



Working Papers

www.cesifo.org/wp

A Spatial Approach to Energy Economics

Juan Moreno-Cruz
M. Scott Taylor

CESIFO WORKING PAPER NO. 4173
CATEGORY 9: RESOURCE AND ENVIRONMENT ECONOMICS
MARCH 2013

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

A Spatial Approach to Energy Economics

Abstract

We develop a spatial model of energy exploitation where energy sources are differentiated by their geographic location and energy density. The spatial setting creates a scaling law that magnifies the importance of differences across energy sources. As a result, renewable sources twice as dense, provide eight times the supply; and all new non-renewable resource plays must first boom and then bust. For both renewable and non-renewable energy sources we link the size of exploitation zones and energy supplies to energy density, and provide empirical measures of key model attributes using data on solar, wind, biomass, and fossil fuel energy sources. Non-renewable sources are four or five orders of magnitude more dense than renewables, implying that the most salient feature of the last 200 years of energy history is the dramatic rise in the use of energy dense fuels.

JEL-Code: Q410.

Keywords: energy, renewables, agglomeration.

Juan Moreno-Cruz
School of Economics
Georgia Institute of Technology
Atlanta / Georgia / USA
juan.moreno-cruz@econ.gatech.edu

M. Scott Taylor
Department of Economics
University of Calgary
Calgary / Alberta / Canada
mstaylor@ucalgary.ca

March 14, 2013

Moreno-Cruz, Assistant Professor, School of Economics, Georgia Institute of Technology, Atlanta, Georgia; Taylor, Professor of Economics and Canada Research Chair, Department of Economics, University of Calgary, Canada and Research Associate at the NBER. In addition, Taylor is a fellow of the Beijer Institute of Ecological Economics and a CESifo associate and a research professor at the Ifo Institute, Germany. Parts of this paper are cannibalized from “Back to the Future of Green Powered Economies” NBER (2012, No. 18236). We are very grateful to seminar participants in several countries, and for the hospitality of the Energy Institute at Haas, University of California Berkeley and the Henri Poincare Institute in Paris, France where a portion of this work was conducted. We would like to thank, without implicating, Bob Allen, John Boyce, Jared Carbone, Oded Galor, Gene Grossman, David Laband, Arvind Magesan, and Jim Wilen for comments on earlier drafts or portions of this work. Excellent research assistance was provided by Nolan Derby, Rui Wan, Mingtao Xu, and Jevan Cherniwchan. The usual disclaimer applies.

1 Introduction

Energy is the most important commodity in the world today. And by almost any metric, the energy industry is impossibly large. Yearly energy sales at over 10 trillion dollars dwarf expenditures on any other single commodity; trade and transport of energy is immense with over 3 trillion dollars in international transactions driving product deliveries through 2 million kilometers of pipelines and 500 million deadweight tons of merchant shipping; 8 of the 10 largest global corporations are energy companies; and a third of the global shipping fleet is occupied shipping oil. Given these figures it may not be surprising that world energy consumption takes the energy equivalent of over 2800 barrels of oil *per second* to quench.¹

Against this background of facts and figures testifying to the enormity of energy in our world today, stands an energy economics literature that is very applied in focus, and often partial equilibrium in nature. While general equilibrium and more aggregative approaches to energy economics are not uncommon, most energy economists today eschew a more abstract and aggregative approach to energy for a decidedly micro-based attention to individual markets, firms and energy supplies. This IO approach to energy is quite natural and appropriate given energy is not a homogenous commodity; energy firms often wield substantial market power; and government regulation is ubiquitous - but one wonders, what questions are left unanswered as a result.²

The purpose of this paper is to introduce a new, and decidedly abstract and aggregative, approach to energy economics that may be useful in answering some of the more macro questions surrounding energy use and economic activity. We think our approach may prove useful to researchers examining long run and large scale questions regarding energy transitions, the role of energy in economic growth, and the impact of energy supplies on economic geography.

Our approach is informed by three key observations. The first is that energy is not physically scarce. The second is that exploiting far flung energy resources and moving energy to markets is primarily what the energy industry does. And third, perhaps the most important attribute of an energy source is its density: its ability to deliver substantial power relative to its weight or physical dimensions. To implement an approach based on these observations we build a very simple spatial model of energy exploitation where the location and density of available energy resources determines the amount of energy available for economic activity in a city center or core.

The assumptions we adopt eliminate any role for physical scarcity by assuming energy

¹See the online appendix for sources and methods of construction.

²A short list of important contributions typifying the IO approach are Borenstein, Bushnell and Wolak (2002), Joskow and Kahn (2002), and Wolfram (1999).

resources are limitless, but still costly to exploit. Exploitation requires transportation to the core, and these costs limit energy use. Since energy resources are located in geographic space, the density of energy resources at any location affects deliveries and ultimate energy supplies. Together these assumptions produce our “Only Energy” model where the location and density of energy resources determines economic outcomes.

We think a spatial approach is valuable for many reasons. Apart from descriptive realism, a spatial approach allows us to focus on the limits to energy use created when resources are sourced from distant or difficult to exploit regions. Many sources of energy available today - solar, wind, coal and non-conventional oil and gas - represent vast, almost limitless, potential supplies. The economic costs of exploiting them however limits their use. For example, the costs of transporting energy today determines the impact local energy sources have on economic development, and transport costs explain why we enlarge markets for natural gas by liquefaction and solar power by concentration.

Several results from our model echo these observations. We find the extent to which we exploit difficult or distant energy sources is determined by their energy density. Low density energy sources like straw or wood must be sourced nearby. Dense energy sources like fossil fuels can be exploited at great distances, and our spatial approach produces a natural definition of a resource frontier. Resources within this frontier constitute the intensive margin of use and generate large rents; resources at the frontier are by definition marginal, earn zero rents and define the extensive margin. For example, arctic oil and some forms of tar sands production may define the resource frontier for oil today, while Saudi crude is part of the intensive margin.

This focus on the resource frontier also reveals an important scaling law linking the extensive margin of exploitation to energy supply. Altering the density of available energy resources creates changes at both intensive and extensive margins. Energy supply rises proportionately with energy density at the intensive margin, but having denser resources to exploit means the extensive margin rises as well. In a spatial setting this implies a scaling law: the area of exploitation (or exploration) scales with the square of the extensive margin in a two dimensional setting.

This scaling law plays out differently for renewables and non-renewables. Common to both is the result that once we combine the adjustments on both margins we find that energy supplies (or potential reserves) are homogenous of degree three in our measure of energy density. This implies, in the case of renewables, that access to an energy source twice as energy dense delivers eight times the energy supply. In the case of non-renewables, we find that every new resource play must first boom and then (optimally) bust. Extractions at first rise and then fall, and peak extractions are also increasing in the energy density of

the resource.

To these impacts we must also add dense energy resources provide strong incentives for improvements in transportation. We first show how the (exogenous) introduction of a low cost transportation option act as if it magnifies the energy density of available nearby resources; placing a road or river near our core is equivalent to endowing the region with resources of higher energy density. Then, and perhaps not surprisingly, we show how the existence of energy dense resources creates large incentives for (endogenous) investments in infrastructure. Together these two results reveal a strong complementarity: energy dense resources beget large investments; and large investments can make any distribution of resources *appear* more energy dense.

These results tell us why the energy infrastructure for oil is the largest and most complicated ever constructed for a single commodity; why villagers in Africa only collect crop residue and dung from nearby sources; why charcoal is created worldwide despite an incredibly inefficient process; and why canals were laboriously developed almost exclusively to move coal in the 19th century but not firewood, grain or people in earlier centuries. They are also suggestive of the strong tendency we see towards natural monopoly in the transmission and distribution of many of today's energy dense resources.

Finally, a spatial approach may well be appropriate in a world soon to contain over 9 billion people. Rising populations are likely to make the areal constraint faced by renewable energies tighter while ongoing and even accelerating urbanization raises the geographic density of energy demands. Indeed the most salient feature of the history of world energy consumption in the last two hundred years is the world's increasing reliance on very dense energy sources delivered to rapidly growing cities in a rapidly urbanizing world. To meet these challenges and those of climate change, both renewable and non-renewable energies are required, and we provide a common framework for thinking about how the supplies of both renewable and non-renewables energy sources may meet this challenge. Moreover, we do so by operating at a resolution level somewhere between the highly detailed industry level studies of the IO approach to energy economics and the highly abstract and aggregative analysis common to resource economics and growth theory.³ Indeed we hope to bridge this gap by showing how detailed discussions of power concepts, transport costs, and resource frontiers can be incorporated into a highly abstract and aggregate setting potentially useful to future work in energy, resource and growth economics.⁴

Introducing new theory requires new measurements as well, and hence we provide meth-

³For a recent treatment see Acemoglu et.al. 2012. A very nice review of several of the energy, environment and growth issues is contained in Smulders (2005).

⁴One such extension is contained in Moreno-Cruz and Taylor (2012).

ods to measure energy densities for both renewable and non-renewable resources. Energy density can be measured in a variety of ways. One method is to rank sources according to their energy content measured in *Joules/kg*. Under this metric, the energy contained in a *kg* of crop residue is less than that in a *kg* of wood, which is in turn less than that of coal, and less than that of oil. While this metric for density is very useful for resources like fossil fuels, it may be less useful for many renewables. Another possible method recognizes that some energy resources take up large areas for their maintenance and exploitation (crops, forest resources, game animals, solar power installations and wind farms) in comparison to what might be called punctiform resources where energy resources are highly concentrated geographically (coal, fast moving rivers and uranium). To capture the areal implications of these resources we can measure the flow of energy per unit area these resources can deliver taking into account the area needed for exploitation or maintenance. This second method is in fact a generalization of the first since it captures both the energy content of sources (key to fossil fuels) and their geographic requirements (key to many renewables).

Since power is the flow of energy per unit time, typically measured in *Watts*, power density measures the flow of energy a source can provide in *Watts* per unit area needed for its exploitation and maintenance.⁵ In the case of renewables power density comes directly from the area requirement for a resource (say timber or biofuels or solar power), together with its ability to provide a steady state flow of energy measured in *Watts*. To facilitate measurement we provide a neat decomposition of the power density of renewable resources into three resource characteristics: its intrinsic growth rate, its physical density in the environment or yield; and its energy content.

In the case of non-renewables surface area is still important, but high energy content non-renewable resources are primarily found subsurface, and energy flows now come from the depletion of energy stocks. To make renewables and non-renewables commensurate we aggregate over the sub-surface deposits using the marginal costs of recovery. Using this method we provide a simple decomposition linking the power density of non-renewables to their energy contents and physical density in the environment. With these methods in hand we then provide estimates of the power density for various food crops, fuel wood, renewable power sources, coal and oil. Although exact measures depend on site specific assumptions regarding deposit quality, in general we find non-renewables are several orders of magnitude more dense than renewables.

Our work is of course related to previous contributions in both energy and resource economics, and has benefitted in perhaps less obvious ways from the contributions of economic

⁵Expending one *Joule* of energy per second provides one *Watt* of power. Power density is typically measured in *Watts* per m^2 .

historians.⁶ Although the Only Energy model is constructed from first principles, it bears some resemblance to von Thunen's model of an Isolated State. In contrast to von Thunen however, transport costs and, by implication, the extensive margin are set by appeal to physical laws governing energy use and work. It seems only appropriate in a model offering a new approach to energy economics to go back to the basics of work, force and friction to derive costs, define concepts, and develop measurement. By doing so we are able to generate a clear extensive margin limiting resource use and link this margin to fundamentals.⁷ In addition, the physics of the problem imbue the model with a form of constant returns that proves useful time and time again in our derivations.

Our work is also related to two streams of the existing energy and resource literature. It bears a family resemblance to other spatial models of resource and energy use where resources and demand centers are treated as points in space (Gaudet, Moreaux and Salant (2001)); where consumers (Kolstad (1994)) or resources (Laffont and Moreaux (1986)) are distributed on line segments; where resource pools are differentiated by costs, suggestive of a spatial setting (Pindyck (1978), Livernois and Uhler (1987), Swierzbinski and Mendelsohn (1989), and Chakravorty, Roumasset, and Tse (1997); and situations where resources themselves move across space (Sanchirico and Wilen (1999)). These contributions are however quite different and focus on very different questions and problems.

Our non-renewables model is also novel but for much of our analysis we employ a reformulation first introduced by Solow and Wan (1976) to translate our spatial model (and its assumptions) into a very tractable equivalent where resources are homogenous but extraction is subject to rising marginal costs. In this setting our finding of a boom and bust pattern for extractions is not novel as this possibility has long been recognized (see for example Pindyck (1978), Livernois and Uhler (1987), and Holland (2008)).⁸ Our innovation here is to tightly tie the boom and bust pattern of extractions to the implications of our scaling law; to show it is a necessary consequence within our spatial setting; and to link the peak of the extraction path to the energy density of resources.

The rest of the paper proceeds as follows. In section two we develop a simple model linking power density to energy supply. This section develops definitions and establishes the production possibility frontier and the optimal extraction rules for both renewable and non-renewable energy sources. We examine the renewables model in a stationary state

⁶See for example Wrigley (2010), Smil (2008), Fouquet and Pearson (1998) and Fouquet (2006).

⁷von Thunen's iceberg costs have the unfortunate and clearly counterfactual implication that any energy resource can be transported anywhere and still yield positive energy rents. See the online appendix A.2 for a proof of this assertion.

⁸For related empirical work see the thoughtful and skeptical review of the existing Peak Oil literature by Boyce (Forthcoming) and the early work by Slade (1982) on U-shaped paths for energy prices.

since recharge rates for many renewables are extremely rapid (solar, wind, hydro). The non-renewables model is examined in a planning context to maximize its connections with the existing literature. Section three discusses methods to measure power density. We derive measures of power density directly from our theory and then construct empirical counterparts using readily available data. A short conclusion ends the paper. An online appendix contains several lengthy calculations and a discussion of data sources.

2 The Only Energy Model

We develop a simple model of energy exploitation where energy is the only input and only output of production. This simple structure allows us to introduce our assumptions on geography and costly transport in a transparent manner and define power and energy density measures very easily. Extensions to introduce additional factors and a market based economy are possible.⁹

We assume consumption and production activities are located at an economic core while potential energy sources are distributed in the surrounding space. The economy's core contains all of its production and consumption units but is zero dimensional — occupies no space. The location of the core is exogenous and fixed throughout. The exploitation zones where energy sources can be found are two dimensional planes allowing us to employ definitions of area, distance, and density. Distance is meant to capture any and all costs incurred when incremental amounts of energy are exploited. Timber resources for example may only be available some distance from the core; and for some renewable energies - solar or wind power, or biofuels - measures of area are important to consider. For other energy sources, like fossil fuels, area considerations are less important, but here increasing distance can reflect the difficulties firms have in accessing incremental resources due to well depth, weather conditions, and non-standard geological formations.

2.1 Renewable Resources

We first consider the case of renewables, since it is simpler and admits a steady state analysis that we can then modify to allow for depletion and exhaustion in the case of non-renewables. We start with a definition. The area exploited to find and collect energy is related to the power obtained measured in Watts [W], and the power density measured in Watts per meter squared [W/m²], of the particular energy source exploited. If the flow of power obtained is W , and the available energy source has power density Δ , then the exploitation zone we need

⁹One such extension is developed in Moreno-Cruz and Taylor (2012).

to meet this demand must equal:

$$EX = W/\Delta \tag{1}$$

where EX is measured in meters squared, $[\text{m}^2]$.

2.1.1 The Geography of Exploitation

If the energy resources are distributed uniformly, then minimum cost search implies energy will be collected from a circular area with the economy's core at its center. Therefore W will be collected from a circular area with size EX defined by a radius R . We assume extraction itself is costless, but transport to the core requires use of energy.¹⁰ To determine the margin of exploitation and energy supply the first step is to relate the distance a resource travels with the density of resources at that distance. The density of resources at any given radii r less than R is simply given by the circumference at this distance, $2\pi r$, divided by the sum of all radii up to and including the limiting radius of R . Since the total area exploited is πR^2 , the density function for energy resources must be given by $f(r) = 2r/R^2$. The average distance any resource is transported is then simply the distance r times the density of resources at that distance. We will refer to this as the average carry distance, ACD , and it is given by:

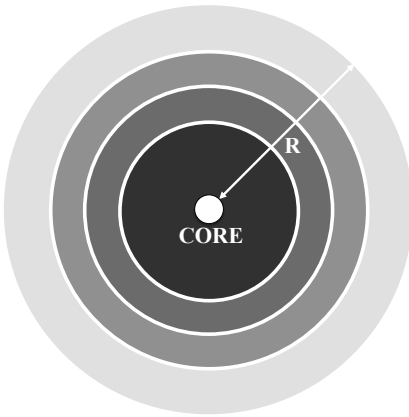
$$ACD = \int_0^R r[2r/R^2]dr = \frac{2}{3}R \tag{2}$$

While we assume energy resources are uniformly distributed across (a two-dimensional) space, it should be clear that other options are easily dealt with at some cost in terms of tractability. In Figure 1 we depict three alternatives. For example, suppose the core was located with resource availability in mind (not an unlikely scenario), then resources may be dense near the core and less dense further away. This is shown in Panel 1.(a) where shading indicates higher power densities. Alternatively, we could allow for the equally likely possibility that a new energy source is located further from the core than was the older established source (think of the potential wind resources on the Great Plains or solar power farms in Arizona). This possibility is shown in Panel 1.(b). And finally, resources could exist in patches or fields within some region, and exploitation may alter the density itself. This possibility is depicted in Panel 1.(c).

