who do not extract immediately. This prediction of a rising price is at odds with the observed price history of non-renewable resources, and some economists e.g. Barnett and Morse (1963), and Simon (1996), have concluded that this price history is evidence of declining rather than increasing scarcity of energy resources. The explanation for this is usually technological advances. However, Hamilton (2011) details the history of global crude oil production over the last century and a half and finds that the production increases have been achieved mainly through the exploitation of new geographic areas, rather than predominantly through technological advances as applied to existing sources. As the scope for adding to production from new geographical areas declines, the suggestion is that the era of rising production could soon end. There are two effects going on: depletion and technological progress; and there is some dispute about which of these effects is "winning".

A set of data that is broadly consistent with Hamilton's interpretation is the energy return on energy invested (EROI) for fossil fuels over the past century (see figure 5.1). EROI can be considered as a *technologically adjusted* index of the cost of obtaining energy resources. So for example oil and gas from 1930 had an EROI of (greater than) $100 : 1$ and so obtaining $100boe$ (barrels of oil equivalent) required *spending* energy (including the energy embodied in the capital used to extract the energy) that contained $\sim 1boe$ so that gross energy production would have had to be $\sim$ 101*boe* to supply the final economy with this 100*boe*. By 2005 oil and gas EROI was $\sim 15 : 1$ and supplying the final economy with 100*boe* would have required gross energy production of $\sim$ 107*boe*. This increase in the cost of supplying the same amount of energy comes despite improvements in technology over the period. Extracting deep water oil in 1930 would not have cost an extra 6% over the oil that was being extracted at that time, rather it would not have been possible at all with the technology available. It is in this sense that EROI can be said to be a technologically adjusted index of the cost of obtaining these resources, and this data suggests that, even allowing for technological advances, the resources that we are extracting are becoming more costly.

Resource	Year	Magnitude (EJ/yr)	EROI (X:1)
Fossil fuels			
Oil and gas	1930	5	>100
Oil and gas	1970	28	30
Oil and gas	2005	9	11 to 18
Discoveries	1970		8
Production	1970	10	20
World oil production	1999	200	35
Imported oil	1990	20	35
Imported oil	2005	27	18
Imported oil	2007	28	12
Natural gas	2005	30	10
Coal (mine-mouth)	1950	n/a	80
Coal (mine-mouth)	2000	5	80
Bitumen from tar sands	n/a	1	2 to 4
Shale oil	n/a	0	5

Figure 5.1: Figure from Murphy and Hall (2010)

However, the total resource of fossil fuels is massive, though of increasingly poor quality. There are enormous quantities of low quality fossil fuel resources, like shale gas, tar sands and brown coal. Figure 5.2 shows that we are a long way from any limits in the availability of fossil fuel resources, notwithstanding any efforts on our part to leave some resources unused because of climate change concerns. Energy resources with higher costs of production (and/or low EROI) tend to be more capital intensive. As an illustrative example we can consider the wooden derricks used for Pennsylvanian oil production in the 19th century against the deep water drilling rigs used today in places like the Gulf of Mexico; or we can compare the pick and shovels used for easily accessible coal seams, to the machinery required for mountain top removal in the Appalachian Mountains. This low quality / high input requirement accords with the intuitive definition of energy quality used in this article. Nuclear energy can also be viewed similarly with finite resources of uranium, but potentially massive resources if 'breeder' reactors are used to recycle the fuel. Again, breeder technology is more expensive and so can be viewed as 'lower quality' in this context. When combined with the possibility of technological advances that lower the cost, and effectively raise the quality, of currently unprofitable marginal resources, these non-renewable resources start to look like renewable resources, with a *regeneration rate* related to the rate of technological progress. This is clearly wrong in the limit, but it may be a good approximation to the fossil fuel energy resources over the next several centuries, and to nuclear energy resources over an even longer term.

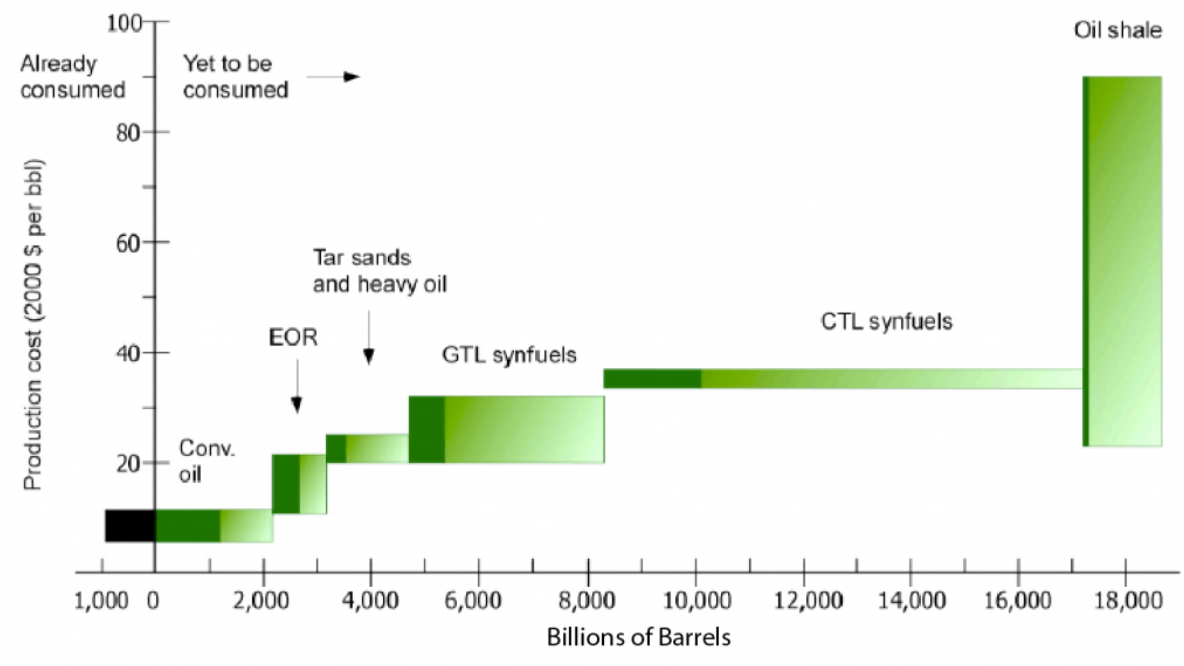

Figure 5.2: Adapted from Brandt and Farrell (2007) by Murphy (2011). Shows resources with their production cost. Proven reserves are dark bands on left, uncertain resources are lighter bands on right.

Given the expectation then (at least from some people) that technology may be sufficient to keep depletion at bay over the medium (and maybe even over the long) term, and the dispute about whether the evidence of the 20th century is consistent with this expectation, we construct a *steady state* Hotellings model of a fossil fuel energy sector in which technological progress and depletion are exactly offset, using the standard framework (see e.g. Beltratti, Chichilnisky, and Heal (1998)) for the lifecycle management of renewable resources like forests and fisheries. At steady state, these resources will again be characterised by an *energy quality* state variable, and we treat the highest quality as a parameter and endogenously allow the exploitation of lower quality resources up to a zero profit limiting case for the marginal resource. The *intellectual experiment* that is then explored is how economies with a different level of highest energy quality available compare with each other.

Again, households own a continuum of energy firms, each indexed by j. Firms j exploits resources labelled S_j by applying inputs Q_j to produce a homogenous good E_j, taking prices as given. There are fixed costs related to the remaining size of the resource that the firm exploits (the rationale being that if high quality resources are exploited first then a large stock means there are high quality resources available, whereas a small stock implies only the poor quality resources remain). The production function and profit flows are given by:

$$\begin{aligned}
E_j(t) &= (Q_j(t) - S_j(t)^{-\alpha})^{\beta} , \quad \beta \in (0,1) , \quad \alpha > 0 \\
\pi_j(t) &= p_E(t)\left(E_j(t) - \frac{1}{z_t}E_j(t)^{\frac{1}{\beta}} - \frac{1}{z_t}S_j(t)^{-\alpha}\right)
\end{aligned}$$

In the renewable stock model, the regeneration rate of the stock is related to the current size of the stock. The simplest analytical way to avoid cornucopian solutions where the stock grows without bound is to have a quadratic equation of motion for the stock:

$$\dot{S}_j(t) = g_{j1}S_j(t) - g_2 S_j(t)^2 - E_j(t)$$

Firms maximise lifetime profits taking prices as given subject to this resource constraint. Profits are discounted at the rate of return available on capital, $r_t = p_K(t) - \delta$. The Hamiltonian of their maximisation problem is therefore:

$$H_t^c = \pi_j(t) + \lambda_t \dot{S}_j(t)$$

With solution conditions

$$\begin{aligned}
\frac{\partial H_t^c}{\partial E_j(t)} = 0 \quad &\Rightarrow \quad \lambda_t = \frac{\partial \pi_t}{\partial E_j(t)} \equiv \pi_E(j,t) \\
\dot{\lambda}_t = r_t\lambda_t - \frac{\partial H_t^c}{\partial S_j(t)} \quad &\Rightarrow \quad \frac{\mathrm{d}\pi_E(j,t)}{\mathrm{d}t} = (r_t - g_{j1} + 2g_2S_j(t))\pi_E(j,t) - \pi_S(j,t)
\end{aligned}$$