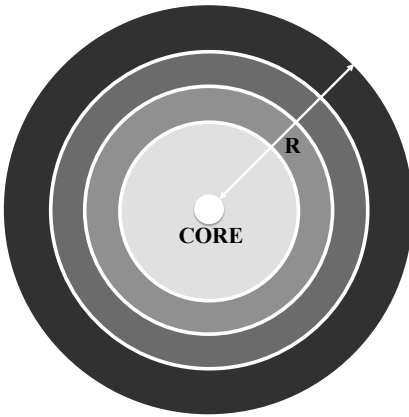
In principle, all of these cases can be examined with an appropriate adjustment of the density $f(r)$ and naturally they will generate average carry distances less than or greater than that given in (2).¹¹ One specific and important possibility we discuss in Section 2.2

¹⁰Adding in constant per unit extraction costs adds little except notation to the analysis.

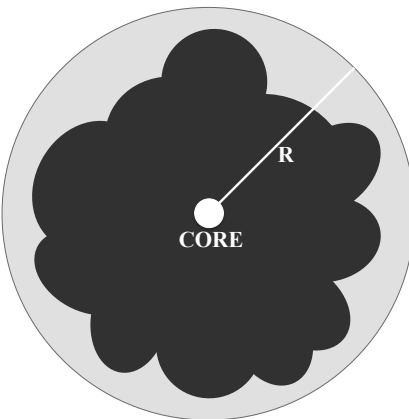
¹¹We analyze situations where resources are either patchy (there are holes) or punctiform (they differ in density) in a subsequent section.



(a) Resources close to the core



(b) Resources away from the core



(c) Uneven resources around the core

Figure 1: Alternative distributions of resources on land

occurs when the exploitation of a non-renewable resource alters the density of resources in a continuous manner hollowing out the exploitation zone over time.

In the uniform case the area of exploitation is a circle with $A = \pi R^2$ and since EX represents the area of exploitation, we can simply link the average carry distance to the energy exploited via $ACD = (2/3)(EX/\pi)^{1/2} = (2/3\sqrt{\pi})(W/\Delta)^{1/2}$ using (1). ACD is measured in meters but to calculate the work done in transporting energy resources we need to find what we will call ‘‘Total Carry,’’ measured in kg-meters. This is found by multiplying the average carry distance by the total weight of resources transported. Total energy resources transported is equal to the area under exploitation times the density of fuel over this area. Therefore, Total Carry is given by:

$$\begin{aligned}\tilde{TC} &= ACD \times EX \times d = [\text{m}][\text{m}^2][\text{kg}/\text{m}^2] = [\text{kg} \times \text{m}] \\ \tilde{TC} &= \left(\frac{2}{3\sqrt{\pi}}\right) \left(\frac{W}{\Delta}\right)^{1/2} EX.d \\ \tilde{TC} &= \left(\frac{2d}{3\sqrt{\pi}}\right) \left(\frac{W}{\Delta}\right)^{3/2}\end{aligned}\tag{3}$$

where we used (1), and d is the (uniform) physical density of this resource in $[\text{kg}/\text{m}^2]$. Total carry is the number of kilogram-meters covered in transporting energy to the core when the power requirements are W and the energy source has density Δ .

Not surprisingly, the greater are the power requirements, the greater are total carrying costs. Less obvious is that marginal (carrying) costs are strictly increasing in power. As power demand rises, larger and larger areas for exploitation must come at the fringe of already sourced areas.

2.1.2 The Physics of Transport

Production possibilities are determined by the amount of power that can be supplied to the core. This section sets out maximum supplied power as a function of energy resource characteristics. A first step in doing so is to recognize that power collected does not equal power available for use in the core. The collection and transportation of resources, like everything else, requires power. The net power supplied to the core is equal to total power collected in the exploitation zone, W , minus the power needed for transportation. In obvious notation, $W^S = W - W^T$.

To calculate the power used in transport, a little high school physics is required. Recall Work is equal to force, f , times distance, x , or work is $W_k = f \cdot x$. Force is in turn equal to mass, M , times acceleration g ; as any mass moved horizontally must overcome the force of

gravity as mediated by friction in transport.¹² All this implies:

$$\begin{aligned} W_k &= \mu[M \cdot g]x = [\text{Newton.meters}] = [\text{Joules}] \\ W_k &= \mu g[M \cdot x] \\ W_k &= \mu g[\text{Total Carry}] \end{aligned} \tag{4}$$

where μ is the coefficient of friction, g is the acceleration of gravity at 9.81m/s^2 , and W_k is then measured in Joules. This work is done per unit time since power is a flow (as is for example the flow of labor and capital services that creates the flow of useful output in standard analyses).

Choosing units is inconsequential and if we measure time in seconds, then the flow of work, W_k , measured in Joules per second is now power requirements measured in Watts.¹³ Putting this all together we find that the total cost of delivering the flow of power W to the core is given by $TC(W)$. Using (3) and (4) we can write this total cost as:

$$W^T = \mu g \tilde{TC}(W) = TC(W) \tag{5}$$

where we note $TC(W)$ is a strictly convex function with $TC(0) = TC'(0) = 0$.

It is now a simple matter to examine how different choices for exploitation generate different supplies of power delivered and available to the core. The calculation of power supplied and delivered to the core, W^S , is made easy by plotting Watts delivered against Watts collected as shown in Figure 2. $TC(W)$ is the strictly convex function shown with the properties just derived. Its vertical height at any W gives the power used in transporting to the core the flow of energy resources represented by the associated power collected on the horizontal axis. By construction, the gap between the 45 degree line and $TC(W)$ represents the power left over to supply the core. Naturally extending the exploitation zone to collect power W^E results in zero power supplied since all of it would be exhausted in transport. In contrast, the choice W^* maximizes supplied power because the vertical distance between W^* and $TC(W^*)$ represents maximum of $W^S = W^* - TC(W^*)$. We will focus throughout on choices that are optimal in this sense.

One important feature of Figure 2 is that it depicts an energy system that is always *productive*; i.e. an energy system that can provide positive supplied power to the economy because there always exists extraction levels for which the power needed for delivery to

¹²We are ignoring static friction encountered when the object first moves. The force that needs to be overcome to keep an object in motion is equal to the normal force times the coefficient of friction. Since the object is moving horizontally, the normal force is just gravity times the mass of the object. The coefficient of friction is a pure number greater than zero; and force is measured in Newtons.

¹³Expending 1 Joule of energy in 1 second means you are delivering 1 Watt of power.

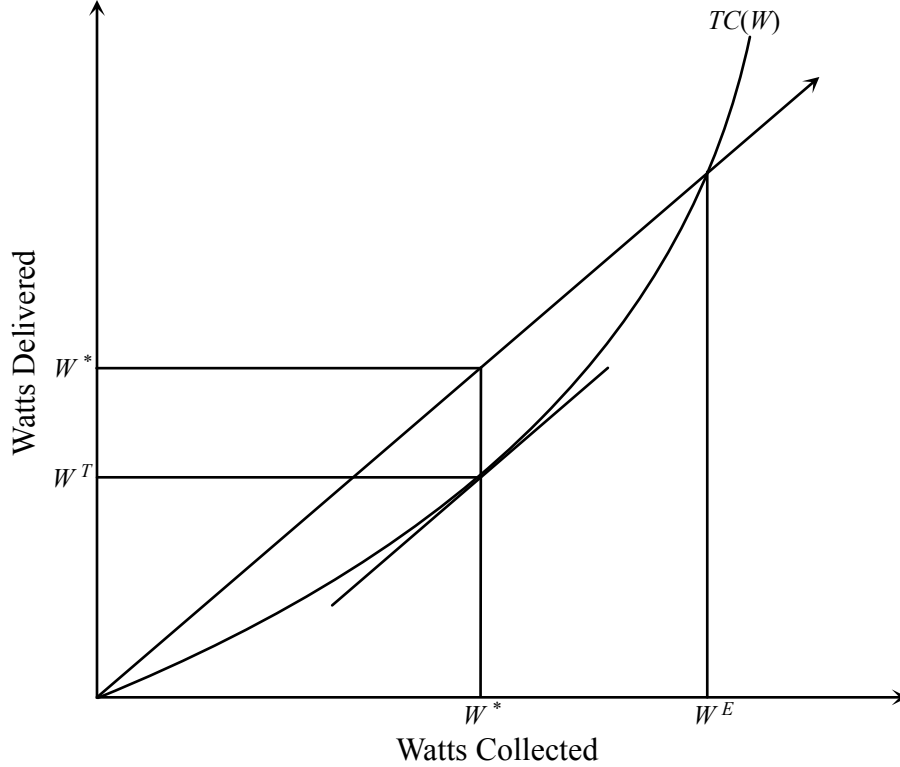


Figure 2: Optimal Energy Exploitation

the core does not exhaust power collected. The system developed here is always productive regardless of the costs of transport. Even if we push the coefficient of friction towards infinity making transport very costly in terms of energy, the slope of the total carry curve at zero Watts, is still zero. In contrast the slope of the 45 degree line remains one even at zero Watts, and hence for any finite coefficient of friction, there will always exist an opportunity to deliver positive power to the core.¹⁴

It is straightforward to show that the maximum supplied power is increasing in power density Δ as this shifts $TC(W)$ outwards, and increasing, for the same reason, when the transportation system becomes more efficient in the physical sense (μ falls). Associated with any W^* is an area of exploitation EX^* . If we focus on efficient (in terms of maximized supplied power) outcomes then we have a simple model of how the density of an energy resource affects power available for use in the core.

When supplied power is maximized, total power is given by:

$$W^* = \frac{\pi \Delta^3}{[\mu g d]^2} \quad (6)$$

¹⁴This does not imply that all landscapes have sufficient resources to support life or a productive economy.

The optimal solution is a cubic function of power density. The intuition is illuminating. Suppose we increase power density but leave the area of exploitation fixed; then supplied power should rise proportionately with power density; i.e. appear with power 1 in an expression like (6) because $W = \Delta EX$. But a higher power density also implies the marginal cost of exploitation falls. With lower marginal costs, exploitation rises and the extensive margin of exploitation moves outwards. Since area is proportional to the square of radius, total power rises with the square of power density because the extensive margin of exploitation changes.

Although it may not be apparent at present, this result is not a function of the circular region of exploitation but follows instead from a well known scaling law. Since area scales with the square of length (here the margin of exploitation), power density will enter as a cube in determining supplied power even when the exploitation zone is not circular. Putting this logic together tells us that the power density of energy sources, and differences in their power densities, is likely to play an important role in determining outcomes.

To find the portion of power used in transport and the remainder that is supplied to the core, evaluate total carry at W^* and subtract to find:

$$W^S = \frac{1}{3}W^*, \quad W^T = \frac{2}{3}W^* \quad (7)$$

Two thirds of the energy collected is spent in transport leaving only one third to supply the power needs of the core. This result *is* linked to the uniform density of resources since an average carry distance of $2/3R^*$ implies only 1/3 of the energy collected will remain for use in the core.

Finally, the area of exploitation is also a function of power density since at the optimum:

$$EX^* = W^*/\Delta = \pi \frac{\Delta^2}{[\mu g d]^2} = \pi [R^*]^2 \text{ where } R^* = \frac{\Delta}{\mu g d} \quad (8)$$

Having a more dense energy source means the energy costs of exploiting far away regions is less: dense energy sources feed further exploitation. The area of exploitation for a fuel twice as dense as another is four times larger. For example, villages in the Middle Ages rarely sourced fuel wood from more than a few miles away.¹⁵ Coal however has been transported hundreds of miles on barges, wagons and trains for much of the last two centuries. Today oil is sourced from some of the most inhospitable climates in the far reaches of the planet and from wells drilled literally miles deep below the ocean surface. Equation (8) tells us this is no

¹⁵Towns on rivers would be an exception, since their transport costs are far less. See our discussion of rivers in section 2.1.3.

surprise - dense energy sources beget large exploitation zones. The reason is straightforward. Since gravity cares nothing about the energy content of any mass transported, higher energy content fuels can in effect fuel their own exploitation by extending the range of (energy) profitable exploitation.

To see why this is true, consider the costs of moving energy resources providing one Watt of power just one meter. Since friction is a constant returns process, the one meter distance calculation is all we really need to understand. One Watt of power, takes up $1/\Delta$ square meters of area which in turn implies the energy resources it represents must weigh d/Δ kilograms. Moving this mass one meter, and overcoming friction, requires a flow of power of $\mu gd/\Delta$. Therefore, $\mu gd/\Delta$ is the number of Watts needed to transport one Watt worth of an energy source with power density Δ , one meter. With this in hand, it is now apparent that (8) identifies marginal energy resources which are located R^* meters from the core and therefore provide zero net energy supply, that is we must have at the margin of exploitation that:

$$1 - R^* \frac{\mu gd}{\Delta} = 0$$

Therefore, the more dense are the energy resources, the lower are the transport costs of obtaining a marginal Watt of power, and not surprisingly the energy exploitation zone must be larger.

This discussion of the margin of exploitation suggests an alternative method of finding net supply that we will sometimes use. Net power supplied must come from adding up, what we could call, “energy rents.” These rents, $1 - \frac{\mu gd}{\Delta} r$, are collected at distances $r \leq R^*$ from the core. To add them up we use a two step procedure. Along any ray from the core, there are Δ Watts of power every meter and transporting these resources from the core yields a density of $\Delta [1 - \frac{\mu gd}{\Delta} r]$ net Watts of power at distance r from the core. The first step is to add up these resources along our ray over all distances less than R^* . The second step is to accumulate these quantities by sweeping across the 2π radians of our circular exploitation zone. By doing so we obtain net power supply to the core as the sum of all energy rents:

$$W^S = \int_0^{2\pi} \int_0^{R^*} v \Delta [1 - \frac{\mu gd}{\Delta} \cdot v] dv d\varphi = 2\pi \int_0^{R^*} v \Delta [1 - \frac{\mu gd}{\Delta} \cdot v] dv \quad (9)$$

$$W^S = \frac{\pi \Delta^3}{3 [\mu gd]^2}$$

Which is of course identical to what we found earlier.

This discussion is also useful in highlighting our relationship to the formulation of von Thunen (1826) and his many successors (for a thorough and excellent treatment see Samuelson (1983)). In von Thunen’s work crops of various types generate economic rent gradients

falling with distance to the core; and the market economy's demands for wood, cereals, or meat determines relative prices, rents, and therefore land use. Despite some strong similarities our formulation and that of von Thunen are not identical. Where we associate any energy source with a finite region of exploitation tied to its power density, the geometric transport costs of von Thunen - cleverly coined iceberg costs by Samuelson - imply an infinite exploitation zone for any and all energy resources. As a result the iceberg assumption of von Thunen leads to the somewhat uncomfortable implication that we can move a barrel of oil (a cord of wood, a bale of hay, a pound of dung, an Ampere of electricity etc.) a billion miles and still reap some energy resources from it. Most importantly it rules out the adjustment in the extensive margin of exploitation that is critical in linking differences in energy density to ultimate supplies; and by implication eliminates the scaling law we find so important.¹⁶ Although von Thunen's rents do fall with distance, the differences across our approaches is subtle and we leave to an online appendix a fuller discussion of these differences.

2.1.3 Incentives for Natural Monopoly? Roads, rivers and canals.

Rivers, roads, and canals were all important components of the energy transportation system in the 19th century just as power lines, oil pipelines, and LNG terminals are important features today. What these components have in common is that they represent low friction and presumably low cost methods of transporting energy to markets. In this section we consider the case of a low friction alternative by examining the impact of a having a river or a road cross the economic core. Analytically, we treat roads as rivers that lower transport costs in both directions.

The analysis proceeds in two steps. First, given the lower transport costs along a river or road energy suppliers may now decide to take a longer route to the core if it offers lower costs. This decision problem affects the shape of our exploitation zone. Second, since a river or road lowers transport costs the overall size of the exploitation zone will also change.

To proceed consider the decision problem of a potential energy supplier located on one meter squared of land with the flow of energy produced equal to Δ Watts. The supplier can take energy directly into the core or deviate to take advantage of a road or river nearby. Rivers and roads help to reduce the amount of work used in transportation, increasing the

¹⁶How then do authors in this literature find adjustments along this margin? While we have not read all related work, a close read of Samuelson's (1983) treatment will show that city size is fixed by assuming there is a minimum finite bound on the labor to land ratio. This finite limit rules out an infinite zone of food production where small dose of labor combine with large quantities of land to produce output at distant locations that can suffer large transport costs to the core. We suspect all analyses adopt, in one way or another, a similar vehicle for limiting exploitation zones. Whether the margin is infinite or exogenous matters little to us since in both cases it is not linked to energy density and therefore cannot play a role in determining energy supplies.

amount of energy delivered to the core. To capture this in our analysis we allow for the coefficient of friction of the river or road to differ from the coefficient of friction of land by a fraction $\rho < 1$. That is, while the coefficient of friction of land is equal to μ , a road's coefficient of friction is $\rho\mu$ in both directions. When traveling with the current a river's coefficient of friction is also $\rho\mu$ but against it μ/ρ . By this assumption, river transport is only useful when you are an energy producer upstream; whereas road transport reduces frictions in two directions and not one.¹⁷

We assume the river or road is a straight line that crosses the core and expands indefinitely.¹⁸ The location of a supplier relative to the core is described by two terms: ι , the distance from the core and θ the angle between the segment formed by the core and the supplier and the river as shown in Figure 3.

We solve the energy producer's problem in two stages. In the first stage transportation costs are minimized by choosing how much distance to cover by land and how much distance to cover by river. In the second stage profits are maximized. The cost minimization problem is given by:

$$\min_{\iota_1, \iota_2} \left(\frac{\mu g d}{\Delta} \right) \Delta \iota_1 + \rho \left(\frac{\mu g d}{\Delta} \right) \Delta \iota_2 \quad (10)$$

subject to

$$\iota^2 = \iota_1^2 - \iota_2^2 + 2\iota_2 \cos \theta \quad (11)$$

where ι_1 is the distance travelled by land and ι_2 is the distance travelled by river or road.

The constraint follows directly from the law of triangles with ι_1 being opposite to the angle θ , as shown in Figure 3. We can replace the constraint in the objective function to find the optimal distances travelled by land and by river:

$$\iota_1^* = \frac{\iota \sin \theta}{(1 - \rho^2)^{1/2}} \quad (12)$$

and

$$\iota_2^* = \iota \cos \theta - \frac{\rho \iota \sin \theta}{(1 - \rho^2)^{1/2}}. \quad (13)$$

If the distance ι_2^* is strictly positive, the supplier deviates to the river, otherwise the supplier goes straight to the core. We can solve for the critical value of θ that separates the suppliers that go straight to the core from those who deviate to the river:

$$\iota_2^* > 0 \text{ if and only if } \theta \leq \cos^{-1} \rho \equiv \bar{\theta} \quad (14)$$

¹⁷We are ignoring the physics of buoyancy which can make travel by a river even upstream quite efficient. The analytics of this case are left to the interested reader.

¹⁸The case where the river does not cross the core is similar and available from the authors.

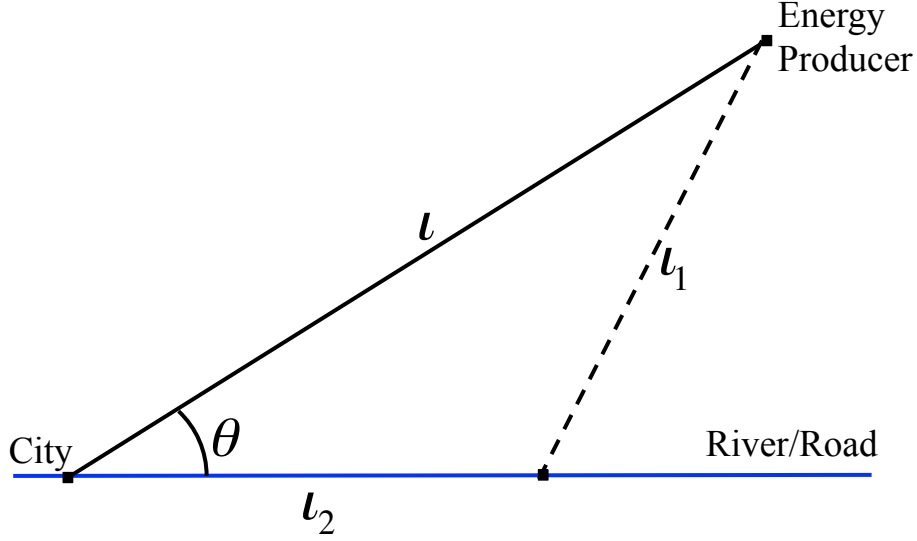


Figure 3: Transport via Road or River

Energy suppliers located at any angle $\theta < \bar{\theta}$ are “close” to the low friction alternative and choose to use it. Since $\rho = \cos(\bar{\theta})$ we know that as $\rho \rightarrow 0$, everyone deviates to the river as they should, since it is so cost effective. Alternatively, as $\rho \rightarrow 1$, the river offers no advantage and no one uses it.

The second part of the energy producer’s problem is to decide whether or not to take its energy to the economic core. An energy producer situated a distance ι from the core and forming an angle θ with the river will go to the core if the energy supplied to the core is positive; i.e., if there are positive energy rents at this location. Energy supplied by this producer is given by

$$W^S = \Delta - \frac{\mu g d}{\Delta} (\iota_1^* + \rho \iota_2^*) \Delta \quad (15)$$

Replacing equations (12) and (13) in the previous equation makes energy rents a function of the distance to the core ι . There is a critical distance ι^* at which suppliers become indifferent between bringing their energy production to the core or supplying zero. This distance is determined by setting $W^S = 0$ and it is given by:

$$\iota^* = \frac{\Delta}{\mu g d} \begin{cases} ((1 - \rho^2)^{1/2} \sin \theta + \rho \cos \theta)^{-1} & \text{if } \theta \leq \bar{\theta} \\ 1 & \text{if } \theta \geq \bar{\theta} \end{cases} \quad (16)$$

which determines the area of exploitation. The zero rent margin, if the producer does not deviate, is exactly as before. But if the producer finds it useful to use the low friction alternative, energy rents are greater and the zero rent margin further from the core.

Total energy supplied to the core can then be written by “adding up” all energy rents over the two regions to obtain:

$$W^S = 2 \times \left[\int_0^{\bar{\theta}} \int_0^{t^*} v \left(\Delta - \frac{\mu g d}{\Delta} (t_1^* + \rho t_2^*) \Delta \right) dv d\theta + \int_{\bar{\theta}}^{\pi} \int_0^{t^*} v \left(\Delta - \frac{\mu g d}{\Delta} v \Delta \right) dv d\theta \right]$$

The integral is multiplied by 2 since our method only adds up over the half circle of π radians whereas total energy comes from the entire exploitation zone. The first integral represents energy coming from suppliers who are close enough to the river to use it in transport. The second integral represents the energy coming from those who travel directly to the core.

Integrating and simplifying gives us a net energy supplied much like that we had before:

$$W^S = \frac{1}{3} \frac{\Delta^3}{(\mu g d)^2} (\pi + g(\rho)) \quad (17)$$

$$g(\rho) = \int_0^{\bar{\theta}} ((1 - \rho^2)^{1/2} \sin \theta + \rho \cos \theta)^{-2} - \bar{\theta} d\theta \geq 0 \quad (18)$$

where the function $g(\rho)$ is positive and monotonic and approaches infinity as ρ goes to zero. We depict the exploitation zone in the river and road case in the two panels of Figure 4 assuming $\rho < 1$.

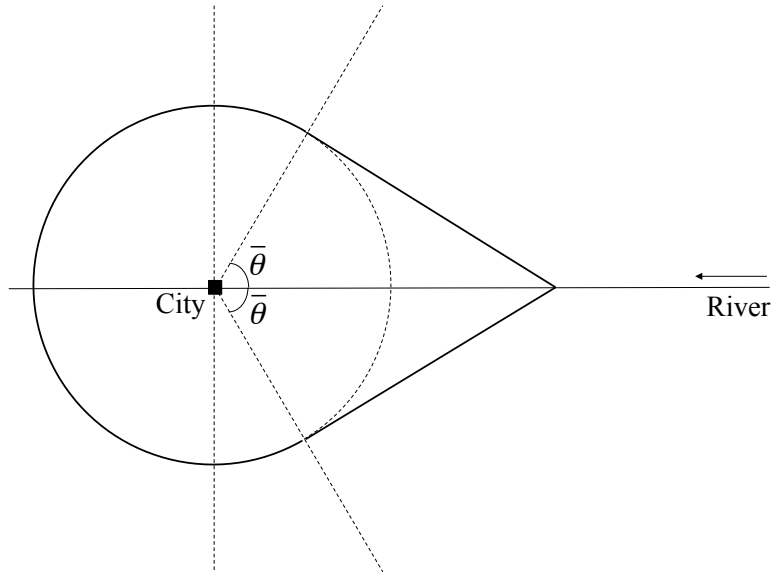
It is now possible to redefine terms slightly to rewrite supply as:

$$W^S = \frac{1}{3} \frac{\pi \tilde{\Delta}^3}{(\mu g d)^2} \quad \text{where } \tilde{\Delta} \equiv \Delta m(\rho)$$

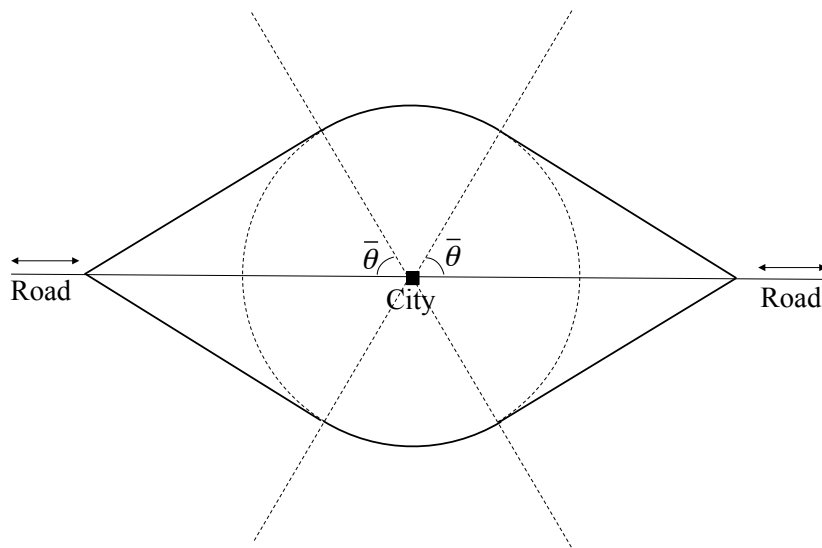
$$m(\rho) = ((1 + g(\rho))/\pi)^{1/3} > 1$$

which is exactly the same form as our earlier supply. The role of improved river transport is identical to being granted a more dense resource base in terms of energy supplied to the core! River location multiplies by $m(\rho)$ the power density of available resources to an extent determined by its capacity for reducing transport costs as reflected in ρ . Setting $\rho = 1$ means the river or road offers no advantage in terms of transportation. This implies $g(\rho) = 0$ since then $\bar{\theta} = 0$ and equation (17) reduces to equation (7). Any $\rho < 1$ effectively magnifies power density; and letting ρ approach zero implies $g(\rho)$ approaches infinity and energy is no longer even economically scarce.

An important feature of this magnification effect is that it is increasing in the power density of available resources. Improved transport is most powerful for very dense resources, and least powerful for low density resources. Any transportation improvement increases the



(a) Area of Exploitation with a River



(b) Area of Exploitation with a Road

Figure 4: Exploitation with Rivers or Roads

extensive margin proportionately, regardless of the energy source; but the resulting change in energy supply is larger for more dense resources because of the scaling law linking changes in the extensive margin to energy supply.

These calculations can be easily expanded to accommodate the case of a road for which the direction of flow is not important. In this case energy suppliers can take advantage of the reduced coefficient of friction in either direction away from the core. Using similar steps, this leads to supply being given by:

$$W^S = \frac{1}{3} \frac{\pi \hat{\Delta}^3}{(\mu g d)^2} \quad \text{where } \hat{\Delta} \equiv \Delta n(\rho)$$

$$n(\rho) \equiv ((1 + 2g(\rho))/\pi)^{1/3} > 1$$

which is of course very similar to the flow of the river case.

The implication of these results is that transport infrastructure is most valuable to convey dense energy resources since the benefits of an existing system are increasing in density. To make this connection clear, suppose we could build roads, canals, or other low cost transportation systems to facilitate energy transport. Suppose further the flow cost of building and maintaining these facilities depend on their efficacy in reducing costs; and we can choose the optimal quality of these investments. Let the cost of building and maintaining transport that delivers an efficiency of ρ be given by $l(\rho)$ which is a decreasing and convex function: $l'(\rho) < 0$, $l''(\rho) > 0$. Then energy supplied net of infrastructure costs are $W^N = W^S(\rho) - l(\rho)$. For concreteness consider the case where $W^S(\rho)$ is given by equation (17) and we are improving river transportation by dredging, locks, canals, maintaining ports etc. The optimal investment problem is simply given by

$$\max_{\rho} W^N = W^S(\rho) - l(\rho) = \frac{1}{3} \frac{\Delta^3}{(\mu g d)^2} (\pi + g(\rho)) - l(\rho) \quad (19)$$

The first order condition that maximizes energy requires

$$\frac{1}{3} \frac{\Delta^3}{(\mu g d)^2} g'(\rho) = l'(\rho) \quad (20)$$

The lefthand side of this equation is the marginal benefit from improved transport and it is again a cubic in power density showing a strong relationship between density and the marginal benefit of further investments. The right hand side is simply marginal costs of improved transportation. Denote the implicit solution to (20) as $\tilde{\rho}(\Delta)$ and total differentiate

it to find¹⁹

$$\frac{d\rho}{d\Delta} = \frac{\frac{\Delta^2}{(\mu g d)^2} g'(\rho)}{l''(\rho) - \frac{1}{3} \frac{\Delta^3}{(\mu g d)^2} g''(\rho)} < 0$$

where the second order conditions ensure the denominator is positive. Therefore, not only do dense resources benefit differentially from existing transportation options like rivers, roads, and sea lanes, but they naturally create large incentives for further improvements magnifying the already large impact differences in transportation make to energy availability in a power dense rich environment.

2.2 Non-Renewable Resources

Using non-renewable energy today precludes you from using it tomorrow. As a result, the exploitation zone will change over time as the resource stock is depleted. This is true because with non-renewables, energy flows come from depleting the resource stock and not from harvesting the perpetual yields flowing from a renewing resource. But stock depletion is not the only way in which renewables and non-renewables differ. Non-renewables are often found beneath the earth's surface in pools, seams, and reservoirs; and naturally their exploitation takes many forms including surface collection (salts, some coals), strip mining (coal, oil sands, peat), and deep mining and drilling (oil, natural gas, uranium). Because non-renewables are often found in topographically concentrated deposits and often at considerable depth, we will need to account for the punctiform nature of these quite different resources in both theory and measurement.

We leave these questions for the moment to examine how the exploitation zone for a non-renewable energy source may evolve over time as energy is supplied from it. One simple and natural way to capture depletion is to have ongoing extraction hollow out the exploitation zone as the resource is extracted. This would for example satisfy a least cost sequence depletion path.²⁰ We present one snap shot of the exploitation zone this path generates in Figure 5. At this point in time, all the energy resources up to r have been extracted, and in order to generate the flow of power W , the new area exploited, EX , would need to be a

¹⁹We have already assumed $l''(\rho) > 0$; it can be shown that $g(\rho)$ is convex for low values of ρ and concave for high values of ρ . Because of the change in concavity there may exist two solutions to the first order conditions, but only one maximum. The second order conditions required for a maximum imply $\frac{1}{3} \frac{\Delta^3}{(\mu g d)^2} g''(\rho) < l''(\rho)$ which simply says the costs are more convex than the benefits from investing in infrastructure.

²⁰There is a small literature examining least cost paths for depletion in situations with multiple deposits or resources. This literature, started by Herfindahl (1967), examines when, and under what conditions, a least cost order of extraction path will be optimal. Chakravorty and Krucic (1994) contains relevant references, some discussion, and a neat result showing the typical least cost path prediction does not hold up when the resources in question are not perfect substitutes in use. This possibility is ruled out in our one energy source set up, but would be relevant in any extension with two, less than perfectly substitutable, resource types.

ring with outer radius R satisfying:

$$EX = \pi (R^2 - r^2) = W/\Delta \quad (21)$$

By rearranging we can neatly link the new exploitation zone with the resource flow extracted $W = \pi\Delta (R^2 - r^2)$.

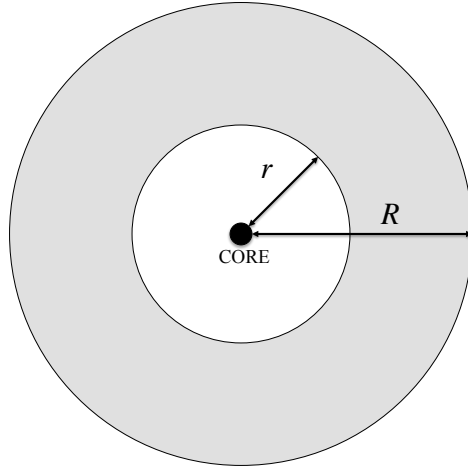


Figure 5: Extraction with Depletion

These W resources have to be brought to the core for use, and we can again calculate the average carry distance by relating the distance a resource travels with the density of resources at that distance. Recalling previous methods, the density of energy resources located a distance v between r and R is given by $f(v) = 2v/[(R^2 - r^2)]$. The average distance resources are transported is again v times the density of resources at that distance. Putting this together we find the average carry distance, ACD , is now given by:

$$ACD(R, r) = \int_r^R v \frac{2v}{(R^2 - r^2)} dv = \frac{2(R^3 - r^3)}{3(R^2 - r^2)} \quad (22)$$

or writing in terms of power W we find:

$$ACD(W, r) = \frac{2\pi\Delta}{3W} \left[\left[\frac{W}{\pi\Delta} + r^2 \right]^{3/2} - r^3 \right] \quad (23)$$

Using now familiar methods, Total Carry is given by:

$$\begin{aligned}\tilde{TC} &= ACD \times EX \times d = [\text{m}][\text{m}^2][\text{kg}/\text{m}^2] = [\text{kg} \times \text{m}] \\ \tilde{TC} &= \frac{2\pi d}{3} \left[\left[\frac{W}{\pi\Delta} + r^2 \right]^{3/2} - r^3 \right]\end{aligned}\quad (24)$$

where again use has been made of (23). Using (24) and (4) we can write this total cost in terms of energy used in transporting energy as:

$$W^T = \mu g \tilde{TC}(W, r) = TC(W, r) \quad (25)$$

This total cost function is identical to our earlier renewables formulation when we evaluate it at $r = 0$. More generally it is increasing and convex in W with $TC(0, r) = 0$, but now marginal extraction costs also depend on r . Differentiating shows:

$$TC_W(W, r) = \frac{2\mu g d}{3\Delta} \left[\frac{W}{\pi\Delta} + r^2 \right]^{1/2} > 0 \quad (26)$$

Whenever $r > 0$, the marginal cost of delivering even the very first unit of power is strictly positive and this marginal cost is increasing in r .