Without further assumptions it is difficult to go any further. If however we assume constant prices then we can describe the behaviour of firms in the fossil fuel sector by the coupled differential equations (where the time subscripts remain only because the current value of production and the current value of the stock may be time varying, prices are constant though):

$$\begin{aligned}
\dot{S}_j(t) &= g_{j1}S_j(t) - g_2S_j(t)^2 - E_j(t) \\
\dot{E}_j(t) &= \frac{1}{\pi_{EE}(j,t)}\left((\rho - g_{j1} + 2g_2S_j(t))\pi_E(j,t) - \pi_S(j,t)\right) \\
\rho - g_{j1} + 2g_2S_j(t) &> 0 \quad \text{(for finite lifetime value)}
\end{aligned}$$

Substituting in for the partial derivatives of the profit function, the $\dot{S} = 0$ and $\dot{E} = 0$ loci are given by the equations:

$$\begin{aligned}
\dot{S} = 0 \quad &\Rightarrow \quad E = g_{j1}S - g_2S^2 \\
\dot{E} = 0 \quad &\Rightarrow \quad E = \left(\beta z \left(1 - \frac{1}{z}\alpha S^{-\alpha-1}\frac{1}{\rho - g_{j1} + 2g_2S}\right)\right)^{\frac{\beta}{1-\beta}}
\end{aligned}$$

Because the profit flow tends to minus infinity as the stock tends to zero, the lifetime profit stream associated with the path that leads to steady state is valued more highly than any paths that involve depleting the resource[5]. The system is therefore saddlepath stable, with a phase diagram of the form shown in Figure 5.3.

There can be more than one steady state for a firm in this economy. (In Figure 5.3, lowering z would lower the vertical asymptote of the $\dot{E} = 0$ locus, without changing the $\dot{S} = 0$ locus. This could lead to 3 steady states, with the actual steady state achieved being a function of the initial state.) However, this is perhaps taking the setup of this model too literally. If we were to use this literal interpretation, then we could not talk about the energy quality as a parameter: the energy quality would be an endogenous variable; the primitive parameters are g_{j1} and g_2 which describe how the resources *regenerate*. The point of this exercise is not to construct a theory of the equilibrium output of a firm supplying fossil fuels given fundamental primitives, rather it is to show that a fossil fuel sector can be broadly consistent with the properties of the abstract energy sector laid in in Section 2. Therefore, from now on we treat S_j as a *parameter* that is a property of the firm operating at this level of energy quality, and so implicitly, g_{j1} is a variable that 'adjusts to keep S_j constant'. The actual parameter that is relevant on an economy wide basis is $s = \max(S_j : \forall j)$ and we allow firms with $S_j < s$ to enter until the marginal firm makes zero profits. We can state, by assumption, that s is such that all the firms j exploiting energy resources with quality $S_j \in (S_{min}, s)$ have a unique steady state (with a phase diagram of the form of Figure 5.3) and the position of this steady state is a smooth function of parameters and prices. Given this set up it is important to reiterate the intellectual exercise that is being undertaken: we compare

[5] Although paths with a cycle of resource overuse followed by shut-down and regeneration have not strictly been ruled out.

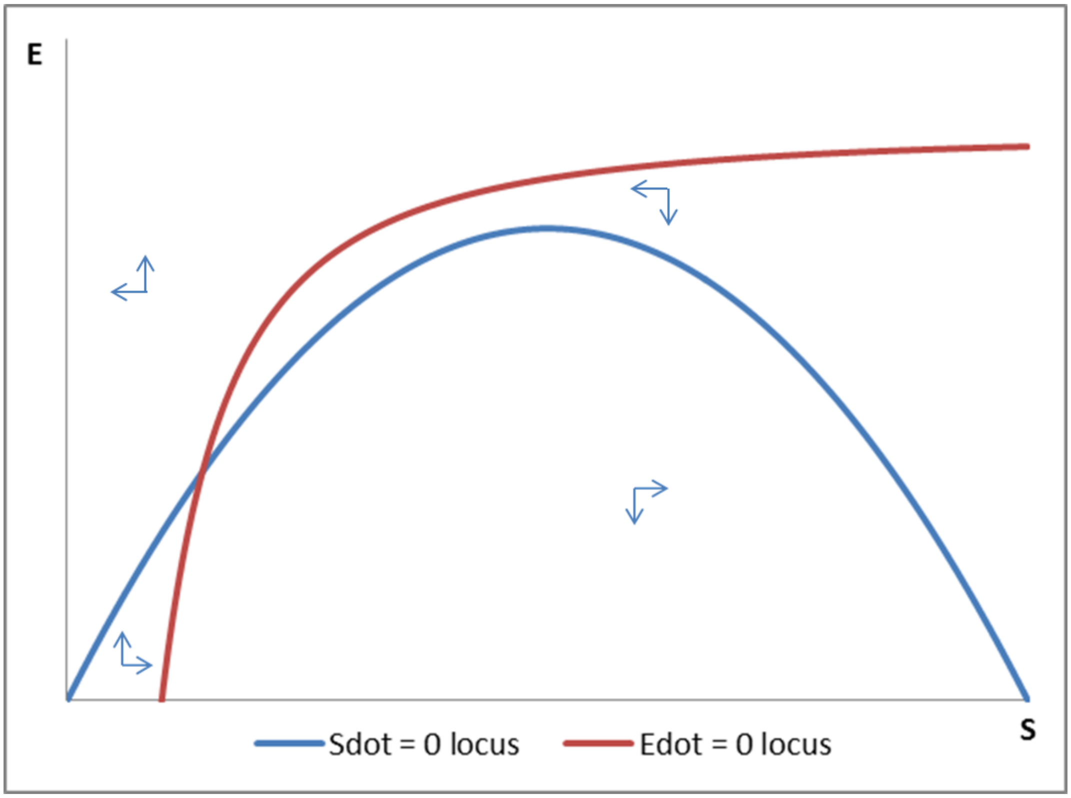

Figure 5.3: Phase diagram for fossil fuel firm, managing its resources under constant prices.

economies with different levels of s, the highest quality available energy. For $s_2 < s_1$ the discussion is framed as if s_2 is the same economy as s_1 but after some *energy quality* shock. However the firms occupying a particular *energy quality location*, S_j are not the same firms (if we assume that the primitive parameters of an individual firm are constant): conditional on g_{j1} a rise in z produces a slight leftwards shift of the $\dot{E} = 0$ locus, and strongly raises its vertical asymptote. In many cases this would produce a fall in E_j. However, when we transform S_j into a parameter that is invariant with respect to prices, we get a well behaved steady state quantity E_j as a function of prices and parameters (including S_j)[6]:

$$G(E_j, S_j, z) = ((\rho + g_2 S_j)S_j - E_j)\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right) - \frac{\alpha}{z} S_j^{-\alpha} = 0 \tag{5}$$

In Appendix 2, as part of proving that this fossil fuel sector satisfies the requirements of the abstract energy sector set out in section 2, it is shown that this relationship implies that the output, E_j, of a firm operating at energy quality, S_j, is an increasing function of the price, z.

The marginal firm in this sector will be the firm that places zero value of operating or not

[6]Just to repeat though, the firm operating at S_j in the economy characterised by s_2 is not the same firm that operated at S_j in the economy characterised by s_1.

operating. Since we are only discussing the steady states of the production of these firms then the zero value condition is the same as the zero profit flow condition i.e. for any given price level, z, there exists a threshold energy quality level:

$$\begin{aligned}
\pi_{min} &= p_E \left(E_{min} - \frac{1}{z} \left(E_{min}^{\frac{1}{\beta}} + S_{min}^{-\alpha} \right) \right) = 0 \\
\text{i.e. } S_{min}(z) &= \left(z E_{min} - E_{min}^{\frac{1}{\beta}} \right)^{-\frac{1}{\alpha}}
\end{aligned}$$

Where E_{min} is related to S_{min} by Equation (5). The aggregate energy sector consists of a continuum of these firms operating over (S_{min}, s) i.e.

$$\begin{aligned}
E &= \int_{S_{min}}^{s} E_j \mathrm{d}S_j \\
Q_E &= \int_{S_{min}}^{s} Q_j \mathrm{d}S_j = \int_{S_{min}}^{s} \left(E_j^{\frac{1}{\beta}} + S_j^{-\alpha} \right) \mathrm{d}S_j \\
\Pi &= \int_{S_{min}}^{s} \pi_j \mathrm{d}S_j = p_E \left(E - \frac{1}{z} Q_E \right)
\end{aligned}$$

It can be shown that S_{min} is a decreasing function of z and so again we have a price taking energy sector that uses intermediate goods such that its output responds endogenously to the relationship between the output energy price and the input intermediates price. Exploitation of lower quality resources is an increasing function of the energy to intermediates price ratio. If it is optimal to exploit a particular resource, then it is optimal to exploit every resource of higher quality, and so the available high quality resources are always exploited. Appendix 2 shows that this fossil fuel energy sector satisfies the properties of the generic energy sector specified in section 2 and hence that the aggregate economy has the steady state behaviour of section 3.

Finally, we consider the behaviour of an infinitesimal (so does not affect the rest of the economy), price taking fossil fuel firm managing a truely non renewable resource. What are the incentives for such a firm as the aggregate economy operates at progressively lower levels of energy quality (though always in the neighbourhood of the steady state if it exists[7])? With constant or weakly increasing returns to scale in the intermediate goods sector, the firm will expect increasing profits (per unit extracted from a given quality of resource) and so will tend to defer extracting resources. Any resources that are not profitable to extract now, will become profitable to extract as the energy

[7]i.e. we consider some exogenous dynamic process for s_t such that $\dot{s}_t < 0$, $z_t \approx z^*(s_t)$ and $r_t \approx \rho$. This putative infinitesimal non renewable resource firm then manages its finite stock of resources under a variable price regime, but the problem is now tractable because it is partial equilibrium since its decisions do not feed back into the economy-wide prices.

quality exploited in the rest of the economy declines. This is in keeping with the sanguine view of non renewable resource economics since Hotelling. However, with strongly increasing returns to scale in the intermediate goods sector, the firm will expect decreasing profits (per unit extracted from a given quality of resource) and so will tend to bring forward the extraction of resources. Any resources that are not profitable to extract now, will never become profitable to extract as the energy quality exploited in the rest of the economy declines. This is a new result not at all in keeping with the usual picture from the non renewable resource literature in economics. Eventually if the economy can no longer maintain a steady state, the interest rate will fall below the rate of time preference. In this circumstance there is some incentive for profitable firms to defer extracting resources since, even though profits (per unit extracted from a given quality of resource) are decreasing, the value placed on future profits is rising as the interest rate falls. It remains the case however that currently unprofitable resources will never be brought 'on-stream'.

6 Illustrative Results

In this section I present illustrative results generated from the model described in sections 2, 3 & 4 i.e. the full macroeconomy with a renewable energy sector. These results illustrate the the mechanism described, but strong increasing returns to scale is contingent upon an extreme parameterisation. To see if it is possible that the phenomena described in this article have any possibility of being quantitatively important, I extend the intermediate goods sector to a monopolistically competitive industry with symmetric firms making heterogenous intermediate goods by renting the capital stock and consuming energy services (i.e. two inputs as opposed to the single input of capital stock in the basic model). These heterogenous goods are aggregated for use in the energy sector and the final goods sector using Dixit-Stiglitz aggregation. This generalised model embeds the basic model[8] and exhibits the same phenomena (but clearly the propositions derived in section 3 only strictly apply to the basic model).

The results are presented here as plots of the steady state value of the index, r, of the marginal, zero profit, energy firm against the energy quality parameter (which is the index of the energy firm exploiting the highest quality resources), $e = 1/s$. Therefore an increase in e on the x-axis corresponds to a reduction in energy quality, and so this is a plot of a comparative static (at steady state) across a continuum of different economies, each having a different energy quality parameter. Different lines on the plot show this comparative static for economies with different levels, ψ, of scale economies in their intermediate goods sectors. The normalisation constant, θ,

[8] The basic model is isomorphic to the generalised model with an elasticity of intermediate sector output with respect to energy equal to zero. However, the generalised model has been set up in a more specific, less general, way, with the microfoundations of monopolistic competition under CES aggregation, rather than just assuming zero profits.

in the intermediate goods sector production function is chosen as a function of the degree of scale economies, $\theta = \theta(\psi)$, so that for $e = 0$ (i.e. $s \to \infty$) the steady state values for the endogenous variables in the model are independent of the level of scale economies, ψ (and therefore the lines on the plot start from a common value, r_0 at $e = 0$) so that the results can be compared on a single plot. If the index of the marginal firm rises (i.e. lower quality resources are exploited) as e rises (i.e. as energy quality falls) then relative energy prices are rising with falling energy quality and the exploitation of lower quality resources expands to (partially) offset the loss of the high quality resources. However, if r falls as e rises then relative energy prices are falling with falling energy quality, and the economy is heading for collapse (in the sense discussed in section 3). This is shown in Figure 6.1 which charts how the steady state of the economy varies with ψ and e. Figure 6.1 does not show timepaths, but we can imagine that if a specific ψ line shows the path of an economy which undergoes a series of depletion shocks[9], then a rising $r(e)$ curve indicates that the use of lower quality resources substitutes for the high quality resources that are no longer available to the economy. A falling $r(e)$ curve indicates that marginal resources are abandoned as high quality resources cease to be available. The parameters used to generate these results, and the results from the following section, are listed in Appendix 3.

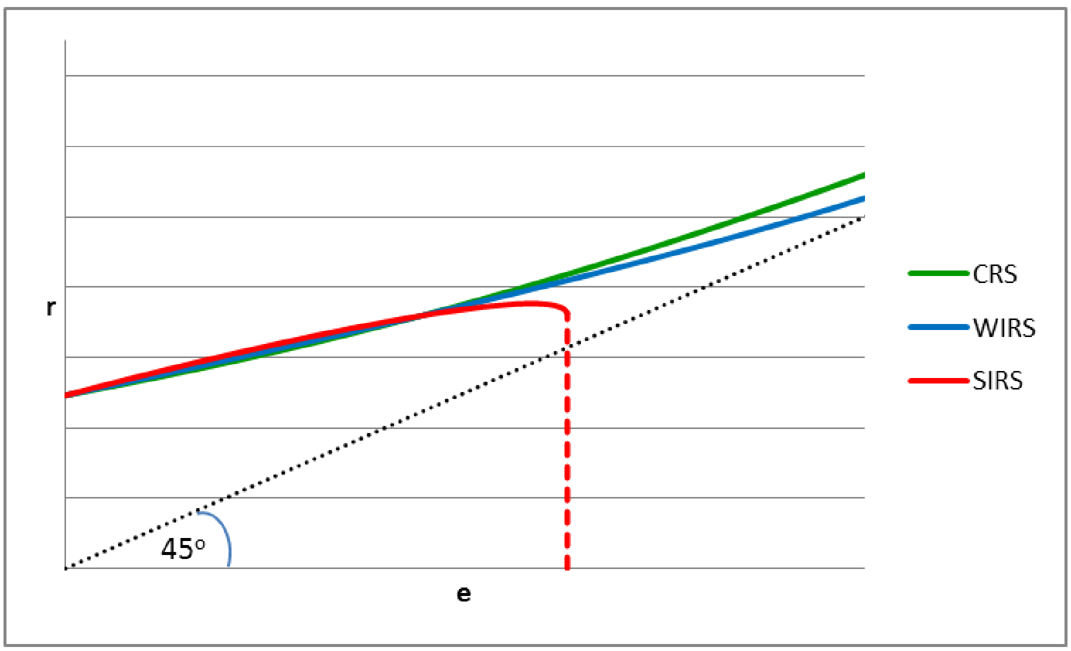

Figure 6.1: $r(e)$ for 3 economies: CRS with $\psi = 1$, WIRS with $1 < \psi < \psi^*$, and SIRS with $\psi > \psi^*$.

Under SIRS, we observe that, initially, the (infinitely) abundant low quality resources provide a substitute for unavailable high quality resources. Eventually as e continues to rise, the $r(e)$ curve

[9] With a point on the graph only generated once the economy has converged to steady state following each depletion shock.

has a turning point and as e rises further the economy starts to abandon the marginal resources despite their abundance, and ultimately the economy collapses. This can be seen more clearly by zooming into figure 6.1 as is shown in Figure 6.2.