This result is very natural and has two straightforward implications. First, if ongoing extractions drive r upwards over time then the marginal cost of extraction rises over time. As a result, if the resource frontier is currently at $r(t)$, then cumulative extractions must be $\Delta\pi r(t)^2$ and marginal extraction costs are an increasing function of cumulative extractions to date. Second, since the marginal benefit of extracting one more Watt is 1, and marginal costs are rising as the resource frontier moves outwards, eventually extractions must cease. To find this point we can set (26) equal to one and solve to find, not surprisingly, that this occurs when r equals $\Delta/\mu g d$. Therefore, when the resource frontier $r(t)$ approaches the zero rent margin at $\bar{R} = \Delta/\mu g d$ the energy system is incapable of delivering any more net energy.

While we have assumed above that the resource frontier would move slowly outwards over time, there is (as yet) no reason for today's energy maximizing planner to save even an iota worth of energy resources for tomorrow. As a result, all resources within the zero rent margin would be extracted as fast as possible with the frontier jumping to the zero rent margin. Once we assume that current and future energy flows are not perfect substitutes the problem becomes far more interesting.

2.2.1 A Solow-Wan Reformulation

To examine optimal depletion paths it is useful to reformulate our problem. From the discussion above it may already be apparent that our spatial model with reserves differentiated by location and hence rent, is isomorphic to a standard zero dimensional problem where there is a fixed and given resource stock that can be exploited subject to marginal extraction costs that rise with cumulative extractions. A similar reformulation was first suggested by Solow and Wan (1976) in an environment where resources were differentiated by their grade.²¹

What we do here is a standard Solow-Wan reformulation, but with one important twist. The difference comes from the fact that our spatial environment carries with it a built in restriction on the cost distribution of resources. In short an implication of our spatial assumptions is that very high quality - high rent - resources are rare while low quality - low rent - resources are abundant. This follows because as the extensive margin moves outwards resource quality falls linearly, while the quantity of resources available rises quadratically.²² As a result, for any zero rent margin \bar{R} , resources within the top 50% of the quality spectrum represent only 25% of available resources. In this sense high quality reserves are scarce and low quality reserves are abundant. This implication of our spatial assumptions has major implications for the extraction cost function and the optimal depletion path.

To reformulate the problem along Solow-Wan lines, the first step is to link current extractions to remaining reserves. To start we recognize that the maximal exploitation zone has radius \bar{R} , and this maximal exploitation zone implies a corresponding limit on recoverable reserves which we denote \bar{X} . It is apparent that these recoverable reserves can be represented by $\bar{X} = \pi\Delta^3/[\mu gd]^2$. Therefore, if the current resource frontier is $\bar{r}(t) < \bar{R}$, then the remaining recoverable reserves at t , which we denote $X(t)$ are given by

$$X(t) = \int_{\bar{r}(t)}^{\bar{R}} f(\iota) d\iota$$

where $f(\iota) = \frac{2\iota}{\bar{R}^2 - \bar{r}^2}$ is defined above and $r(0) = 0$ since no resources have been extracted at the start of time. This allows us to write remaining reserves as

$$X(t) = \bar{X} - \Delta\pi r(t)^2 \tag{27}$$

²¹Solow and Wan (1976) suggested this reformulation in a short footnote; for a more illuminating treatment see section 2 of Swierzbinski and Mendelsohn (1989).

²²While it is not apparent now we show subsequently that the implications of the scaling law still hold when the resource exploitation zone is patchy (has heterogeneous power densities), punctiform (has holes in it), or when resources are present in certain locations only probabalistically.

where we note cumulative extractions at t , are simply $\Delta\pi r(t)^2$.

Now if we differentiate (27) with respect to time we find the needed link between remaining recoverable reserves and today's rate of extractions:

$$\dot{X} = -2\Delta\pi r(t)\dot{r}(t) = -W(t) \quad (28)$$

The intuition is simple. As extraction proceeds, the resource base is drawn down and the frontier expands. The frontier expands at rate $\dot{r}(t)$ and resources with power density Δ are reaped from a ring with density $2\pi r(t)$ per unit time. The last equality in (28) follows because the instantaneous change in the stock must equal $W(t)$, the flow of energy extracted at t measured in Watts. This completes the first step of the reformulation.

The second step in the reformulation is to find the associated cost function for extractions. To find the cost function recall that at any t when $W(t)$ is extracted, we use $W^S(t)$ for consumption and $W^T(t)$ for fuel extraction. At this t there is also a unique $r(t)$ and therefore extraction costs must be $W^T(t) = r(t)[\mu g d / \Delta]W(t)$ since $W^S(t) = W(t) - r(t)[\mu g d / \Delta]W(t)$. Consequently, $r(t)[\mu g d / \Delta]$ is one measure of the per unit extraction cost. To eliminate the reliance on r , note that equation (27) gives us a relation between remaining reserves $X(t)$, total reserves \bar{X} , and cumulative extractions to date $\Delta\pi r(t)^2$. Use this expression to substitute for $r(t)$ in terms of remaining and recoverable reserves. With some simplification, we can now write the relationship between energy services supplied, $W^S(t)$, current extractions, $W(t)$, remaining reserves $X(t)$, and recoverable reserves \bar{X} , as follows:

$$W^S(t) = [1 - C(X(t))]W(t) \quad (29)$$

$$C(X) = \left(1 - \frac{X}{\bar{X}}\right)^{1/2} \quad (30)$$

where we can now interpret $C(X(t))W(t)$ as the cost of extracting $W(t)$ units of energy from a homogenous pool of recoverable reserves \bar{X} , when remaining reserves equal X . $C(X(t))$ is therefore the unit extraction cost function (where we have suppressed its reliance on recoverable reserves, \bar{X}).

As we mentioned above our spatial formulation imposes further restrictions on the unit extraction cost function. In particular, it has several properties directly tied to our spatial environment.

First consider the role played by remaining reserves X . The unit cost of extraction, when there has been no extraction at all, is zero, $C(X = \bar{X}) = 0$. The very first units of the resource extracted are costless since they are proximate. But as extraction proceeds and remaining

reserves fall, the costs of extraction increase. This is reflected in the finding $C''(X) < 0$. With enough extraction the remaining reserves approach zero, and unit extraction costs hit their upper bound. When all reserves depleted $X = 0$ and we have $C(0) = 1$. The number one has special significance because at the zero rent margin it takes one unit of energy to extract one unit of energy. Therefore, the spatial structure implies unit costs start at zero, rise with extraction, but are always less than one: $C(X) < 1$ for $X < \bar{X}$.

Second, consider the role of total recoverable reserves \bar{X} . An increase in total reserves raises unit extraction costs. This, at first blush, is perplexing. But holding remaining reserves constant, an increase in total reserves implies we are currently exploiting inferior and very costly reserves. As a result cumulative production must be greater and costs higher. By increasing \bar{X} we are in effect adding reserves at the extensive - more expensive - margin and costs rise. In contrast, an increase in remaining reserves X (with given total reserves) adds reserves at the intensive margin. Adding reserves at this margin implies there are more yet to exploit high quality reserves remaining. Unit extraction costs fall. Therefore, the spatial structure allows us to distinguish between the intensive and extensive margins.

Third, and perhaps most importantly, the form of the unit extraction cost function reflects the scaling law we have found so important elsewhere. Since the costs of extraction rise linearly in distance (or what is the same reserve quality falls linearly) while reserves obtained by incurring these costs rise with its square, costs must rise with the square root of cumulative extractions. This is nothing but the dual (cost) implications of the scaling law.

2.2.2 Optimal Depletion

To highlight our connections with the existing literature we adopt a standard planning perspective. We assume a social planner maximizes the welfare of a representative consumer who values the energy services available for consumption in the core. By choosing service units appropriately, utility is defined over net energy supplied. Very little is lost if we adopt a CRRA instantaneous utility function given by:

$$U(W^S) = \frac{(W^S)^{1-\sigma} - 1}{1-\sigma} \quad (31)$$

where $\sigma > 0$ and write social welfare as simply:

$$\max_{W(t)} \int_0^\infty e^{-\rho t} U(W^S(t)) dt \quad (32)$$

The planner maximizes (32) subject to the constraints (28) and (29). We write the current value Hamiltonian as

$$\mathcal{H} = U[(1 - C(X(t)))W(t)] - \lambda(t)W(t) \quad (33)$$

where $\lambda(t)$ is the co-state variable associated with the stock of resources. The optimality conditions are given by:

$$\frac{\partial \mathcal{H}}{\partial W(t)} = U'(W^S(t))(1 - C(X(t))) - \lambda(t) = 0 \quad (34)$$

$$\frac{\partial \mathcal{H}}{\partial X(t)} = -U'(W^S(t))C'(X)W = \rho\lambda(t) - \dot{\lambda}(t) \quad (35)$$

with transversality condition

$$\lim_{t \rightarrow \infty} e^{\rho t} \lambda(t) X(t) = 0 \quad (36)$$

Equation (34) is the static optimization condition. It says to equate the marginal value of an extra unit of energy consumed to the shadow value of resources needed to deliver it. To deliver one unit of consumption requires $1/(1 - C(X(t)))$ units of the resource be extracted at a shadow cost $\lambda(t)$ per unit extracted; that is:

$$U'(W^S(t)) = \frac{\lambda(t)}{(1 - C(X(t)))} \quad (37)$$

Equation (35) is the intertemporal asset market condition. Rearranging slightly:

$$\rho\lambda(t) = \dot{\lambda}(t) - U'(W^S(t))C'(X)W \quad (38)$$

Leaving a marginal unit of the resource *in situ* is an investment of $\lambda(t)$ and this requires a return of ρ . The return provided is composed of expected capital gains at rate $\dot{\lambda}(t)$ plus the marginal reduction in the costs of providing current consumption or $-U'(W^S(t))C'(X)W$.

Using (31) and taking the time derivative of (34), substituting in (35), and rearranging we find one differential equation linking the current rate of extractions to cumulative extractions:

$$\frac{\dot{W}(t)}{W(t)} = -\frac{\rho}{\sigma} - C'(X)W^S(t) \quad (39)$$

A second differential equation is provided by (28) while one initial condition and the transversality condition close the system.

Before proceeding to an examination of the extraction path in the phase plane, it is useful to understand the two quite different motivations at work in (39). This equation is especially important to understand since it is the only behavioral equation governing extractions. First, and not surprisingly, there are the standard Hotelling motives arising from the finiteness of the resource stock and the impatience of our planner. To isolate this motive suppose extraction costs are zero, then using (37) and (38) the shadow value of the resource in situ rises at the rate of time preference. Energy extracted equals energy supplied and we would simply have

$$\frac{\dot{W}(t)}{W(t)} = -\frac{\rho}{\sigma} < 0$$

The values of marginal utilities discounted to time zero must be equalized across all periods and this is achieved by energy consumption falling at a rate proportional to time preference and the elasticity of marginal utility. This motivation follows from the finiteness of the reserves; it predicts a declining path for extractions, and it reflects the forces identified in Hotelling's classic work (Hotelling 1931).

The second motivation at work is a Ricardian one, as it follows from the fact that reserves differ in their Ricardian rents. To see this connection return to our spatial structure for a moment. We know that very proximate resources have large rents and are relatively scarce; very distant ones have very little rent and are relatively abundant. Therefore, resources differ in the Ricardian rent created by their location relative to the core. Once we translate this spatial structure - via the Solow-Wan reformulation - into an implication on extraction costs, it implies that differences in Ricardian rent across reserves are now reflected in extraction costs that rise with cumulative extraction. Since any unit extracted today raises the cost of all future extractions, it pays to shift extraction into later periods. This Ricardian motivation - which follows from differences in energy rents across resources - argues for a delay in extractions or what is the same, a rising path of extractions over time. Ignoring the first Hotelling term in (39) we would find an extraction path that reflects only Ricardian considerations given by

$$\frac{\dot{W}(t)}{W(t)} = -W^S(t)C'(X) > 0$$

Since both cumulative extractions and energy consumed vary with time, the strength of the Ricardian motives vary over time. In contrast with time separable preferences and CRRA instantaneous utility, what we have described as the Hotelling motives are pinned down only by parameters. As we show below, the interplay of these two forces will always

produce a boom and bust path for energy production.²³

2.2.3 Boom and Bust

The behavior of the dynamic system is presented in Figure 6. The $\dot{W}(t) = 0$ isocline is depicted by the solid curve in Figure 6. This curve is positive for values of $X(t) \in [0, \bar{X}]$ and has a maximum value when cumulative extraction is one quarter of total reserves; i.e. with remaining reserves $X(t) = 3\bar{X}/4$ and cumulative extraction $\bar{X} - X(t) = \bar{X}/4$. The $\dot{X}(t) = 0$ isocline is coincident with the horizontal axis. At all points above the $\dot{X}(t) = 0$ isocline, movement must be rightwards to extract all reserves, giving arrows of motion in the positive direction parallel to the horizontal axis. At points above the $\dot{W}(t) = 0$ isocline, extractions must be increasing since costs are currently too low; below the isocline just the opposite is true. This information is captured by the arrows of motion shown.

Imagine we start with a new resource and hence cumulative extractions are zero. Since the arrows of motion near the origin imply all movement must be upwards and to the right, we know the system must move immediately to an initial extraction point like that shown by $W(0) = W_0$. From this initial point, the arrows of motion indicate we move upwards and to the right and cut the $\dot{W}(t) = 0$ isocline at zero slope. Once we cross this isocline, the arrows of motion tell us the extraction path must turn downwards and the transversality condition requires the path slowly approach \bar{X} on the horizontal axis. Working backwards it is now apparent the transversality condition chooses the initial W_0 and this choice has to feature less extraction than that given by the peak of the $\dot{W}(t)$ isocline.

The key result is of course the peak in energy extraction.²⁴ Descriptively, the result

²³One other way to isolate these two forces is to divide up extractions into those going to consumption and those allocated to extraction costs. Hotelling forces must rule the former; Ricardian the latter. To do so take the time derivative of (29). to find:

$$\dot{W}^S(t) = (1 - C(X(t)))\dot{W}(t) - C'(X(t))W(t)\dot{X}(t) \quad (40)$$

and using equations (28) and (39), canceling terms and rearranging obtain:

$$\dot{W}^S(t) = -\frac{\rho}{\sigma}W^S(t) \quad (41)$$

with solution given by $W^S(t) = W_0^S e^{-(\rho/\sigma)t}$ where $W_0^S = W_0$. Therefore, energy consumption follows a very simple Hotelling like rule ensuring that discounted marginal utilities are equal at time zero. It is now immediate then that the profile for energy used in extraction must reflect the Ricardian forces highlighted above. These forces argue for a rising profile for extractions over time, and since they rely on endogenous magnitudes the strength of this force should vary over time. The reader can verify that this component of energy extraction does initially rise quite steeply with time, but eventually falls since energy extracted itself also falls.

²⁴It is well known that the typical Hotelling's prediction that the extraction path is always declining only applies when, among other things, the extraction costs are independent of cumulative extraction or what is equivalent remaining reserves; that is, $C'(X) = 0$. When costs are increasing in cumulative extraction then

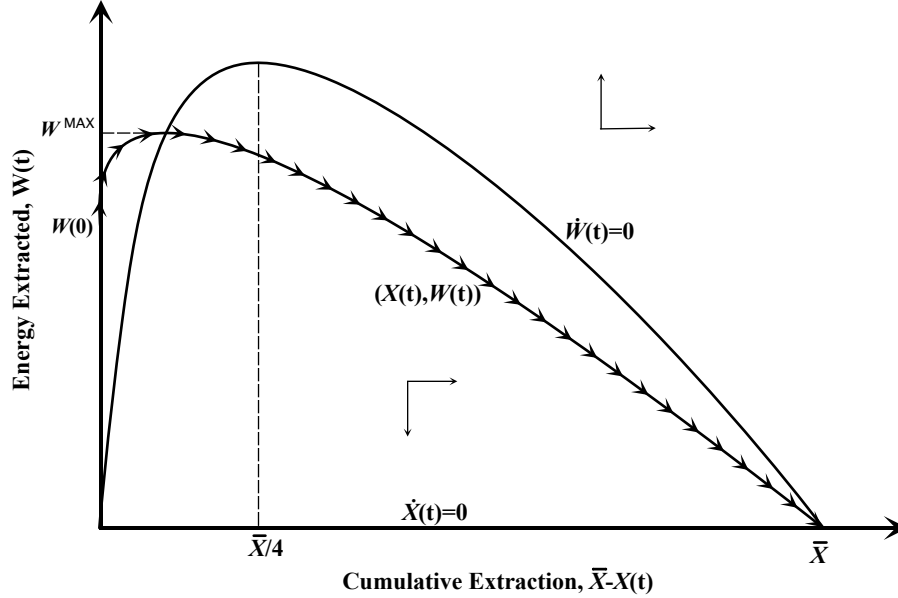


Figure 6: Optimal Extraction Rule

follows because the Ricardian motivations dominate the Hotelling ones up to the point where the system crosses the $\dot{W}(t) = 0$ isocline. Thereafter the Hotelling forces dominate. Analytically, it follows because at the very first instant of time, energy consumption must be positive $W^S(0) > 0$, and $C'(X(0)) = -\infty$ implying $\dot{W}(t) > 0$ at least initially. And as extraction proceeds $W^S(t)$ must approach zero (the resource is finite) and $C'(X(t))$ increases to a finite bound. Therefore, the Ricardian forces fall over time and are eventually dominated by the Hotelling ones.