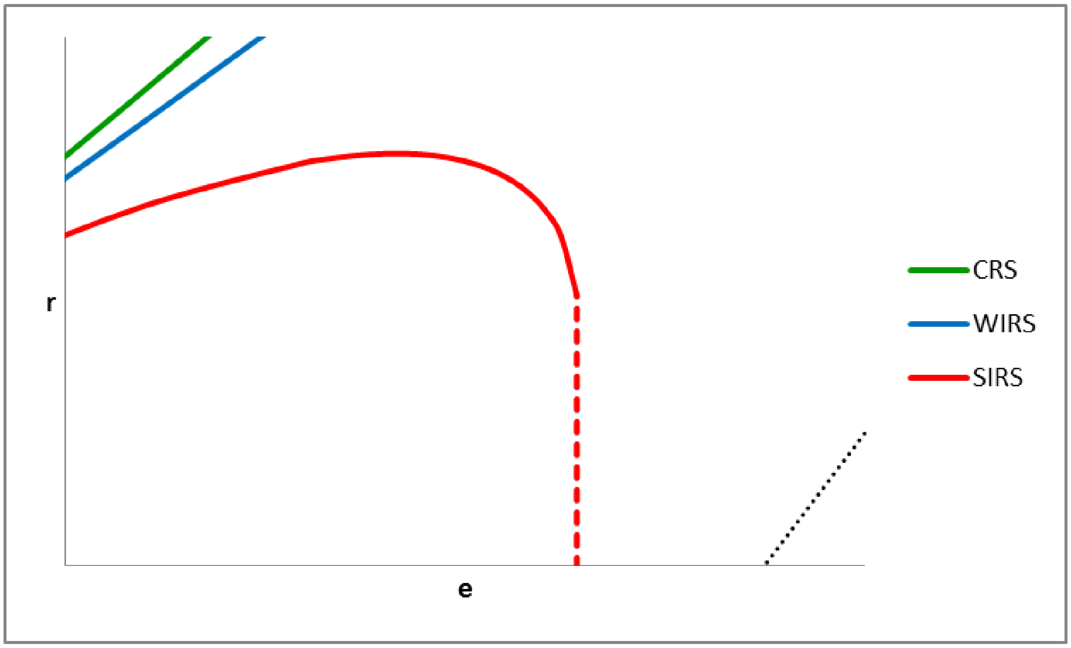

Figure 6.2: As Figure 6.1 but zoomed in.

These results are somewhat unsatisfactory since proposition 4 gives us that $\psi^* = 1/(\gamma + \mu)$. This implies that, given standard estimates from incomes shares, $\psi^* \approx 3$ - vastly higher than any plausible estimate of the degree of scale economies in the real world. Determining whether the phenomena described in this article is an irrelevant feature which real world parameters do not remotely approach, or whether it is worthy of investigating quantitatively, is the purpose of the following model generalisation which essentially adds an energy input requirement to the operation of the intermediate goods sector.

The intermediate goods sector is generalised by splitting it into two. A perfectly competitive *aggregation* sector buys the output of a monopolistically competitive sector which produces heterogenous goods. The aggregation sector has production function and profits as follows:

$$
\begin{aligned}
Q + Q_E &= \left(\int_0^n q_i^{\frac{\sigma-1}{\sigma}} \mathrm{d}i \right)^{\frac{\sigma}{\sigma-1}} \quad , \quad \sigma > 1 \\
\pi_A &= p_Q \left(Q + Q_E \right) - \int_0^n p_i q_i \mathrm{d}i = 0
\end{aligned}
$$

The monopolistically competitive sector consists of measure n (endogenous) firms each producing a differentiated good with some monopoly pricing power. The demand schedule that each monopolist faces, and their production and profit functions are:

$$\begin{aligned} q_i &= (Q + Q_E)\, p_i^{-\sigma} p_Q^{\sigma} \\ q_i &= \phi(E_i^{\eta} K_i^{1-\eta} - f) \\ \pi_i &= p_i q_i - p_E E_i - \rho K_i \end{aligned}$$

Where f is a fixed cost. There is free entry so the profits of each monopolist are driven to zero. The equilibrium conditions for this sector are:

$$\begin{aligned} E_Q &= \Big(\frac{\eta}{1-\eta}\frac{p_K}{p_E}\Big) K \\ p_Q &= \frac{\sigma}{\sigma-1}\frac{p_K}{\phi(1-\eta)}\Big(\frac{\eta}{1-\eta}\frac{p_K}{p_E}\Big)^{\frac{\eta\sigma}{1-\sigma}} (\sigma f)^{\frac{1}{\sigma-1}} K^{\frac{1}{1-\sigma}} \\ Q + Q_E &= (\sigma-1)\phi f \Big(\frac{\eta}{1-\eta}\frac{p_K}{p_E}\Big)^{\frac{\eta\sigma}{\sigma-1}} (\sigma f)^{\frac{\sigma}{1-\sigma}} K^{\frac{\sigma}{\sigma-1}} \end{aligned}$$

The only other change from the model presented previously is that the total output of the energy sector now has to be split across the final and intermediate goods sectors, $E + E_Q$. The fixed cost, f here performs normalisation role as the parameter θ in the basic model: defining $f = f(\sigma)$ allows us to normalise the economies with different σ's so that they all coincide for $e = 0$. Simulating this model (with the parameters detailed in Appendix 3) produces very similar results to the basic model (see Figures 6.3 & 6.4) but now, as detailed in Appendix 3, the degree of returns to scale needed for SIRS and collapse is much lower[10] than in the basic model[11]. This suggests that the SIRS mechanism is not obviously ruled out by the parameterisation needed to observe it, and so this is a phenomenum that is worthy of quantitative investigation.

7 Policy & Innovation

The collapse of a SIRS economy as energy quality declines is due to prices and not to any fundamental limits. By construction, low quality resources are infinitely abundant, and a high relative energy price will ensure that they are profitable to exploit. An energy subsidy will therefore bring resources into production and will increase the scale of intermediates sector, improving the allocation across the economy. The economy does not collapse under CRS, and the market allocation

[10] Depending of the value of the parameter η. The generalised model exhibits SIRS so long as $\sigma < \sigma^* = 1/((1-\eta)(1-\mu-\gamma))$ - see Appendix 4.

[11] The SIRS result, as described in Appendix 3, is generated with a returns to scale, $\psi = 1.3 \iff \sigma = 4.3333$, which is the number estimated by Caballero and Lyons (1989).

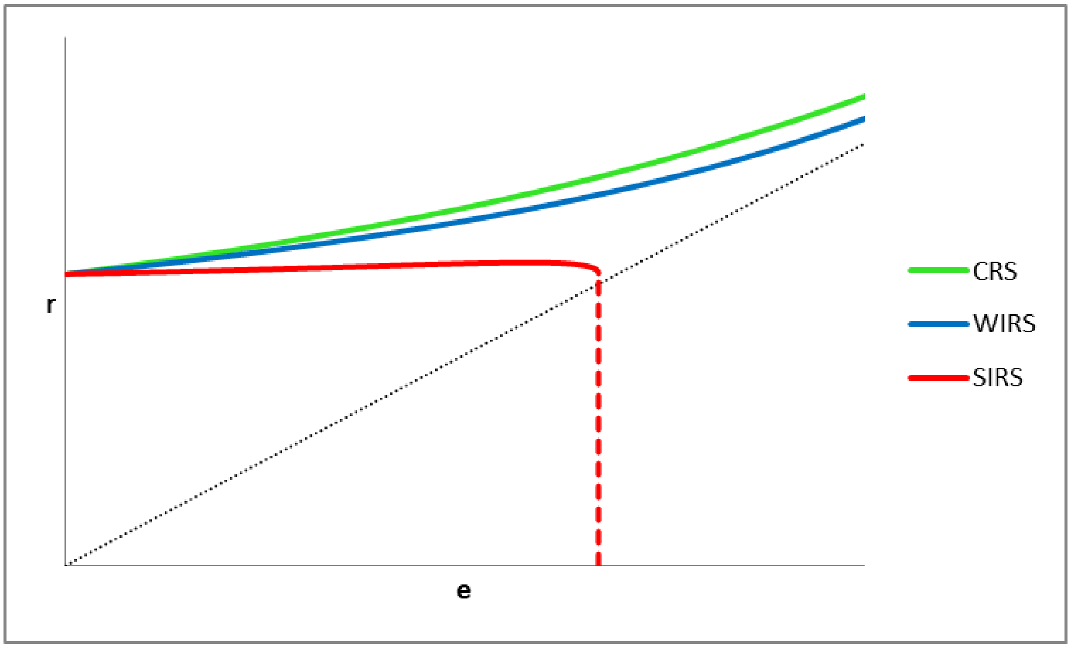

Figure 6.3: $r(e)$ for 3 economies: CRS, WIRS, & SIRS; from the generalised model.