More deeply, the impact of using up the very first unit of resources on subsequent extraction costs is extremely high, $C'(X(0)) = -\infty$, not because these resources have the greatest rents (which they do) but because they are so scarce in relation to the resource pool whose extraction costs are now raised. Scarcity drives the result; and high rent resources are so scarce because of our scaling law.²⁵

$C'(X) < 0$ and the second term in the righthand side of equation (39) is positive, and if this term is large enough then it is possible for $\dot{W} > 0$, at least for a period of time. This observation is often credited to Livernois and Uhler (1987).

²⁵As a check on our intuition and our scarcity measure we present in the online appendix a simplified model with spatially distributed resources, a Solow-Wan reformulation, and an identical planning problem. We assume however that the extensive margin can only expand in one direction, by letting the length of a fixed width rectangular exploitation zone adjust to economic conditions. Under this assumption, the area of exploitation rises only proportionately with the extensive margin and energy supply does not scale with the square of the extensive margin: our scaling law is absent. As well, we now have a cost distribution over resources given by $F(r) = [r/\bar{R}]$ - one quarter of reserves are now within the lowest one quarter of costs since $F(.25) = .25$. We find the unit extraction cost function is $C(X) = (1 - X/\bar{X})$ which reflects the absence of our scaling law. And $C'(0) = -1/\bar{X} < 0$ is finite and constant. Perhaps not surprisingly, the Hotelling

2.2.4 Two Small Mysteries

Two questions remain unanswered. How is the extraction path related to resource characteristics - most importantly, its power density? And secondly, what is the division of energy across consumption and extraction services over time?

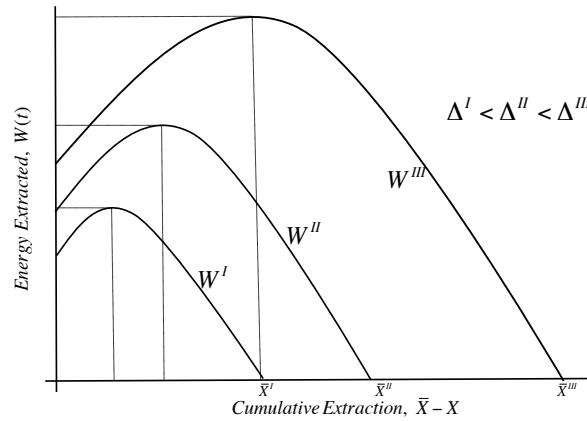
To answer the first, we present in Panel 7.(a) the time paths of reserves and extraction for three different values of power density, $\Delta^I < \Delta^{II} < \Delta^{III}$. An increase in power density increases total initial reserves, so that $\bar{X}^I < \bar{X}^{II} < \bar{X}^{III}$. The initial level of extraction also increases when power density increases, $W_0^I < W_0^{II} < W_0^{III}$. And the boom and bust pattern we have found is accentuated when the resource in question is of higher density. Therefore, while all new resource plays must boom and then bust this pattern may be most visible in the case of very productive, and power dense, fossil fuels. Moreover the peaks in energy extraction occur later in the life of the resource when cumulative extractions are much larger.

The intuition follows from our earlier discussions. An increase in power density raises the quality of all reserves we would have exploited (an intensive margin effect) and adds new ones (an extensive margin effect). Giving the economy more resources would normally call for a level shift in energy consumption in all periods via the Hotelling motive. In fact, if only Hotelling motives were at play it is easy to show an increase in power density would be met with a uniform shift in energy consumption in all periods. But higher power density also moves the resource frontier outwards and this expansion along the extensive margin makes quite costly reserves now economic to recover. The Ricardian motive strengthens and this argues for a shift in the extraction path. Together we find higher extractions but a path shifted forward.

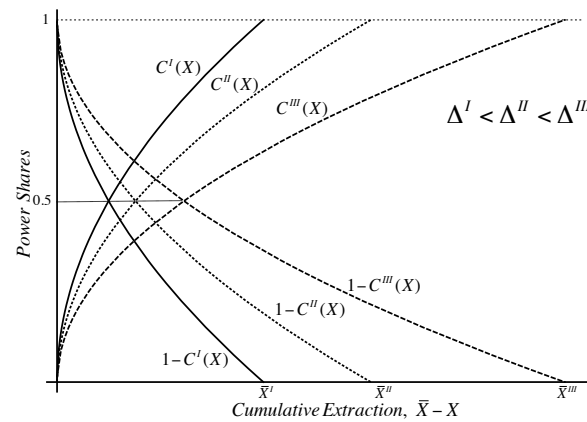
One useful way to examine how the optimal division of energy across consumption and extraction services is to examine how the shares of energy going to each activity varies as the resource is exploited. The share of energy allocated to consumption is simply $W^S(t)/W(t) = 1 - C(X(t))$ which starts at one and approaches zero in the limit. Similarly, the share of energy used in extraction is simply $W^T(t)/W(t) = C(X(t))$ which starts at zero and approaches one. The two shares are equal when $C(X(t)) = .5$ which occurs when remaining reserves are .75 of total reserves and the $\dot{W}(t) = 0$ isocline peaks. We plot several of these shares in Panel 7.(b) showing how these profiles change as the power density of the underlying resource changes. Naturally with more power dense resources extractions are higher than otherwise, but so too is consumption and this implies that the share of consumption in energy use remains larger later in the life of the resource. Power dense resource forestall the day

motivations always outweigh the Ricardian ones and the extraction path is everywhere declining.

when energy costs of extraction limit energy consumption. In all cases the figure shows the share of energy going to extraction costs rises as the resource is developed but this impact is delayed with power dense resources. It is immediate that the energy costs of extraction only rises above 50% after the peak in extraction occurs.



(a) Different Power Densities



(b) Power Shares

Figure 7: Optimal Extraction and Power Density

2.3 Patchy, Punctiform and Probabilistic Resource distributions

Thus far we have assumed resources are uniformly distributed, the space containing resource plays is connected, and there is no uncertainty regarding whether the energy resources in question are present. These are strong assumptions, but for some resources they seem innocuous. For example, crops and woodlands typically satisfy these constraints at least over fairly large areas. But for other resources they fit less well and it is unclear how our analysis would change under these assumptions. For example, the available locations for resource exploitation may be patchy (containing holes) because of land use restrictions, habitat conservation, or noise considerations. The siting decisions for wind and solar farms certainly fit this description. In other cases, most notably fossil fuels, there are often a few very significant deposits surrounded by areas with little if any resource potential (the space contains resource “plays” with widely different power densities). We will refer to this case as one where the resource distribution is punctiform. Finally in some cases it is not clear *ex ante* whether resources are present in any specific location although there maybe a well defined probability distribution over them (oil and gas deposits come to mind). We refer to this case as one where the distribution of resources is probabilistic. We will show that often very little of substance changes with alternative resource distributions although the calculations become more lengthy and the expressions less transparent.

To understand why these complications rarely matter, recall our discussion of energy rents which allowed us to define the extensive margin R^* , for a resource of given power density Δ . Let this reliance of the extensive margin on the power density of resources be written as $R^*(\Delta)$. Then since all resources within this margin provide positive energy rents it should be apparent that they will be exploited even if the resource base is not connected nor homogenous. And if we locate all such potential resources, identify their extensive margins, and then integrate over their relevant regions this (more complicated) sum of energy rents will equal the energy supply just as before. Apart from mathematical complications, patchy and punctiform resource distributions pose no special problem. Alternatively if we assume resources are present in specific locations with given probabilities, we can again identify $R^*(\Delta)$ and integrate over this space to find what would now be expected energy supply. And if the space defined by $R^*(\Delta)$ can be divided into many resource plays with identical and independent success distributions, then a law of large numbers result could be invoked to render expected energy supply equal to *ex post* energy supply. At bottom the reason why these complications do not matter much is the constant returns built into transport costs by the physics of the underlying problem. Moving an object twice as far is twice the work; moving an object with twice the mass is twice the work; and if movement is output and work (energy) is the input, this production function is CRS. The CRS feature of the problem

allows us to aggregate easily, define boundaries simply, and replace patchy, punctiform and probabilistic resource distributions with much simpler connected and homogenous ones in many cases.

To see exactly how we to incorporate complicated resource distributions, we construct two examples.

2.3.1 Patchy and Punctiform

It may be clear from the description above that the key complication is locating the various resources in space. To make the analysis tractable and transparent, we construct discrete resource distributions. Consider a division of the space surrounding the core into concentric circles that are then divided further into wedges created by extending rays from the core. The result, shown in Figure 8, is a sequence of land parcels we will refer to as resource plays.

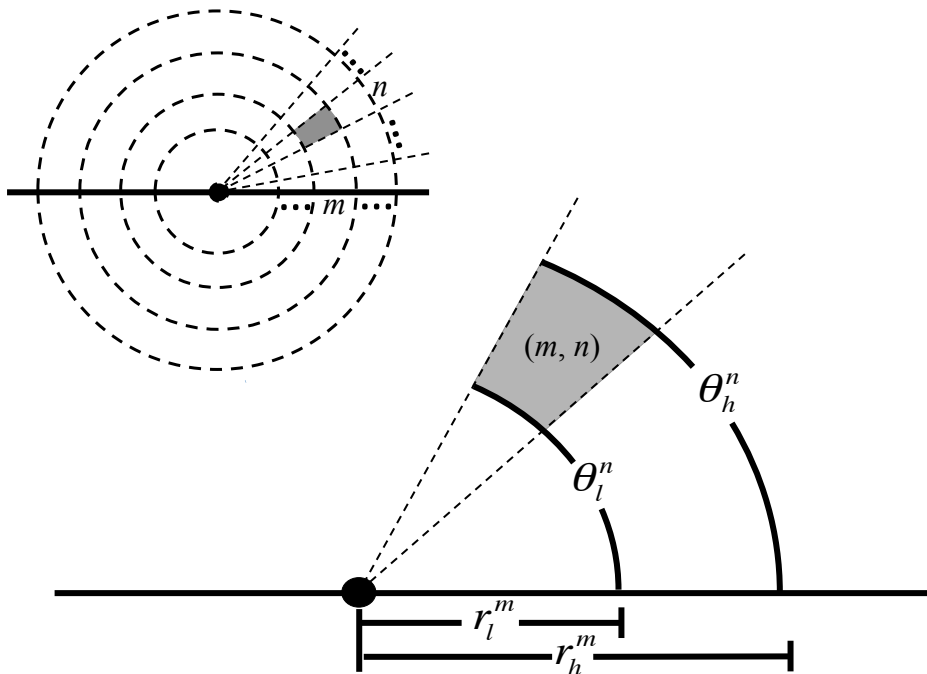


Figure 8: Uneven distribution of power density into parcels

Let there be $n = 1, \dots, N$ rays and $m = 1, \dots, M$ circles, then there are $N \times M$ resource plays each uniquely identified by the duple (m, n) . Suppose each play has an associated power density Δ_{mn} with geometrical shape characterized by its width $r^m = r_h^m - r_l^m$ and the angle of the wedge $\theta^n = \theta_h^n - \theta_l^n$. Where h and l refer to both the higher and lower radius

bounds defining the play; and the higher and lower angles (measured in radians) that define its location in the plane. Using the same approach as in equation (9) we can find write the (maximum) energy supplied by any given resource play:

$$W^{mn} = \int_{\theta_l^n}^{\theta_h^n} \int_{r_l^m}^{r_h^m} v \Delta_{mn} \left(1 - \frac{\mu g d}{\Delta_{mn}} v \right) dv d\varphi$$

$$W^{mn} = \frac{1}{2} (\theta_h^n - \theta_l^n) \left((r_h^m)^2 - (r_l^m)^2 \right) \Delta_{mn} \left(1 - \frac{2}{3} \frac{\mu g d}{\Delta_{mn}} \frac{(r_h^m)^3 - (r_l^m)^3}{(r_h^m)^2 - (r_l^m)^2} \right) \quad (42)$$

Since an energy supplier with play (m, n) supplies energy if the play provides positive energy rents, we need to account for this complication by noting that each density Δ_{mn} has an associated energy margin $\bar{R}_{mn} = \frac{\Delta_{mn}}{\mu g d}$. This implies the actual energy supplied to the core by any resource play must be such that:

$$W_{mn}^S = \begin{cases} W_{mn} & \text{if } r_h^m \leq \bar{R}_{mn} \\ \bar{W}_{mn} & \text{if } r_l^m \leq \bar{R}_{mn} < r_h^m \\ 0 & \text{if } \bar{R}_{mn} < r_l^m < r_h^m \end{cases} \quad (43)$$

where \bar{W}_{mn} has the same form as equation (42) but where r_h^m is replaced by \bar{R}_{mn} .

To find the aggregate energy supplied we add the n plays of each annulus m and then add all the annulus. Without further restrictions, the possibilities are very numerous. Therefore consider the case where each play within an annulus m has the same power density Δ_m . As well, order the power densities from lowest to highest such that $\Delta_0 = 0 < \Delta_1 < \dots < \Delta_m < \dots < \Delta_M$ so distant resources are the most power dense, and in order to eliminate potential gaps in our distribution we assume the width of each annulus is determined by the energy margins of the neighboring annulus. That is assume $r_l^m = \bar{R}_{m-1}$ and $r_h^m = \bar{R}_m$. Alternate assumptions are readily investigated. Using these assumptions we can now replace the definition of \bar{R}_m back in equation (42) to find the energy supplied to the core :

$$W^{mn} = \frac{1}{2} (\theta_h^n - \theta_l^n) \left(\left(\frac{\Delta_m}{\mu g d} \right)^2 - \left(\frac{\Delta_{m-1}}{\mu g d} \right)^2 \right) \Delta_m \left(1 - \frac{2}{3} \frac{\mu g d}{\Delta_m} \frac{\left(\frac{\Delta_m}{\mu g d} \right)^3 - \left(\frac{\Delta_{m-1}}{\mu g d} \right)^3}{\left(\frac{\Delta_m}{\mu g d} \right)^2 - \left(\frac{\Delta_{m-1}}{\mu g d} \right)^2} \right)$$

$$W^{mn} = \frac{1}{2} (\theta_h^n - \theta_l^n) \frac{1}{(\mu g d)^2} \left(\frac{\Delta_m^3}{3} - \Delta_m \Delta_{m-1}^2 + \frac{2}{3} \Delta_{m-1}^3 \right)$$

Add over all wedges in the annulus m to find

$$W^m = \frac{\pi}{(\mu g d)^2} \left(\frac{\Delta_m^3}{3} - \Delta_m \Delta_{m-1}^2 + \frac{2}{3} \Delta_{m-1}^3 \right)$$

and add over all annulus M to find:

$$W^S = \frac{\pi}{3(\mu gd)^2} \sum_{m=1}^M \Delta_m^3 \left(1 - 3 \frac{\Delta_{m-1}^2}{\Delta_m^2} + 2 \frac{\Delta_{m-1}^3}{\Delta_m^3} \right)$$

Two observations are in order. First, since the summation is over primitive determinants of the model, we could just as well replace this complicated sum with $\tilde{\Delta}$, where $\tilde{\Delta}$ is the power density of a hypothetical connected and uniformly distributed resource base yielding the same energy supply. $\tilde{\Delta} > 0$ by virtue of our ordering of power densities, and we can write it simply as:

$$\tilde{\Delta} = \left[\sum_{m=1}^M \Delta_m^3 \left(1 - 3 \frac{\Delta_{m-1}^2}{\Delta_m^2} + 2 \frac{\Delta_{m-1}^3}{\Delta_m^3} \right) \right]^{1/3}$$

Second, if we alter the power density of our hypothetical resource base, $\tilde{\Delta}$, by $\lambda > 0$ this is equivalent to uniform scaling by λ of all power densities in the heterogenous resource zone. A moment's reflection will show that energy supply is homogenous of degree three in all power densities taken together. Therefore, for many purposes we can simply write

$$W^S = \frac{\pi \tilde{\Delta}^3}{3(\mu gd)^2}$$

and ignore the fact that the exploitation zone in question is both patchy and punctiform.

2.3.2 Probabilistic

Here we assume the power density of the resource is uniform across space and it is given by Δ_o . This implies all the resources found inside the margin of extraction given by $R_o = \Delta_o/\mu gd$ are going to be exploited. Divide this space as we did before using N rays and M circles to identify $N \times M$ resource plays but now assume each play has a probability q of having a resource with power density Δ_o in place and a probability $1 - q$ of being empty. Given Δ_o is uniform and constant, our previous assumptions imply the width of each annulus is equal to R_o/M . Thus, the geometrical shape of the parcel (m, n) is characterized by its boundaries set by $r_m^h = (m + 1) \frac{R_o}{M}$ and $r_m^l = m \frac{R_o}{M}$ and the angle of the wedge $\theta^n = \theta_h^n - \theta_l^n$. In the case where parcel (m, n) is not empty, we can calculate the same double integral we calculated for the case of patchy distributions and replace the values for r_m^l and r_m^h to find

$$W^{mn} = \frac{1}{2} (\theta_h^n - \theta_l^n) \left(\frac{R_o}{M} \right)^2 \Delta_o ((m + 1)^2 - m^2) \left(1 - \frac{2}{3} \frac{1}{M} \frac{((m + 1)^3 - m^3)}{((m + 1)^2 - m^2)} \right) \quad (44)$$

Replacing the definition for R_o and noting the wedges are of equal size given by $\theta^n = 2\pi/N$ we find:

$$W^{mn} = \frac{\pi}{NM^2} \frac{\Delta_o^3}{(\mu g d)^2} ((m+1)^2 - m^2) \left(1 - \frac{2}{3} \frac{1}{M} \frac{((m+1)^3 - m^3)}{((m+1)^2 - m^2)} \right) \quad (45)$$

As we mentioned above, the power collected from parcel (m, n) is W^{mn} with probability q and it is zero with probability $1 - q$. Therefore, the expected value of energy provided by parcel (m, n) is:

$$E[W^{mn}] = q \cdot W^{mn} + (1 - q) \cdot 0 \quad (46)$$

We can now aggregate across parcels and use the linearity of the expected value operator to find:

$$E[W^S] = q \cdot \bar{M} \frac{\pi \Delta_o^3}{3(\mu g d)^2} \quad (47)$$

where \bar{M} is a constant given by:

$$\bar{M} = \sum_{m=0}^M \frac{((m+1)^2 - m^2)}{M^2} \left(3 - 2 \frac{1}{M} \frac{((m+1)^3 - m^3)}{((m+1)^2 - m^2)} \right) \quad (48)$$

In this more complicated case very little seems to change. Power density again enters as a cubic as before since now the area of *exploration* rises with the square of the extensive margin and success is proportional to this exploration zone. As well, as mentioned earlier if the number of plays were large a variety of assumptions are available on the joint distribution across the plays that would render a law of large numbers result. The simplest case being the one employed above where each play is treated as an independent and identically distributed Bernoulli random variable.