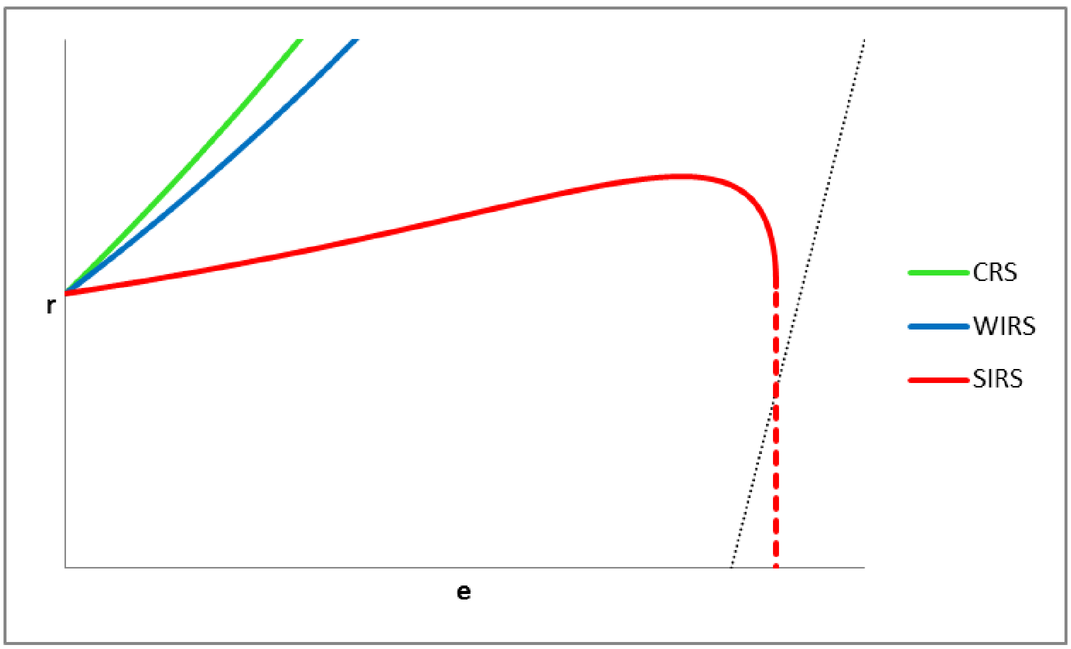

Figure 6.4: As Figure 6.3 but zoomed in.

cannot be improved upon[12]. If we were to impose lump sum taxes on the households and use the proceeds to subsidise energy production, then the pareto-optimal tax rate is zero for CRS and is increasing in the degree of scale economies: the more at risk of collapse the economy is from the interaction of increasing returns and energy quality, the more amenable this situation is to policy

[12]This follows from The First Welfare Theorem.

intervention.

When we imagine technological solutions to energy scarcity, we often think of high technology goods that use energy more efficiently. This is the situation that Aghion, Howitt, Brant-Collett, and Garcaia-Peanalosa (1998) abstract from in their model of endogenous growth with non-renewable resources. A real world example of this could be modern cars with computer optimised engines, as opposed to simpler vehicles that use petrol inputs much less efficiently. In Aghion & Howitt's model, as energy became scarce and expensive there was an increasing incentive to develop this technology. However, in the model presented in this paper I suggest that, if we live in a world of SIRS, then energy scarcity may not motivate us to use this advanced energy efficient technology, because the price of the high-tech computer optimised capital goods could rise by more than the price of energy, and consumers will substitute away from such goods towards goods that use intermediates efficiently but energy inefficiently.

Specifically, we could formulate two alternative production technologies for final goods: an intermediates intensive technology that uses energy very efficiently, and an energy intensive technology that used intermediates efficiently. There would be some price ratio at which these technologies used energy and capital services in the same ratio to produce the same output level, and this price ratio would be the price ratio that the economy switched from one technology to the other. In the results presented in Section 6, rising e always causes rising r in the WIRS economies, and causes rising r in the SIRS economy initially. Rising r occurs because of a rising energy to intermediates price. The switch price will eventually be reached for the WIRS economies and they will ultimately use the energy efficient technology. The switch price may or may not be reached under SIRS, but even supposing that it is, further declines in high quality energy availability could see the switch price being reached again on the way down i.e. the SIRS economy may choose never to take up the energy efficient technology, and even if the economy does adopt it, it may then abandon it. This is intuitive - we can well imagine that productivity in advanced sectors depends on sufficient scale, and if scale is hit hard enough by a shortage of energy, then advanced energy efficient products may not be available.

The same issue arises for a putative backstop technology. We could suppose that some non-depletable backstop was available at some relative energy price. For expositional purposes let us suppose it is large scale deployment of solar panels in deserts. Again the productivity of the sectors that can produce the solar technology depends upon the scale at which it operates. At low levels of scarcity there is large demand for semiconductor technology so this sector is large and productive. It appears that the solar backstop is feasible but the relative energy price is too low to justify its deployment. Energy scarcity rises and energy prices rise (relative to wages). However intermediate goods sectors across the economy contract and productivity falls. The pricing is such that despite the rise in energy prices, we are no closer to profitably deploying the backstop technology. The economy eventually collapses for lack of energy, and at no point was it profitable to deploy the backstop technology. The description here only applies to the SIRS economy, under WIRS, the backstop technology will eventually be deployed.

In general, this story applies to any innovation effort that may allow an economy to grow or continue at the same level under resource restrictions. If the benefits to innovating are positively related to the energy price, but the costs are positively related to the intermediates price, then it will eventually be optimal to undertake the innovation effort under WIRS. It may be the case that it is never be optimal to undertake the innovation effort under SIRS. This problem is amenable to policy intervention though: subsidies can support the scale of industry so that innovations or technologies are within the reach of a SIRS economy, whereas they may be out of reach without policy interventions. This is therefore (theoretical) support (though not necessarily support in any specific case, or for our *real world* economy) for subsidies, e.g. renewables feed-in-tariffs, which may create an industry of sufficient scale to be profitable.

8 Conclusion

I find that the price movements caused by declining energy resources may not be conducive to the exploitation of more marginal resources. Such price movements can lead to macroeconomic collapse, for lack of energy, before all the technologically available energy resources have been exploited. This result is in contrast with the basic Hotellings model and almost all of the non-renewable natural resources and energy literature. Increasing returns to scale, as estimated as occurring in, and often assumed for, manufacturing sectors and industrial economies, is a sufficient condition for this phenomena to be manifest. Innovative or technological solutions to future energy shortages are also adversely affected by this phenomena. However, the more that this phenomena is a real problem, the more it is amenable to policy intervention - which does allow society to mitigate the problem through activist policy.

The mechanism underlying this interaction effect between scale economies and energy quality is that energy supply decisions are positively related to the energy price, and negatively related to input prices. Scale economies can cause productivity in the intermediate sectors, that manufacture inputs for the energy sector, to fall as their scale falls. This can mean that an energy quality shock, which is a supply shock to the whole economy, causes the supply of intermediates to fall by more than the energy supply, and so the scarce commodity is the intermediates. Hence the energy price rise is less than the intermediates price rise and energy sector profitability falls. The existing literature only considers the positive relationship of energy supply with the energy price, and does not consider intermediate inputs at all.

This mechanism may be quantitatively important: the recent experience of historically high energy prices (likely caused by the supply of high quality, low marginal cost, oil resources failing to rise to match rising demand, predominantly from China, see e.g. Kilian and Murphy (2010)) has not led to a uniform pursuit of alternative energy resources. Renewables investment has been volatile[13], worldwide nuclear electricity generation has seen absolute declines[14], and for OECD countries the recent experience of high energy prices is associated with demand destruction rather than supply increases[15]. Obviously there are multiple reasons for these outcomes such as the short run demand side effects of the global financial crisis, as well as policy decisions due to the Fukushima disaster. However, along with high energy prices there have also been some price increases in the industries which supply the renewable energy industry due to higher commodity costs and supply bottlenecks[16]. Energy sector costs rising with the energy price, such that the profitability of marginal suppliers not necessarily improving with a scarcity induced rise in the energy price, is the basic mechanism underlying this article; and so real world experience may be consistent with the phenomena described by this paper.

Future research must develop techniques for testing whether the interaction of scale economies and energy quality is quantitatively important. What are the returns to scale of real manufacturing sectors, especially of those sectors which supply components for extrative industries? Do, for example, growth accounting exercises suggest that we live in a world of constant or weakly increasing returns to scale, or do we live in a world of strongly increasing returns to scale?

[13] See e.g. The GWEC (2011).

[14] See e.g. EPI (2012).

[15] From EIA (2012) OECD petroleum consumption fell by 10% from 2007 to 2011.

[16] See figure 0.3 of EWEA (2009) which shows cost reductions from 1987 to 2004 as technology improved, with cost increases for the final data point in 2006.

References

Aghion, P., P. Howitt, M. Brant-Collett, and C. Garcaia-Peanalosa (1998): *Endogenous Growth Theory.*

Barnett, H., and C. Morse (1963): *Scarcity and Growth.* Johns Hopkins Univ Pr.