3 Measuring Power Density

Thus far we have taken power density as a primitive. While this is a natural starting point, it is important to develop a deeper understanding of the determinants of power density in order to provide estimates useful for empirical work, for calibration exercises, etc. We start by developing the theory required for measurement. While there exist in the literature estimates of power densities for many energy sources, how these figures are constructed is unclear and rarely documented adequately. Measuring power density for some renewable resources is fairly straightforward; for example, crops dedicated to biofuels or human consumption can be turned into energy equivalents and then power flows by taking account of crop cycle,

length, and area planted; similarly coppiced forests can provide stable flows of wood products for heating and cooking needs and similarly occupy well defined areas. In these cases, the renewable flows are captured by the physical quantity of fuel reaped from a resource stock.

In other cases the renewable flow does not have mass but provides either kinetic or electromagnetic energy we capture and exploit directly. In these cases the measurement is straightforward and represents the potential of these flows. For example, the power density of solar is easily estimated once we are armed with knowledge of insolation potential at a location. Similarly, wind or wave farms provide useful kinetic energy and we again find the potential power flow from the resource per unit area.

In the case of non-renewables, measurement is generally more challenging. One common method is to calculate the actual physical footprint of a energy facility's size and divide this by the current energy output. So for example, if a coal based generating station produces a constant flow of 1MW, and the mine and generating station takes up 1 km², then its power density is simply $[1 \times 10^6 \text{W}]/[1 \times 10^3 \text{m}]^2 = 1 \text{W/m}^2$. There are several obvious problems with this method. First, the measure of power density is technology dependent. Improvements in generation technology will affect power density, and therefore power density will not be a characteristic of an energy source but rather reflect current technology in place. Second, it is difficult to know which "inputs" we should include in the measurement. For example, by including mining, crushing, and generating facilities in the calculation we make implicit decisions about which facilities to include and which to exclude. Should we also include the area taken up by transmission lines, relay stations, and other parts of the grid? In the case of oil, do we include pipelines, refineries and gas stations? If pipelines are buried and transmission lines are above ground, how do we deal with this?

How these decisions are made will materially affect the calculation. Replication of any measure produced will be almost impossible. As a result any comparison across measurement attempts will be far to reliant on the individual judgment of the researcher.

In order to resolve these issues, we present a method to measure power density for both renewables and non-renewables that is independent of technology, easy to replicate, and allows for a comparison of power densities across energy types. To do so we use restrictions from economic theory to help aggregate resources along the dimensions on which they differ, and we provide measures assuming an ideal environment where resource stocks are homogeneous and where the only costs of exploitation are those arising from energy costs. Our goal is to develop measures that reflect only the physical properties of the resource and not our current, past, or future ability to reap these energy flows.

We start our discussion with the case of renewables since it is simpler to understand and relatively straightforward in practice.

3.1 Power Density for Renewables

3.1.1 Theory

Any renewable flow of energy resources produced from a renewable resource \mathcal{F} [Watts] can be written as the product of the current (and steady state) physical size of the resource stock S [kg] multiplied by the energy content of the resource, e [Joules/kg] and a growth rate r [1/time]. This implies the flow is given by $\mathcal{F} = reS$.²⁶ The physical size of the stock can be similarly written as the product of the physical density of the resource, d [kg/m²] times the area actually used by the resource a [m²]. Making this substitution we obtain the flow of energy as $\mathcal{F} = (red)a$. Power density is just the flow of energy per unit area or $\Delta = \mathcal{F}/a = red$ [W/m²].

Power density is proportional to the product of the maximal rate of regeneration, r , which measures the percentage rate of growth of the resource in an unconstrained environment. Perhaps not surprisingly, a renewable energy source that grows twice as fast has twice the power density. It is also proportional to the energy content of the fuel, e , measured in [Joules/kg], again perhaps not surprising that energy density matters but the specific form is of course not obvious. Finally power density also depends on a fuel's physical density, d , measured in [kg/m²]. All else equal a fuel that produces a greater output in terms of harvest weight gives more energy.

Two special cases deserve attention. The first case applies to resources like wind or solar where there is no associated physical product. In this case we replace the stationary harvest of the resource that we used above by a measure of an average flow per unit time, and then apply energy equivalents to obtain a measure in terms of energy production per unit time. For example, average wind flow per unit area in a given location can be transformed into its kinetic energy equivalent per unit time; average solar insolation in a location is already measured in Watts per unit area terms, and measures of the energy in wave motion can likewise be measured in power density terms. In these three cases while there is no physical resource reaped, power density is simply measured by the potential energy flow these resources deliver per unit area, per unit time.

The second case arises when the harvest from a resource affects the resource stock size, and in turn the power density of the energy source. For example this is most likely to occur when resources are crops, forest land, or hydro power. In these situations it is natural to pin down the power density of the resource by assuming the energy supplier manages the resource to maximize the discounted flow of energy over time. In the online appendix we

²⁶For example, a 100kg forest growing at 10% per year generates 10 kg of firewood per year. Firewood contains 15 MJ per kg; and there are $31,536 \times 10^3$ seconds in a year. The forest provides 4.75 W on average for the year.

Table 1: Power Density Crops

Crop	Yield [kg/m^2]	Energy Content [MJ/kg]	Power Density W/m^2
Sugar cane	7.16	17.01	3.86
Sugar beet	4.87	9.75	1.50
Cassava	1.25	6.69	0.27
Bananas	2.05	3.73	0.24
Rice, paddy	0.43	15.5	0.21
Potatoes	1.78	3.22	0.18
Sweet Potatoes	1.25	3.60	0.14
Wheat	0.30	15.07	0.14
Barley	0.26	14.74	0.12

Note: Data are shown in the online appendix.

show that even in this case, the decomposition of power density is very simple and given by:

$$\Delta = \gamma red \quad \gamma > 0$$

where the new term γ captures the intensity of harvesting and is tied directly to the energy suppliers discount rate. Taking this complication into account we find power density is simply linked to four potentially observable magnitudes: the rate of resource regeneration, the energy content of the fuel provided, its physical density, and the discount rate applied to future rents by energy suppliers.

3.1.2 Empirics

Renewable energy has two final uses: food and fuel. One use is supplying energy to maintain bodily functions, while the other use is supplying energy for heating, light and power applications. In Table 1, we present figures on the power density of various staple crops from around the world. The figures presented for yields are estimates of “typical” yields for these crops in a system with sustainable rotation (fallow periods). As shown food crops, even staples, offer relatively small power density. Even the powerful potato offers only 0.18 W/m^2 in terms of food for fuel, but some tropical crops such as cassava (.26 W/m^2) and bananas (.24 W/m^2) provide much more. Crops that have found use as biofuels such as sugar cane (3.87 W/m^2) and sugar beet (1.50 W/m^2) have power densities one order of magnitude larger than other crops.

Table 2 presents figures on the power density of forests for six regions of the U.S. Since the productivity of forests and their composition varies so too does their power density. For

example, the South Central forests have the highest growth rates (column three) whereas the North Central forests have the greatest percent of Hardwoods (column five). The power density figures are again relatively small and on the order of 0.1 to 0.15 W/m².

Table 2: Power Density of Wood

Forest Region	Total Acres (1,000s)	Ave. Prod. (ft ³ /acre/yr)	Percent Softwood	Percent Hardwood	Average Power Dens. (W/m ²)
Northeast	79,803	57.10	25.21%	74.79%	0.10
North Central	84,215	66.54	18.72%	81.28%	0.12
Southeast	85,665	80.22	41.00%	59.00%	0.14
South Central	118,364	84.69	35.20%	64.80%	0.15
Rocky Mountain	70,969	52.00	90.29%	9.71%	0.08
Pacific Coast	75,197	81.71	89.01%	10.99%	0.12

Note: Data are shown in the online appendix.

Comparing Table 1 to Table 2 we see that in general wood provides lower power density than crops. This is true even though wood density, d , and energy content, e , are much higher than the standard crop. The result follows, because forest's rate of growth, r , is very low relative to that provided by annual crops.

Finally in Table 3 we present estimates for solar and wind energy for six regions in the US. In all cases, we measure the potential provided by the resource rather than measures of our current ability to reap the resources in question. The power density of solar energy captures the yearly average amount of radiation collected by one squared meter of surface with tilt equal to the latitude of the point of measurement. Solar radiation data is reported in the first column on Table 3.

The amount of power that can be extracted from wind is a cubic function of the speed of the wind (column 2), and it is proportional to the area of the cross-section perpendicular to the velocity of the wind. Assuming one squared meter cross-section give us the amount of power extracted at any given velocity. If we further assume one meter squared of land is used by one meter squared of cross-section then we can find the power density of wind. This is what we show in the last column of Table 3.

It is interesting to note the huge differences across these tables measuring the power density of crops or timber versus the raw energy flows in solar. Photosynthesis is, even in the best environments, a very inefficient process taking solar power from the sun and then via nature's capital equipment turning it into biomass. Estimates on this efficiency vary but a common estimate is below 1% efficiency, and this is similar to what a very naive comparison of these tables might suggest.

Table 3: Power Density of Wind and Sun

Region	Solar	Wind	
	Power Density [W/m ²]	Speed [m/s]	Power Density [W/m ²]
Northeast	164	7	122
North Central	226	9	281
Southeast	217	7.5	136
South Central	219	7	115
Rocky Mountain	256	9.5	296
Pacific Coast	229	7.5	144

Note: Data are shown in the online appendix.

3.2 Power Density for Non-renewables

3.2.1 Theory

When exploiting a non-renewable energy source of homogenous quality, the flow of energy obtained is proportional to the change in the resource stock over time. If we measure energy in Joules, and time in seconds, then the flow is in Watts; if we measure the resource stock in kilograms then the coefficient of proportionality, e , should be in units of [Joules/kg] while the change in the stock must be in units of kilograms per second. If the resource was distributed uniformly on a two dimensional plane, then the change in the stock is equal to the physical density of the resource, which we will denote by d [kg/m²], times the area exploited, a [m²]. This implies the power density of a non-renewable resource, if distributed on a two dimensional plane, would be written as:

$$W = eda \quad (49)$$

$$\Delta = W/a = ed \quad (50)$$

This makes perfect sense: the flow of energy comes from the change in a stock; the magnitude of the flow is determined by the stock's energy content; the change in the stock is equal to the product of the area extracted times the physical density of the energy source; and dividing by the size of the area exploited, we obtain the result.

This measure of power density may be useful but it only holds in those cases where non-renewables are distributed very thinly across Earth's surface. For example, shallow deposits of the non-renewable Peat may well fit this description. In the vast majority of cases however non-renewables are not thinly distributed as surface deposits. Their key characteristic is in fact their punctiform nature. Our task is to develop a method to allow the density of resources available at the surface to reflect an appropriate aggregate of available resources

lying below. Perhaps not surprisingly, our method of aggregation uses relative prices which in this case reflect the marginal energy costs of exploiting underlying deposits.

The Ideal Deposit: Consider a hypothetical resource owner with resource rights to one meter squared of surface area in our exploitation zone of homogenous resource quality. If this resource owner extracts a 1m^3 cube of energy resources with energy content $e[\text{J}/\text{kg}]$ and volumetric density $d_v[\text{kg}/\text{m}^3]$, then this cubic meter has mass of $d_v[\text{kg}]$ and weight of $d_v g[\text{N}]$. The total energy contained in this cube would be simply $ed_v[\text{J}]$ and if it was exhausted in one second the power delivered would be ed_v [Watts].

Now consider resources contained beneath this 1m^2 as shown in Figure 9.(a). If resources are located at distance $\eta[\text{m}]$ from the surface, then the work needed to bring them to the surface would be just $gd_v\eta[\text{J}]$ since work must be done to offset gravity.²⁷ The energy contained in any cube mined is just $ed_v[\text{J}]$ as before, therefore the net energy this cube represents at the surface is what we have referred to previously as energy rents. Energy rents in this case are just $ed_v - gd_v\eta$. Resources where the energy cost of extraction equals their entire energy content produce zero energy rents and are located at depth $\eta^* = e/g$. These are the resources at the margin of profitability. Resources above this depth provide positive energy rents; resources below it provide negative energy rents. Hence, any resources located at depth $\eta > \eta^*$ will not be recovered. The depth of marginal resources is determined by the energy cost of extraction — it is proportional to g , and is rising in the energy content of the resource itself, e .

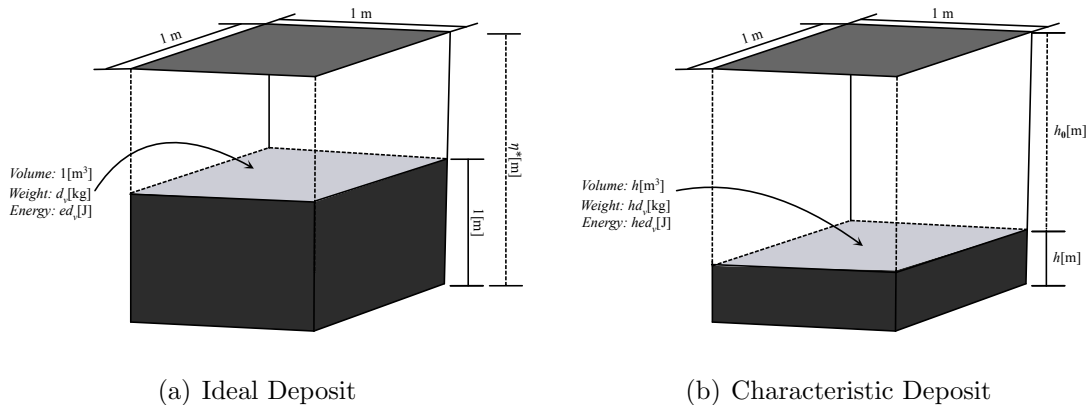


Figure 9: Non-renewable deposit

Putting all of this together shows the energy extracted below 1m^2 of surface area is given by the volume of the mined area $\eta[\text{m}^3]$, the implied mass of mined materials per unit volume,

²⁷We assume here the resource is not under pressure as many liquid or gaseous forms of hydrocarbons are, nor is the extraction subject to friction. Adding either of these complications is relatively easy, but require deposit specific information.

d_v , and the energy content of this extraction, e . This extraction represents the change in the stock over time. The net energy available must account for the energy costs of extraction and is found by the following simple integration.

$$\Delta = \int_0^{\eta^*} [ed_v - gd_v\eta] d\eta = \int_0^{\eta^*} ed_v [1 - (g/e)\eta] d\eta \quad (51)$$

$$\Delta = e^2 d_v / 2g \quad (52)$$

The power density of non-renewable resources in units of $[\text{W}/\text{m}^2]$ is given by $e^2 d_v / 2g$. This is a very important result. It links the power density of non-renewables to two observable and commonly measured characteristics of energy resources – their energy content e and volumetric density d_v – and one parameter g . The power density of exhaustible sources is a quadratic in the energy content of a resource, and linear in its (volumetric) density. It is a quadratic in energy content because doubling energy content doubles the energy contained in any given extraction, while also doubling mine depth. Here again we see the role played by changes in both the intensive and extensive margins. While doubling volumetric density doubles the energy contained in any extraction - since mass rises linearly with density - it has no effect on mine depth. The reason is simply that increased density doubles energy per unit volume, but also doubles the energy costs per unit volume extracted. Since costs and benefits both double, mine depth is independent of volumetric density.

Not surprisingly, this result is quite similar to the preliminary calculation contained in equation (50) and quite similar to our earlier decomposition for renewables. In fact these measures are identical at some level. To see why, it is necessary to recognize the density of deposits (measured in kg/m^2 or d) brought to the surface net of the energy costs of extraction is equal to $ed_v/2g$. This is true since e/g is depth of the ideal deposit, and $d_v/2$ is average carry. Therefore, it is now easy to see how our earlier formulation can account for the sub-surface aggregation proposed above, and we have for nonrenewables:

$$\Delta = ed \quad \text{where } d = ed_v/2g \quad (53)$$

3.2.2 Empirics

Non-renewable resources are found in all continents. Reservoirs can be shallow or deep; they can be large or small; and fuels can be of different qualities. For example, the “quality” of coal as measured by its energy content depends on the depth of burial. Lignite coal is the coal with the lowest energy content and it is formed when Peat is buried between 200 and

1500 meters during the coalification process. The process of Lignite coalification increases the energy content from around 13 MJ/kg for Peat to 16 MJ/kg but it almost doubles its volumetric density from around 355 kg/m³ to 700 kg/m³. The formation of Bituminous coal occurs at greater depths (between 2500 and 6000 meters), and in the process the energy content increases to 32 MJ/kg and the volumetric density again increases to over 900 kg/m³. Anthracite coal is formed deeper than other ranks of coal (between 6000 and 7500 meters). It is the purest form of coal (up to 96% pure carbon) with an energy content of 35 MJ/kg and volumetric density as high as 950 kg/m³. Erosion, earthquakes and volcanic activity can expose these deposits or bury them even more. For example, the maximum depth of deposits in Argentina is 600 meters for sub-bituminous coal and minimum seam thickness of 1.8 meters, Ukraine has deposits of bituminous coal that are 1600 meters deep and have a minimum seam thickness of 0.5 meters. Australia has Lignite coal deposits at a maximum depth of 300 meters and minimum seam thickness of 3 meters (WEC 2010). Many places in the world also feature surface deposits.

Oil and Natural Gas deposits can also be found at various depths and the quality of the resources also depends on the depth of burial, although the process of formation is quite different. The formation of petroleum occurs anywhere between 2500m and 4500m (Deffeyes 2005). The temperature and pressure at this depth combine to transform decayed organic matter into oil. Just as with coal, erosion and other geological forces can bury the oil deposit further or expose it. Oil deposits can be right at the surface of Earth's crust or they can be as deep as 12,000 meters; extreme temperatures below this point are likely to bake most of the crude in the deposit (WEC 2010).

These complications mean that our theoretical construct of a continuous ideal deposit running η meters in depth from the surface is rarely obtained. To amend our calculation for any specific deposit, suppose the upper limit of the deposit is located at a distance h_0 from the surface, and suppose the thickness of the deposit is h , so the maximum depth of the deposit is $h_0 + h$; as shown in Figure 9.(b). Because the resource rights are over one meter squared the dimensions of the deposit are simply $h[m^3]$. In this case to measure the power density of this deposit, we use:

$$\Delta = \int_{h_0}^{\min\{h_0+h, \eta^*\}} [ed_v - gd_v\eta] d\eta = \int_{h_0}^{\min\{h_0+h, \eta^*\}} ed_v [1 - (g/e)\eta] d\eta$$

$$\Delta = \min\{h [d_v e - gd_v(h_0 + h/2)], e^2 d_v / 2g\} \quad (54)$$

Given the large variation in deposit size, depth, and thickness across energy sources we present several possible configurations in Table 4. Several observations are in order. First, non-renewable resources have very high energy contents and correspondingly large power

Table 4: Power Density Non-Renewables

Resource	Energy Content [MJ/kg]	Volumetric Density [kg/m ³]	Power Density [GW/m ²]			
			$h_0=0$ m		$h_0=10000$ m	
			$h=1$ m	$h=10$ m	$h=1$ m	$h=10$ m
Bituminous Coal	33	913	30.1	301	30	300
Lignite	16	865	13.8	138	13.8	138
Peat	14	310	4.34	43.4	4.31	43.1
Qinhuangdao Crude	44	960	42.2	422	42.1	421
Brent Crude	46	830	38.2	382	38.1	381

densities. Even a cubic meter of Peat contains a huge energy supply as shown by the deposit at zero depth and 1m thick which contains 4340 MJ. Part of this feature comes from the fact that a Watt is a very small unit of power; and part from the high energy content of non-renewables.

Second, energy content and volumetric density are positively correlated. Energy resources dense in energy content Joules/kg also tend to be dense in a volumetric sense kg/m³. As a result, as “density” increases these fuels contain increasingly large amounts of energy.

Third, the term $hd_v e$ in equation (54) dominates the energy costs $hgd_v(h_0 + h/2)$ of any extraction. While depth matters, it matters relatively little to overall energy delivered. As a result, power density can be approximated by a linear function of thickness, h . This is apparent by looking across any row of the table at deposits of the same thickness buried at different depths.

Finally, it is very natural to compare the estimates in Tables 1-3, and Table 4 and conclude non-renewables are far superior to renewables. This conclusion is premature. We would start by reminding the reader that the power densities in Tables 1-3 are flows provided *indefinitely* by the resources in question whereas the power densities reported in Table 4 are flows coming from the exhaustion of a quantity of non-renewable resources in a second. In addition it is unclear what superior means. In an economic context superior means of greater social value, and we are abstracting from differences across energy resources in terms of their storability, intermittency and environmental impact.²⁸

²⁸See Borenstein (2012) for an excellent description and analysis of the methods and pitfalls in comparing renewable and non-renewable energy technologies for electricity generation.

4 Conclusions

This paper set out a simple model where energy resources differ in their power and energy density; where the collection, exploitation and transport of any energy resource is costly; and where energy resources are distributed in space. We did so to highlight a hitherto under appreciated characteristic of energy supplies - their densities. Although everyone knows a kilogram of straw offers less energy than one of wood or coal, it is less obvious how these differences can affect the extent to which we exploit these energy resources, the production plans we make, and the surpluses we reap. Although we enjoy the benefits of exploiting dense energy sources everywhere in our daily lives - when we drive to work, turn on a space heater, or use our cell phones - a clear and explicit examination of the implications of energy density is nowhere to be found.

To understand how density may matter to economic outcomes, we built a very simple spatial model where energy is the only input and output in the economy and all outcomes are efficient. A spatial approach is necessary if we are to define and employ measures of power and energy density; a cost to exploiting energy resources is required if we are to understand how energy densities may affect economic outcomes; a one factor energy-in-energy-out framework is the simplest possible setting in which to introduce, develop and discuss key results; and a planning perspective is innocuous in a world where all technologies are CRS and ownership of resources is assumed to be diffuse.

The model has two key characteristics that drive many results: energy resources are distributed spatially and marginal energy resources require more energy inputs to exploit. The spatial setting delivers a scaling law, linking the margin of exploitation of a resource to its supply. Supplies scale with the square of the margin of exploitation and this has large economic consequences. Since dense energy resources are in some sense cheaper to move we are likely to search far and wide to exploit them. The very density of the energy resource we seek, fuels our efforts to obtain more. Therefore differences in power density across energy resources create large differences in energy supply. We find that roads, rivers and other low friction alternatives effectively magnify the power density of available resources turning what could be a resource poor region into a resource rich one. And while our energy system today is extremely complex and features an incredible transport of energy from producers to consumers that suggests locations and energy density may be irrelevant, we turn this logic on its head by showing how the incentive to create such an infrastructure relies on the existence of dense energy sources in the first place.

Our assumption of rising marginal energy costs of exploitation is also key to many results. Even though we assume the potential supply of energy resources is limitless, we show how

the energy supply available for use will always be constrained by the physical necessity of using energy in the costly exploitation and transportation of energy resources. A rising marginal cost to energy exploitation interacted with our scaling law to produce a boom and bust pattern for non-renewable production.

Finally, since there can be no useful theory without measurement we developed ideal measures of power and energy densities for both renewable and non-renewable resources. The measures came directly from our theory, are independent of available technology, and are relatively easy to implement. We then showed how to implement these formulas by providing measures for a range of both renewable and non-renewable energy sources. Our measures show non-renewables are often 4 or 5 orders of magnitude more power dense than renewables implying that the most salient feature of the history of energy use in the last two hundred years is our increasing reliance on extremely energy dense fuel sources.

References

- Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous.** 2012. “The Environment and Directed Technical Change.” *American Economic Review*, 102(1): 131-66.
- Boyce, J.R.** “Prediction and Inference in the Hubbert-Deffeyes Peak Oil Model.” *The Energy Journal*, forthcoming.
- Borenstein, S.** 2012. “The Private and Public Economics of Renewable Electricity Generation.” *Journal of Economic Perspectives*, 26(1): 67-92.
- Borenstein, S. J. Bushnell, and F. Wolak.** 2002. “Measuring Market Inefficiencies in California’s Restructured Wholesale electricity market.” *The American Economic Review*, 92(5): 1376-1405.
- Chakravorty, U. and D.L. Kruice.** 1994. “Heterogenous Demand and Order of Resource Extraction.” *Econometrica*, 62(6): 1445-1452.
- Chakravorty, U., J. Roumasset, and K. Tse.** 1997. “Endogenous substitution among energy resources and global warming.” *Journal of Political Economy*, 195: 1201-1234.
- Fouquet, R. and R. Pearson.** 1998. “A Thousand Years of Energy Use in the United Kingdom.” *The Energy Journal*, 19(4): 1-41.
- Fouquet, R.** 2006. “Seven Centuries of Energy Services: The price and Use of Light in the United Kingdom (1300-2000).” *The Energy Journal*, 1: 139-177.
- Gaudet, G., M. Moreaux, and S.W. Salant.** 2001. “Intertemporal Depletion of Resource Sites by Spatially Distributed Users.” *American Economic Review*, 91(4): 1149-1159.
- Gipe, P.** 2004. *Wind Power: Renewable Energy for Home, Farm, and Business*. USA: Chelsea Green Publishing Company.

- Herfindahl, O.C.** 1967. "Depletion in Economic Theory." In *Extractive Resources and Taxation*, edited by Mason Gaffney, 63-90. Madison, Wisconsin: University of Wisconsin Press.
- Holland, S.P.** 2008. "Modeling Peak Oil." *The Energy Journal*, 29(2): 61-79.
- Hotelling, H.** 1931. "The Economics of Exhaustible Resources." *Journal of Political Economy*, 39(2): 137-175.
- International Atomic Energy Agency.** 2009. *World Distribution of Uranium Deposits (UDEPO) with Uranium Deposit Classification* Vienna. IAEA-TECDOC-1629
- Joskow, P. and E. Kahn.** 2002. "Quantitative Analysis of Pricing Behavior in California's Electricity Market During the Summer 2000." *The Energy Journal*, 23(4): 1-35.
- Krautkraemer, J.A.** 1998. "Nonrenewable Resource Scarcity." *Journal of Economic Literature*, 36(4): 2065-2107.
- Kolstad, C.D.** 1994. "Hotelling Rents in Hotelling Space: Product Differentiation in Exhaustible Resource Markets." *Journal of Environmental Economics and Management*, 26: 163-180.
- Laffont, Jean-Jacques and Moreaux, Michel.** 1986. "Bordeaux Contre Gravier: Une Analyse par les Anticipations Rationnelles," in Gerard Gaudet and Pierre Lasserre, eds., *Ressources naturelles et theorie economique*. Quebec: Presses de l'Universite Laval, 1986, pp. 231- 53.
- Livernois, J.R. and R.S. Uhler.** 1987. "Extraction Costs and the Economics of Nonrenewable Resources." *Journal of Political Economy*, 95(11): 195-203.
- Moreno-Cruz, J. and M.S. Taylor.** 2012. "Back to the Future of Green Powered Economies," *The National Bureau of Economic Research*, Working Paper No: 18236.
- Neher, P.A.** 1990. *Natural Resource Economics: Conservation and Exploitation*. New York: Cambridge University Press.
- Pindyck, R.S.** 1978. "The Optimal Exploration and Production of Nonrenewable Resources." *Journal of Political Economy*, 86(5): 841-861.
- Sanchirico, J.N. and J.E. Wilen.** 1999. "Bioeconomics of Spatial Exploitation in a Patchy Environment." *Journal of Environmental Economics and Management*, 37: 129-150.
- Samuelson, P.A.** 1983. "Thunen at Two Hundred." *Journal of Economic Literature*, 21(4): 1468-1488.
- Slade, M.E.** 1982. "Trends in Natural-Resource Commodity Prices: An Analysis of the Time Domain." *Journal of Environmental Economics and Management*, 9: 122-137.
- Smil, V.** 2008. *Energy in Nature and Society: General Energetics of Complex Systems*. USA: MIT Press.

- Smith, M.D., J.N. Sanchirico, and J.E. Wilen.** 2009. "The Economics of Spatial-dynamic Processes: Applications to Renewable Resources." *Journal of Environmental Economics and Management*. 57: 104-121.
- Smulders, S.** 2005. "Endogenous Technological Change, Natural Resources, and Growth." In *Scarcity and Growth Revisited: Natural Resources and the Environment in the New Millennium*, edited by R.D. Simpson, M.A. Toman, and R.U. Ayres, Chapter 5. Baltimore: The John Hopkins University Press for Resources for the Future.
- Solow, R.M. and F.Y. Wan.** 1976. "Extraction Costs in the Theory of Exhaustible Resources." *The Bell Journal of Economics*, 7(2): 359-370.
- Swierzbinski, J.E. and Mendelsohn, R.** 1989. "Exploration and Exhaustible Resources: The Microfoundations of Aggregate Models." *International Economic Review*, 30(1): 175-186.
- Von Thunen, J.H.** 1826. "Der Isolierte Staat," In *Beziehung auf Landwirtschaft und Nationaleconomie*. Hamburg.
- Wolfram, C.D.** 1999. "Measuring Duopoly Power in the British Electricity Spot Market." *The American Economic Review*, 89(4): 805-826.
- World Trade Organization.** 2011. "World Commodity Profiles for 2011." *WTO*, Geneva.
- World Energy Council.** 2010. *Survey of Energy Resources*. ISBN: 978 0 946121 021.
- Wrigley, E.A.** 2010. *Energy and the English Industrial Revolution*. Cambridge, U.K.: Cambridge University Press.

A Online appendix to accompany “A Spatial Approach to Energy Economics”

Juan Moreno-Cruz and M. Scott Taylor

A.1 Data and Data Manipulation

A.1.1 How large is the Energy Industry

The aggregate 10 trillion dollar number is an estimate calculated as follows: annual energy consumption was 12,274.6 million tons of oil equivalent in 2011 (British Petroleum, 2012, 40). A ton of oil is equivalent to 7.33 barrels, and the average spot price of a barrel of Brent crude oil, a common benchmark for the world price of oil, was \$111.26 (US) in 2011 (Energy Information Administration, 2012). Calculating annual global energy sales in 2011 using this information yields an estimate of \$10.010 trillion US.

Global energy consumption in barrels of oil equivalent per second in 2011 is calculated by once again taking annual energy consumption of 12,274.6 million tons of oil equivalent, multiplying by 7.33 barrels of oil per ton and then dividing that figure by $31,536,000 = 60 \times 60 \times 24 \times 365$ seconds in 2011 to get 2853.02 barrels of oil equivalent consumed per second.

The World Trade Organization valued 2011 global fuels exports (f.o.b.) at \$3.171 trillion US (2012, 3). Fuels are classified by the United Nations Statistics Division (2006) in section 3 of the Standard International Trade Classification, Rev. 4 as: coal, coke and briquettes; petroleum, petroleum products and related materials; gas, natural and manufactured; and electric current.

Data on global pipeline infrastructure listed by country is accessible through the Central Intelligence Agency’s World Factbook (2012). Simple addition yields global oil and gas pipeline infrastructure totaling 1,934,298 km, with natural gas pipelines accounting for 1,217,324 km. Oil, refined products and liquid petroleum gas pipelines account for the remaining 716,974 km.

The global oil tanker fleet has a total capacity of 474.846 million deadweight tons (DWT), corresponding to a 34.0 percent share of the total capacity of the global merchant fleet. Liquefied natural gas (LNG) carriers have a combined total capacity of 43.339 million DWT, a 3.1 percent share of the deadweight tonnage of the global merchant fleet. Combining these numbers yields a total global oil and gas merchant vessel capacity of 518.185 DWT. All Information on capacity and percentage share of global fleet capacity is taken from the United Nations Conference on Trade and Development (2011, 37).

Fortune Magazine’s Global 500 list (2012) ranks the largest companies in the world, by revenue. Included in the list’s top 10 are energy companies Royal Dutch Shell, Exxon Mobil, British Petroleum, Sinopec Group, China National Petroleum, State Grid, Chevron and ConocoPhillips.

A.1.2 Table 1: Crops.

Data on 2011 global crop yields in Table 1 is provided by the Food and Agriculture Organization of the United Nations’ (FAO) FAOSTAT Production database (2013). Global crop yield is calculated automatically by the FAO database by dividing annual crop production by area harvested. Production and harvest figures are reported to the FAO by individual countries via questionnaire or national agriculture publications. Detailed information on the collection methods for this data can be found in the Metadata section of the FAO Statistical Yearbook (2012, 357-358) as well as in the entry for agricultural production on the FAO’s Methods & Standards Webpage.

Data on energy content for the crops given in Table 1 is provided by the United States Department of Agriculture, Agricultural Research Service’s National Nutrient Database, Release 25. Energy content is given in kilocalories (kcal) per 100 grams, which is converted to MJ/kg by multiplying by 4.1868 kJ/kcal and then dividing by 100 to convert kJ/100g into MJ/kg.

A.1.3 Table 2: Forests.

Average power density for a forest in each of the United States regions listed in Table 2 is calculated by the formula:

$$\begin{aligned} & (\text{Average Productivity in Region (hg/m}^2\text{/year)} \times \text{Power Density of Average Hardwood} \\ & (\text{J/m}^2\text{/year)} \times \text{Percentage Hardwood in Average Forest}) + (\text{Average Productivity (kg/m}^2\text{/year)} \\ & \times \text{Power Density of Average Softwood} \times \text{Percentage Softwood in Average Forest}) \end{aligned}$$

The power density of average hardwoods and softwoods is calculated from data provided by Engineering Toolbox. To be precise, power density for individual tree species is calculated by dividing the recoverable heat value of a dry cord of wood (million BTU/cord) by the weight of a dry cord (lb/cord) given by Engineering Toolbox to get the recoverable heat value per pound of wood (million BTU/lb), which is then converted to recoverable heat value per kilogram of wood (MJ/kg), also called potential heat value per kilogram of wood, using conversion factors of 1 lb = 0.4536 kg and 1 million BTU = 1055.06 MJ. Calculations in Table 2 are based on an average dry hardwood density of 35.52 lb/ft³ with a 14.89 MJ/Kg potential heat value, and an average dry softwood density of 27.45lb/ft³ with a 14.87 MJ/Kg

potential heat value.