Basu, S., and J. Fernald (1997): "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy.*

Beltratti, A., G. Chichilnisky, and G. Heal (1998): "Sustainability: Dynamics and Uncertainty," chap. Sustainable Use of Renewable Resources. Springer.

Brandt, A., and A. Farrell (2007): "Scraping the bottom of the barrel: greenhouse gas emission consequences of a transition to low-quality and synthetic petroleum resources," *Climatic Change.*

Caballero, R., and R. Lyons (1989): "The Role of External Economies in U.S. Manufacturing," *NBER Working Paper.*

Dasgupta, P., and G. Heal (1974): "The optimal depletion of exhaustible resources," *Review of Economic Studies.*

EIA (2012): "O.E.C.D. Petroleum Consumption Data," Data query, The US Energy Information Administration, www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=5&pid=5&aid=2&cid=CG5,&syid =2007&eyid=2011&unit=TBPD.

EPI (2012): "World Nuclear Electricity Generation Down 5 Percent Since 2006," Data hightlights, The Earth Policy Institute, www.earth-policy.org/data_highlights/2012/highlights33.

EWEA (2009): "The Economics of Wind Energy," Report, European Wind Energy Association, www.windenergie.nl/sites/windenergie.nl/files/documents/the_economics_of_windenergy_ewea.pdf.

GWEC (2011): "Global Wind Report 2010," Report, The Global Wind Energy Council, www.gwec.net/publications/global-wind-report-2/global-wind-report-2011/.

Hall, R. (1989): "Invariance Properties of Solow's Productivity Residual," *NBER Working Paper.*

Hamilton, J. (2011): "Oil Prices, Exhaustible Resources, and Economic Growth," *Prepared for 'Handbook of Energy and Climate Change'.*

HOLLAND, S. (2008): "Modeling Peak Oil," *The Energy Journal*.

HOTELLING, H. (1931): "The Economics of Exhaustible Resources," *The Journal of Political Economy*.

KILIAN, L., AND D. MURPHY (2010): "The role of inventories and speculative trading in the global market for crude oil," *CEPR Discussion Papers*.

MOHLER, L., AND M. SEITZ (2010): "The Gains from Variety in the European Union," *Munich Discussion Paper http://epub.ub.uni-muenchen.de/11477/*.

MURPHY, D., AND C. HALL (2010): "Year in review - EROI or energy return on (energy) invested," *Annals of The New York Academy of Sciences*.

MURPHY, K., A. SCHLEIFER, AND R. VISHNY (1989): "Industrialization And The Big Push," *Journal of Political Economy*.

MURPHY, T. (2011): "Peak Oil Perspective," Blog post, physics.ucsd.edu/do-the-math/2011/11/peak-oil-perspective.

SIMON, J. L. (1996): *The Ultimate Resource 2*. Princeton University Press.

VENTURA, J. (2005): "A global view of economic growth," chap. in: P. Aghion and S. Durlauf (ed.) "Handbook of Economic Growth".

A Appendix 1: The renewable energy sector satisfies requirements of section 2

Throughout we refer to:

$$\begin{aligned} e(s) &= s^{-1} \\ r(z) &= (\beta z)^{\frac{\beta}{1-\beta}}(1-\beta) \\ E &= \frac{1}{2}\left(1-\beta^2\right)(\beta z)^{\frac{2\beta}{1-\beta}} - e(\beta z)^{\frac{\beta}{1-\beta}} + \frac{1}{2}e^2 \\ Q_E &= (1-\beta)(\beta z)^{\frac{1+\beta}{1-\beta}} - e(\beta z)^{\frac{1}{1-\beta}} \end{aligned}$$

Proposition A. 1. 1. *$\exists z_{min}(e)$ s.t. $E(z_{min}(e), e) = 0$ and $Q_E(z_{min}(e), e) = 0$ with $\frac{\mathrm{d}z_{min}(s)}{\mathrm{d}s} = \frac{\mathrm{d}z_{min}(e)}{\mathrm{d}e}\frac{\mathrm{d}e}{\mathrm{d}s} < 0 \iff \frac{\mathrm{d}z_{min}(e)}{\mathrm{d}e} > 0$ (since $e = s^{-1}$) and $z_{min} \to \infty$ as $e \to \infty$.*

Proof. z_{min} is defined by the marginal firm being the only firm i.e. $r(z_{min}) = (1-\beta)(\beta z)^{\frac{\beta}{1-\beta}} = e$, so that no energy industry exists.

$$z_{min} = \frac{1}{\beta}\left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}} e^{\frac{1-\beta}{\beta}}$$

Plugging this expression into the equations for E and Q_E gives zero as required (though this is by definition), and we can immediately see that z_{min} is an increasing function of e and that $z_{min} \to \infty$ as $e \to \infty$. □

Proposition A. 1. 2. $\frac{\partial E}{\partial z} > 0$

Proof.

$$\begin{aligned} \frac{\partial E}{\partial z} &= \frac{\beta^2}{1-\beta}(\beta z)^{\frac{2\beta-1}{1-\beta}}\left((1-\beta^2)(\beta z)^{\frac{\beta}{1-\beta}} - e\right) > 0 \\ \text{so long as } z &> \frac{1}{\beta}\left(\frac{e}{(1+\beta)(1-\beta)}\right)^{\frac{1-\beta}{\beta}} \\ \text{true, since: } z &> z_{min} = \frac{1}{\beta}\left(\frac{e}{1-\beta}\right)^{\frac{1-\beta}{\beta}} > \frac{1}{\beta}\left(\frac{e}{(1+\beta)(1-\beta)}\right)^{\frac{1-\beta}{\beta}} \end{aligned}$$

□

Proposition A. 1. 3. $\frac{\partial Q_E}{\partial z} > 0$

Proof.

$$\begin{aligned}
\frac{\partial Q_E}{\partial z} &= \frac{\beta}{1-\beta}(\beta z)^{\frac{\beta}{1-\beta}}\left((1-\beta^2)(\beta z)^{\frac{\beta}{1-\beta}}-e\right)>0 \\
\text{so long as } z &> \frac{1}{\beta}\left(\frac{e}{(1+\beta)(1-\beta)}\right)^{\frac{1-\beta}{\beta}} \\
\text{true, since: } z &> z_{min}=\frac{1}{\beta}\left(\frac{e}{1-\beta}\right)^{\frac{1-\beta}{\beta}}>\frac{1}{\beta}\left(\frac{e}{(1+\beta)(1-\beta)}\right)^{\frac{1-\beta}{\beta}}
\end{aligned}$$

□

Proposition A. 1. 4. $\lim_{z\to\infty} E = a(e)z^x,\ x\geq 0$

Proof. Clearly $\lim_{z\to\infty} E=\frac{1}{2}\left(1-\beta^2\right)\beta^{\frac{2\beta}{1-\beta}}z^{\frac{2\beta}{1-\beta}}$, so $x=\frac{2\beta}{1-\beta}>0$ and $a(e)=\frac{1}{2}\left(1-\beta^2\right)\beta^{\frac{2\beta}{1-\beta}}$. □

Proposition A. 1. 5. $\lim_{z\to\infty}\left(\frac{Q_E}{E}\right)=b(e)z^y,\ 0\leq y\leq 1$

Proof.

$$\begin{aligned}
\text{Have } \lim_{z\to\infty} Q_E &= (1-\beta)\,(\beta z)^{\frac{1+\beta}{1-\beta}} \\
\text{so } \lim_{z\to\infty}\left[\frac{Q_E}{E}\right] &= \frac{2\beta}{(1+\beta)}z
\end{aligned}$$

So $b(e)=\frac{2\beta}{1+\beta}$ and $y=1\in[0,1]$. □

Proposition A. 1. 6. $\lim_{z\to z_{min}} E = c(e)g(z),\ g(z_{min})=0$

Proof.

$$\begin{aligned}
\text{Have } z_{min} &= \frac{1}{\beta}\left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}}e^{\frac{1-\beta}{\beta}}\Rightarrow e=(1-\beta)(\beta z_{min})^{\frac{\beta}{1-\beta}} \\
\text{so } E &= \frac{1}{2}\left(1-\beta^2\right)\left((\beta z)^{\frac{\beta}{1-\beta}}-(\beta z_{min})^{\frac{\beta}{1-\beta}}\right)\left((\beta z)^{\frac{\beta}{1-\beta}}-\frac{1-\beta}{1+\beta}(\beta z_{min})^{\frac{\beta}{1-\beta}}\right) \\
\text{i.e. } \lim_{z\to z_{min}} E &= \beta e\left((\beta z)^{\frac{\beta}{1-\beta}}-(\beta z_{min})^{\frac{\beta}{1-\beta}}\right)
\end{aligned}$$

So $g(z)=(\beta z)^{\frac{\beta}{1-\beta}}-(\beta z_{min})^{\frac{\beta}{1-\beta}}$, with $g(z_{min})=0$, and $c(e)=\beta e$. □

Proposition A. 1. 7. $\lim_{z\to z_{min}} Q_E = d(e)g(z),\ g(z_{min})=0$

Proof.

$$\begin{aligned}
Q_E &= (1-\beta)(\beta z)^{\frac{1}{1-\beta}}\left((\beta z)^{\frac{\beta}{1-\beta}}-(\beta z_{min})^{\frac{\beta}{1-\beta}}\right) \\
\text{i.e. } \lim_{z\to z_{min}} Q_E &= \left(\frac{e}{1-\beta}\right)^{\frac{1}{\beta}}\left((\beta z)^{\frac{\beta}{1-\beta}}-(\beta z_{min})^{\frac{\beta}{1-\beta}}\right)
\end{aligned}$$

So $g(z) = (\beta z)^{\frac{\beta}{1-\beta}} - (\beta z_{min})^{\frac{\beta}{1-\beta}}$, with $g(z_{min}) = 0$, and $d(e) = \left(\frac{e}{1-\beta}\right)^{\frac{1}{\beta}}$. $\square$

B Appendix 2: The fossil fuel energy sector satisfies requirements of section 2

Lemma A. 2. 1. $\frac{\mathrm{d}S_{min}}{\mathrm{d}z} < 0$.

Proof. For a given price level, $p_E(1)$ & z_1, $S_{min}(z_1)$ satisfies:

$$\pi_{min} = p_E(1)\left(E_{min} - \frac{1}{z_1}E_{min}^{\frac{1}{\beta}} - \frac{1}{z_1}S_{min}^{-\alpha}\right) = 0$$

If prices change to $z_2 > z_1$ then clearly the firm operating at S_{min} clearly has the option of keeping quantities of inputs constant, continuing to produce E_{min}, and so make strictly positive profits. Consequently there will be at least one firm with $S_M < S_{min}$ which now finds it profitable to start producing i.e. with a rise in z there is a new, lower quality firm. Equivalently, $\frac{\mathrm{d}S_{min}}{\mathrm{d}z} < 0$. □

Proposition A. 2. 1. *$\exists z_{min}(s)$ s.t. $E(z_{min}(s), s) = 0$ and $Q_E(z_{min}(s), s) = 0$ with $\frac{\mathrm{d}z_{min}(s)}{\mathrm{d}s} < 0$ and $z_{min} \to \infty$ as $s \to 0$.*

Proof. z_{min} is defined by the marginal firm being the only firm i.e. $S_{min}(z_{min}(s)) = s$, so that no energy industry exists. Aggregate industry inputs and outputs are defined as an integral over the range $[S_{min}, s]$. At $z = z_{min}$ this interval has zero length and so $E(z_{min}(s), s) = 0$ and $Q_E(z_{min}(s), s) = 0$ are trivially satisfied.

$$\begin{aligned} \text{Have } S_{min}(z_{min}(s)) &= s \\ \text{so } \frac{\mathrm{d}S_{min}(z_{min}(s))}{\mathrm{d}s} &= \left.\frac{\mathrm{d}S_{min}(z)}{\mathrm{d}z}\right|_{z=z_{min}} \frac{\mathrm{d}z_{min}}{\mathrm{d}s} = 1 \\ \text{i.e. } \frac{\mathrm{d}z_{min}}{\mathrm{d}s} &= 1\Big/\left.\frac{\mathrm{d}S_{min}(z)}{\mathrm{d}z}\right|_{z=z_{min}} \end{aligned}$$

Lemma 1 shows that $\frac{\mathrm{d}S_{min}}{\mathrm{d}z} < 0$ in general, and so in particular, $\frac{\mathrm{d}S_{min}}{\mathrm{d}z}|_{z=z_{min}} < 0$. Therefore $\frac{\mathrm{d}z_{min}(s)}{\mathrm{d}s} < 0$ as req. Taking the limit of the Equation (5) for the marginal firm as $s \to 0$ (the z to use for the marginal firm is z_{min} and the energy output is zero in order to be on the $\dot{S} = 0$ locus at $S_j = s = 0$) gives:

$$\frac{\alpha}{z_{min}} = s^{\alpha}\left(\left(\rho + g_2 s\right) s\right) \times 1 \to 0 \text{ as } s \to 0 \Rightarrow z_{min} \to \infty \text{ as } s \to 0$$

i.e. $z_{min} \to \infty$ as $s \to 0$. □

Lemma A. 2. 2. $G_z \equiv \frac{\partial G}{\partial z} > 0$

Proof.

$$\begin{aligned} G(E_j, S_j, z) &= ((\rho + g_2 S_j)S_j - E_j)\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right) - \frac{\alpha}{z} S_j^{-\alpha} \\ \frac{\partial G}{\partial z} &= \beta\left((\rho + g_2 S_j)S_j - E_j\right) E_j^{\frac{1-\beta}{\beta}} (\beta z)^{-2} + \frac{\alpha}{z^2} S_j^{-\alpha} > 0 \end{aligned}$$

□

Assumption A. 2. 1. *Let*

$$0 < \beta < \frac{1}{1 + \frac{z^*(s_i)}{\alpha} S_j^{\alpha} E_j \left(1 - (\beta z^*(s_i))^{-1} E_j^{\frac{1-\beta}{\beta}}\right)} = B(i,j)$$

$\forall S_j \in [S_{min}(z^*(s_i)), s_i]$ *and* $\forall s_i \in (0, s]$ *where* E_j *is given by* $G(E_j, S_j, z^*(s_i)) = 0$ *and* $z^*(s_i) \equiv$ *equilibrium price ratio given the energy quality parameter* s_i.

This is an uncontroversial regularity assumption. β is the curvature of the energy production function which is already assumed to be in the interval $(0, 1)$. This assumption just narrows the interval somewhat, since we can easily show that the denominator is greater than 1:

$$\begin{aligned} G(E_j, S_j, z) = 0 \quad &\Rightarrow \quad ((\rho + g_2 S_j)S_j - E_j)\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right) = \frac{\alpha}{z} S_j^{-\alpha} > 0 \\ \text{Have } \rho - g_{j1} + 2g_2 S_j(t) > 0 \quad &\text{and} \quad g_{j1} = \frac{E_j}{S_j} + g_2 S_j \\ \text{Combining} \quad &\Rightarrow \quad (\rho + g_2 S_j)S_j - E_j > 0 \\ &\Rightarrow \quad 1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}} > 0 \\ &\Rightarrow \quad 0 < B(i,j) = \frac{1}{1 + \frac{z^*(s_i)}{\alpha} S_j^{\alpha} E_j \left(1 - (\beta z^*(s_i))^{-1} E_j^{\frac{1-\beta}{\beta}}\right)} < 1 \end{aligned}$$

Lemma A. 2. 3. $G_E \equiv \frac{\partial G}{\partial E_j} < 0$

Proof.

$$\begin{aligned} G(E_j, S_j, z) &= ((\rho + g_2 S_j)S_j - E_j)\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right) - \frac{\alpha}{z} S_j^{-\alpha} \\ \frac{\partial G}{\partial E_j} &= -1 - \frac{1}{\beta^2 z} E_j^{\frac{1-\beta}{\beta}} \left((\rho + g_2 S_j)(1 - \beta)\frac{S_j}{E_j} - 1\right) \end{aligned}$$

$$\text{i.e. } (\rho + g_2 S_j)(1 - \beta)\frac{S_j}{E_j} > 1 \Rightarrow \frac{\partial G}{\partial E_j} < 0$$

$$
\begin{aligned}
(\rho + g_2 S_j)(1-\beta)\frac{S_j}{E_j} &> 1 \\
\Rightarrow (\rho + g_2 S_j)S_j(1-\beta) - E_j &> 0
\end{aligned}
$$

$$
\begin{aligned}
\text{Have } (\rho + g_2 S_j)S_j - E_j &= \frac{\alpha}{z} S_j^{-\alpha}\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right)^{-1} \\
\text{so } (\rho + g_2 S_j)S_j(1-\beta) - E_j &= \frac{\alpha}{z} S_j^{-\alpha}\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right)^{-1} - \beta(\rho + g_2 S_j)S_j
\end{aligned}
$$

$$
\begin{aligned}
\text{i.e. } \beta &< \frac{\alpha}{z} S_j^{-\alpha}\left(\left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right)((\rho + g_2 S_j)S_j)\right)^{-1} \\
&= \frac{1}{1 + \frac{z}{\alpha} S_j^{\alpha} E_j \left(1 - (\beta z)^{-1} E_j^{\frac{1-\beta}{\beta}}\right)}
\end{aligned}
$$

$\Rightarrow (\rho + g_2 S_j)S_j(1-\beta) - E_j > 0$ i.e. Assumption 1 ensures that $\frac{\partial G}{\partial E_j} < 0$. □