Average hardwood power density is then calculated by averaging the potential heat value per kilogram of wood for five species (aspen, cottonwood, red oak, red maple and white oak) considered common in American forests by the USDA (2007, 62). An identical calculation is done for four species of softwood (hemlock, ponderosa pine, balsam fir and white pine) also considered common in American forests by the USDA (2007, 62) to get an average softwood power density. It should be noted that the variation in potential heat value per kilogram between different species of tree is always less than 1%, so average hardwood and softwood heat values do not vary much if different species of tree are chosen for the calculation.

Data on percentage of softwoods and hardwoods in an average forest is calculated using data on hardwood and softwood volumes provided by the United States Department of Agriculture (USDA) (2007, 206-208). Data on average productivity classifications for forests is also taken from the USDA (2007, 160-162). Forests are categorized by the USDA by cubic feet of wood per acre per year (cu. ft.) into one of five different classifications: 120+ cu. ft., 85-119 cu. ft., 50-84 cu. ft., 20-49 cu. ft. and 0-19 cu. ft. A simple average of the extreme values in each productivity class is taken to represent average productivity for forests within that class, while forests with a productivity exceeding 120 cu. ft. are capped at a productivity of 120 cu. ft. The productivity of an average forest in any region of the United States can then be calculated after making these assumptions. Forests on reserved land, or with an average productivity between 0-19 cubic feet per acre per year, are omitted from the stocks of total hardwood and softwood by the USDA, and hence are also omitted from the calculation of average forest productivity.

A.1.4 Table 3: Solar and Wind.

Solar energy calculations based on data from the National Renewable Energy Laboratory NREL 2012. The data can be downloaded from http://www.nrel.gov/gis/data_solar.html. We use the annual average direct normal irradiance for the lower 48 states and Hawaii PV 10km Resolution 1998 to 2009. The data are originally in kWh/m²/day. We transform them to W/m² multiplying by 1000 to get Watts and dividing by 24 to eliminate day from the calculation.

Wind energy calculations are also from NREL and can be downloaded from http://www.nrel.gov/gis/data_wind.html and it is using wind speeds at a height of 50 meters. The exact relation between speed and power is given by the following equation $W = \frac{1}{2}\rho v^3$, where ρ is the density of air and v is the speed. Increasing the height of the tower increases speed. For example increasing the height of the wind tower from 10 to 50 meters increases speed by approximately 25% which in turn would increase power density. Here we also

assume wind speeds at 50 meters. To calculate power we assume air density at sea-level and temperature of 15°C which is $\rho = 1.225\text{kg}/\text{m}^3$. Gipe P. (2004) “Wind Power: Renewable Energy for Home, Farm, and Business.” Chelsea Green Publishing Company ISBN: 978-1-931498-14-2. We assume a cross-sectional area of 1m^2 sitting on 1m^2 of land. That is, we assume the radius of the wind turbine is approximately 55% of the area it sit on.

For both solar and wind energy we use GIS to aggregate at the state level calculating the area-weighted average of the cells contained in each state. We then obtain the regional area weighted average using the same regions described above.

The six regions of the United States listed in Table 2 and Table 3 are defined by the USDA (2007, 1) as follows:

- Northeast: Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, West Virginia
- North Central: Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, Wisconsin
- Southeast: Florida, Georgia, North Carolina, South Carolina, Virginia
- South Central: Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Oklahoma, Tennessee, Texas
- Rocky Mountain: Kansas, Nebraska, North Dakota, South Dakota, Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
- Pacific Coast: Alaska, Oregon, Washington, California, Hawaii

A.2 von Thunen and Iceberg Costs of Transport

The transport cost assumptions adopted in von Thunen are subtly, but importantly, different from what we have assumed here. In short, iceberg costs require the energy costs of transportation to fall immediately and completely as energy is expended. At the practical level this rules out containers for fuel storage or combustion, residues left from incomplete combustion, and no mass of the vehicle carrying the load. In terms of the oat eating horse example both Von Thunen and Samuelson used, the horse cannot have any mass, the oats cannot remain resident in the horse, and there is of course no wagon to pull. It is fair to say that while the iceberg assumption is tractable it is also a knife edge assumption as we will show below. If even an epsilon of the mass of energy is wasted in moving containers, in moving engines or left in incompletely combusted particles then the transport process produces a result significantly different than von Thunen’s but qualitatively the same as in

our specification. Specifically it will lead to a formulation where there is a maximum zone of exploitation tied to the power density of energy.

To be precise, assume there is some fixed cost associated with moving energy a distance dx . Then transportation costs over this increment are given by:

$$W^T(x) = \left(C(W_0) + \frac{\mu g d}{\Delta} W(x) \right) dx$$

Total energy remaining at distance $x+dx$ is given by $W(x+dx) = W(x) - (C(W_0) + \frac{\mu g d}{\Delta} W(x)) dx$. Rearranging terms we can rewrite this expression as

$$\frac{W(x+dx) - W(x)}{dx} = \frac{dW(x)}{dx} = - \left(C(W_0) + \frac{\mu g d}{\Delta} W(x) \right)$$

The solution to this differential equation is:

$$W(x) = \left(W_0 + \frac{\Delta}{\mu g d} C(W_0) \right) e^{-\frac{\mu g d}{\Delta} x} - \frac{\Delta}{\mu g d} C(W_0)$$

Define R as the radius for which $W(R) = 0$; that is, the energy supplied to the core by any energy source further away than R is zero. The solution for R is:

$$R = \frac{\Delta}{\mu g d} \ln \left(1 + \frac{\mu g d}{\Delta} \frac{W_0}{C(W_0)} \right)$$

The iceberg assumption occurs when $C(W_0) = 0$ since in this case R goes to infinity. The case we consider in the text occurs arises when $C(W_0)$ is proportional to the mass of energy transported; specifically that $C(W_0) = \frac{\mu g d}{\Delta} \frac{W_0}{e-1}$ since then we obtain $R = \frac{\Delta}{\mu g d}$.

Two observations are in order. First, there exists a finite margin of exploitation for any $C(W_0) > 0$. Thus, iceberg costs represent a knife edge assumption; any value other than $C(W_0) = 0$ generates a qualitatively different result. Any and all energy sources sharing the same - infinite - margin of exploitation when $C(W_0) = 0$; they have finite and different margins of exploitation for any $C(W_0) > 0$. Second, iceberg costs have proven tractable in general equilibrium models because they allow us to model the transportation system without introducing another economic activity complicating predictions in small dimensional models. The formulation in the body of the paper does however respect this constraint. Note the only costs of transport come from moving energy (and not containers, equipment or engines even though the result is consistent with formulations with these fixed costs), the key to our result is our assumption that the mass of energy is transported even as it is used (converted)

in transport. In Von Thunen and Samuelson terms, we are either a) still feeding the horse, but only when it gets to town; or b) keeping the oats resident in the horse throughout.

A.3 Renewables with Stock Effects

To investigate further it is necessary to be specific about the dynamics of resource growth. Suppose the energy source is a renewable fuel with natural growth given by $G(S)$ with $G(0) = G(K) = 0$. As is standard let G be strictly concave and let S denote the stock in physical units. For example, the energy source could be a forest, an area dedicated to biofuels, or even an area dedicated to solar or wind power. It is helpful to first take a specific example with explicit units. To that end, let $G(S) = rS[1 - S/K]$ and the maximum sustainable yield harvest, H^{msy} , as one possible plan for taking from the resource. Then in perpetuity this harvest is given by:

$$H^{msy} = G(S^{msy}) = G(K/2) = rK/4 \quad (\text{A.1})$$

where r is the intrinsic resource growth rate and K is the carrying capacity (i.e. a growth rate times a stock level). Now consider units explicitly. If K is measured in kilograms and $K = 100$ kg, and $r = 10\%$ per unit time, then the sustainable harvest is $(.1)(100)/(4)=2.5\text{kg/unit time}$. If we multiply this quantity by the energy content of the fuel in [Joules/kg] denoted by e , we obtain a measure of Joules per unit time that could be harvested from the resource. Choosing to measure time in seconds, we obtain *Watts*. The final step is to divide this flow of power by the area of exploitation needed to maintain it. Since the carrying capacity is K kg, and if the fuel has a physical density, d , measured in [kg/m²], then the total area needed for this resource flow is K/d . All this implies we can write power density for this renewable energy resource, Δ , as:

$$\begin{aligned} H^{msy}e &= \mathcal{F}[\text{Joules/second}] = [\text{Watts}] \\ \Delta &= \frac{(rK/4)e}{K/d} = \left[\frac{\text{Watts}}{\text{m}^2} \right] \\ \Delta &= \gamma red \quad \gamma > 0. \end{aligned} \quad (\text{A.2})$$

Power density is the simple product of three fundamental, commonly used, and potentially observable characteristics of an energy source, plus one behavioral component captured in the parameter γ . The remaining term in power density is the factor of proportionality γ which captures the intensity of harvesting. To see this note that if harvesting results in a steady state stock equal to a fraction of the carrying capacity given by κK , then $\gamma =$

$\kappa(1-\kappa) < 1$. Harvesting zero implies $\kappa = 1, \gamma = 0$ and $\Delta = 0$; harvesting sufficiently high to cause extinction implies $\kappa = 0, \gamma = 0$ and again $\Delta = 0$. The example given above has $\kappa = 1/2$ and $\gamma = 1/4$. We chose this example for a particular reason: if the resource owner was interested in maximizing total energy collected over an indefinite future, then the owner would adjust their take to match that of the maximum sustainable yield. This is obvious, but more generally, if the owner discounts the value of future versus current energy flows, the optimal stationary harvest maximizing this objective would lead to $\delta = G'(S^*)$ where δ is the discount rate on future periods. This is an application of a well known result in resource economics.¹

A.4 A One-dimensional model of Energy Exploitation

Consider due to geographical limitations, the area of exploitation is restricted to a two dimensional stripe that can only expand along the one dimension shown by the horizontal axis as shown in Figure 1.a. This stripe is h meters wide. The core is centered in the origin. Here the exploitation zone is given by

$$EX = \frac{W}{\Delta} = h.R \quad (\text{A.3})$$

where R is the distance away from the core that is needed to collect W Watts. That is $W = \Delta.h.R$. The density of resources at a distance r is given by $f(r) = h/h.R$ and using familiar methods we find the average carrying distance is given by $ACD = W/2\Delta h$. Assume the resource is renewable, then total energy extracted is given by $W^* = \frac{\Delta^2}{\mu g d} h$, the energy used in transportation is $W^T = W^*/2$ and the energy supplied to the core is $W^S = W^*/2$. The zero energy margin can be found by setting $EX = hR = W^*/\Delta$ so that $\bar{R} = \frac{\Delta}{\mu g d}$. Note with one less dimension supply is no longer a cubic in power density.

Now assume the resource is non-renewable. Then the stock of non-renewable resources is $\bar{X} = \Delta h \bar{R} = \frac{\Delta^2}{\mu g d} h$. Using the same techniques as in section 5 we find $X(t) = \bar{X} - \Delta h r(t)$ so that $r(t) = \frac{\bar{X} - X(t)}{\Delta h}$. The cost function is then given by

$$C(X) = \frac{r(t)}{\bar{R}} = 1 - \frac{X}{\bar{X}}. \quad (\text{A.4})$$

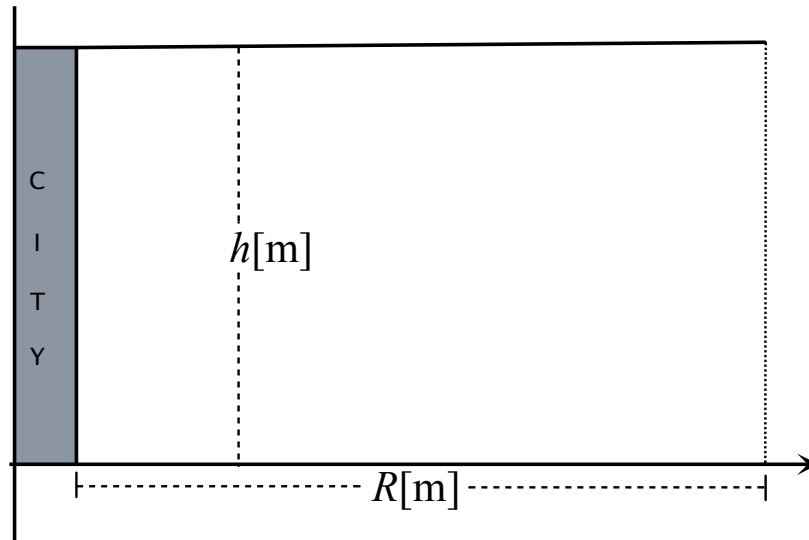
Note from the figure it should be clear that one half the resources are in the first one half of resource quality, unlike our previous case and this is why the cost function appears as it

¹ The optimal stock and attendant harvesting is set only by impatience and is independent of prices. In our simple example with logistic growth, power density is simply $\Delta = \gamma red$, and $\gamma = [1 - \delta/r]^2/4$ is positive as long as suppliers discount rate is not too high.

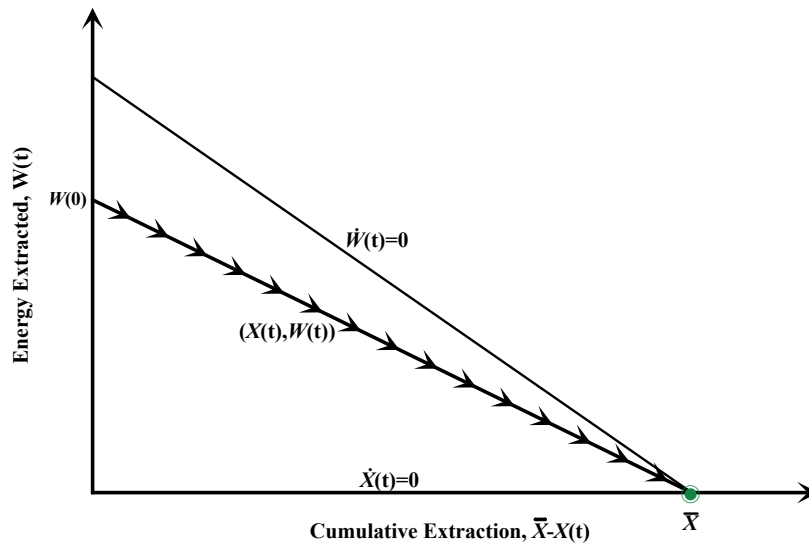
does. Now replacing this equation in equation ?? we find:

$$\dot{W} = -\frac{\rho}{\sigma}W + \frac{W^2}{X}; \quad (\text{A.5})$$

which implies the $\dot{W} = 0$ locus is a straight line of the form $W = \frac{\rho}{\sigma}X$. The phase diagram is shown in Figure 1.b. It shows the transition path towards exhaustion is monotonic.



(a) Exploitation Zone



(b) Phase-Diagram

Figure 1: One-directional energy search model

References for Online Appendix

- British Petroleum.** 2012. *BP Statistical Review of World Energy*. London: British Petroleum.
- Central Intelligence Agency.** 2012. *The World Factbook: Field Listing: Pipelines*. Available at: <https://www.cia.gov/library/publications/the-world-factbook/fields/2117.html>, Accessed March 10, 2013.
- Exxon Mobil.** 2010. *The Outlook for Energy: A View to 2030*.
- Energy Information Administration.** 2012. *2012 Brief: Average 2012 Crude Oil Prices Remain Near 2011 Levels*. Available at: <http://www.eia.gov/todayinenergy/detail.cfm?id=9530>.
- Engineering Toolbox.** *Wood and Combustion Heat Values*. Available at: http://www.engineeringtoolbox.com/wood-combustion-heat-d_372.html, Accessed March 12, 2013.
- Food and Agriculture Organization of the United Nations.** 2013. *FAOSTAT Production*. United Nations. Available at: <http://faostat.fao.org/site/339/default.aspx>, Accessed March 12, 2013.
- Food and Agriculture Organization of the United Nations.** 2012. *FAO Statistical Yearbook 2012 World Food and Agriculture*. Rome: United Nations.
- Food and Agriculture Organization of the United Nations.** 2012. *Methods & Standards*. Available at: <http://faostat3.fao.org/home/index.html#METADATA>, Accessed March 12, 2013.
- Fortune Magazine.** 2012. *Global 500*. Available at: http://money.cnn.com/magazines/fortune/global500/2012/full_list/, Accessed March 10, 2013.
- United Nations Conference on Trade and Development.** 2011. *Review of Maritime Transport*. Geneva: United Nations.
- United Nations Statistics Division.** 2006. *Standard International Trade Classification, Revision 4*. Available at: <http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=28>, Accessed March 10, 2013.
- United States Department of Agriculture.** 2007. *Forest Resources of the United States*. Available at: http://www.fs.fed.us/nrs/pubs/gtr/gtr_wo78.pdf, Accessed March 12, 2013.
- United States Department of Agriculture.** 2012. Agricultural Research Service. *USDA National Nutrient Database for Standard Reference, Release 25*. United States Department of Agriculture. Available at: <http://ndb.nal.usda.gov/ndb/search/list>, Accessed March 12, 2013.
- World Trade Organization.** 2012. *World Commodity Profiles for 2011*. Available at: http://www.wto.org/english/res_e/statis_e/world_commodity_profiles11_e.pdf.