Proposition A. 2. 2. $\frac{\partial E}{\partial z} > 0$

Proof.

$$
\begin{aligned}
G(E_j, S_j, z) = 0 \quad \Rightarrow \quad \frac{\partial E_j}{\partial z} &= -\frac{G_z}{G_E} \\
\text{i.e. } \frac{\partial E_j}{\partial z} &> 0 \quad \text{by Lemmas 2 \& 3}
\end{aligned}
$$

$$
\text{Aggregate energy production, } E(z) = \int_{S_{min}(z)}^{s} E_j(z)\mathrm{d}S_j
$$

Every element of this sum is rising as z rises. The upper limit of the sum is constant with respect to z. The lower limit falls with rising z (by Lemma 1). Therefore we clearly have $\frac{\partial E}{\partial z} > 0$. □

Proposition A. 2. 3. $\frac{\partial Q_E}{\partial z} > 0$

Proof.

$$
\begin{aligned}
Q_j(z) &= E_j^{\frac{1}{\beta}} + S_j^{-\alpha} \\
\text{i.e. } \frac{\partial Q_j}{\partial z} &= \frac{1}{\beta} E_j^{\frac{1-\beta}{\beta}} \frac{\partial E_j}{\partial z} > 0 \\
Q_E &= \int_{S_{min}(z)}^{s} Q_j(z)\mathrm{d}S_j
\end{aligned}
$$

Again, every element of this sum is rising as z rises. The upper limit of the sum is constant with respect to z. The lower limit falls with rising z (by Lemma 1). Therefore we clearly have $\frac{\partial Q_E}{\partial z} > 0$. $\square$

Proposition A. 2. 4. $\lim_{z\to\infty} E = a(s)z^x,\ x \geq 0$

Proof. Clearly $\lim_{z\to\infty} E_j = (\rho + g_2 S_j)S_j$ (from Equation (5)) and $\lim_{z\to\infty} S_{min} = 0$. Therefore:

$$\lim_{z\to\infty} E = \int_0^s (\rho S_j + g_2 S_j^2)\mathrm{d}S_j = \frac{1}{2}\rho s^2 + \frac{1}{3} g_2 s^3$$

i.e. $a(s) = \frac{1}{2}\rho s^2 + \frac{1}{3} g_2 s^3$ and $x = 0$. $\square$

Proposition A. 2. 5. $\lim_{z\to\infty} \left(\frac{Q_E}{E}\right) = b(s)z^y,\ 0 \leq y \leq 1$

Proof.

$$\begin{aligned} \text{Have } Q_j &= E_j^{\frac{1}{\beta}} + S_j^{-\alpha} \to (\rho S_j + g_2 S_j^2)^{\frac{1}{\beta}} + S_j^{-\alpha} \text{ as } z \to \infty \\ \text{so } \lim_{z\to\infty}\left[\frac{Q_E}{E}\right] &= \frac{\int_0^s((\rho S_j + g_2 S_j^2)^{\frac{1}{\beta}} + S_j^{-\alpha})\mathrm{d}S_j}{\frac{1}{2}\rho s^2 + \frac{1}{3} g_2 s^3} = b(s) \end{aligned}$$

So $b(s)$ is as above and $y = 0 \in [0, 1]$. $\square$

Proposition A. 2. 6. $\lim_{z\to z_{min}} E = c(s)g(z),\ g(z_{min}) = 0$

Proof.

$$\lim_{z\to z_{min}} E = E_s(z_{min}) \times (s - S_{min}(z))$$

So $g(z) = s - S_{min}(z)$ gives the result. $E_s(z_{min})$ is a strictly positive constant giving the energy output defined by the zero profit condition for the firm operating at s. $\square$

Proposition A. 2. 7. $\lim_{z\to z_{min}} Q_E = d(s)g(z),\ g(z_{min}) = 0$

Proof.

$$\lim_{z\to z_{min}} Q_E = Q_s(z_{min}) \times (s - S_{min}(z))$$

So $g(z) = s - S_{min}(z)$ gives the result. $Q_s(z_{min})$ is a strictly positive constant giving the intermediate inputs defined by the zero profit condition for the firm operating at s. $\square$

C Appendix 3: Parameters of simulated economies

Parameter	Basic Model	Generalised Model
A	1	1
β	0.5	0.5
μ	0.111111111	0.04
γ	0.222222222	0.293333333
δ	0.1	0.1
ρ	0.05	0.05
η	n/a	0.7
φ	n/a	1
ψ	Used	Equivalent
CRS	1	1
WIRS	2.9	1.111111111
ψ*	3	1.25
SIRS	3.1875	1.3
ψ**	3.375	1.777251185
σ	Equivalent	Used
CRS	∞	∞
WIRS	1.526315789	10
σ*	1.5	5
SIRS	1.457142857	4.333333333
σ**	1.421052632	2.286585366

D Appendix 4: SIRS in generalised model

Start with zero profits in the intermediate goods market and substitute:

$$\begin{aligned}
p_K K + p_E E_Q &= p_Q(Q+Q_E) \\
\text{i.e. } \frac{p_K K}{1-\eta} &= p_Q(Q+Q_E) \\
\text{but } K &= \sigma f\left[(\sigma-1)\phi f\right]^{\frac{1-\sigma}{\sigma}}\left(\frac{\eta}{1-\eta}\frac{p_K}{p_E}\right)^{-\eta}(Q+Q_E)^{\frac{\sigma-1}{\sigma}} \\
\text{so } p_K &= \frac{1-\eta}{\sigma f}\left[(\sigma-1)\phi f\right]^{\frac{\sigma-1}{\sigma}}\left(\frac{\eta}{1-\eta}\frac{p_K}{p_E}\right)^{\eta}(Q+Q_E)^{\frac{1-\sigma}{\sigma}}p_Q(Q+Q_E) \\
&= \frac{1-\eta}{\sigma f}\left[(\sigma-1)\phi f\right]^{\frac{\sigma-1}{\sigma}}\left(\frac{\eta}{1-\eta}\right)^{\eta}\left(\frac{p_K}{z}\right)^{\eta}p_Q^{1-\eta}(Q+Q_E)^{1+\frac{1-\sigma}{\sigma}} \\
\text{i.e. } p_K^{1-\eta} &= \frac{1-\eta}{\sigma f}\left[(\sigma-1)\phi f\right]^{\frac{\sigma-1}{\sigma}}\left(\frac{\eta}{1-\eta}\right)^{\eta}z^{-\eta}p_Q^{1-\eta}(Q+Q_E)^{\frac{1}{\sigma}} \\
\text{i.e. } p_K &= c_0 z^{\frac{\eta}{\eta-1}}p_Q(Q+Q_E)^{\frac{1}{\sigma(1-\eta)}} \\
&= c_1 z^{\frac{\eta}{\eta-1}+\gamma-1+\frac{1}{\sigma(1-\eta)}}E^{\mu+\gamma-1+\frac{1}{\sigma(1-\eta)}}\left(\frac{\gamma}{\mu}+\frac{Q_E}{zE}\right)^{\frac{1}{\sigma(1-\eta)}}
\end{aligned}$$

Taking limits:

$$\begin{aligned}
\lim_{z\to z_{min}} p_K &= c_2 E(z_{min})^{\mu+\gamma-1+\frac{1}{\sigma(1-\eta)}} \\
\text{i.e. } \lim_{z\to z_{min}} p_K = 0 \quad \text{if} &\quad \mu+\gamma-1+\frac{1}{\sigma(1-\eta)} > 0 \\
\text{i.e. } \quad \text{if} &\quad \sigma < \frac{1}{(1-\eta)(1-\gamma-\mu)} \\
\text{i.e. } \quad &\quad \sigma^* = \frac{1}{(1-\eta)(1-\gamma-\mu)}
\end{aligned}$$

Likewise:

$$\begin{aligned}
\lim_{z\to\infty} p_K &= c_3 z^{\frac{\eta}{\eta-1}+\gamma-1+\frac{1}{\sigma(1-\eta)}+x\left(\mu+\gamma-1+\frac{1}{\sigma(1-\eta)}\right)} \\
\text{i.e. } \lim_{z\to\infty} p_K = 0 \quad \text{if} &\quad \frac{\eta}{\eta-1}+\gamma-1+\frac{1}{\sigma(1-\eta)}+x\left(\mu+\gamma-1+\frac{1}{\sigma(1-\eta)}\right) < 0 \\
\text{i.e. } \quad \text{if} &\quad \sigma > \frac{1}{\frac{\eta}{1+x}+(1-\eta)(1-\gamma-\frac{x}{1+x}\mu)} \\
\text{i.e. } \quad &\quad \sigma^{**} = \frac{1}{\frac{\eta}{1+x}+(1-\eta)(1-\gamma-\frac{x}{1+x}\mu)}
\end{aligned}$$