

Thermodynamics of long-run economic innovation and growth

Timothy J. Garrett

Abstract

This article derives prognostic expressions for the evolution of globally aggregated economic wealth, productivity, inflation, technological change, innovation and growth. The approach is to treat civilization as an open, non-equilibrium thermodynamic system that dissipates energy and diffuses matter in order to sustain existing circulations and to further its material growth. Appealing to a prior result that established a fixed relationship between a very general representation of global economic wealth and rates of global primary energy consumption, physically derived expressions for economic quantities follow. The analysis suggests that wealth can be expressed in terms of the length density of civilization's networks and the availability of energy resources. Rates of return on wealth are accelerated by energy reserve discovery, improvements to human and infrastructure longevity, and a more common culture, or a lowering of the amount of energy required to diffuse raw materials into civilization's bulk. According to a logistic equation, rates of return are slowed by past growth, and if rates of return approach zero, such "slowing down" makes civilization fragile with respect to externally imposed network decay. If past technological change has been especially rapid, then civilization is particularly vulnerable to newly unfavorable conditions that might force a switch into a mode of accelerating collapse.

1 Introduction

Like other natural systems, civilization is composed of matter, and its internal circulations are maintained through a dissipation of potential energy. Oil, coal, and other fuels "heat" civilization to raise the potential of its internal components. Frictional, resistive, radiative, and viscous forces return the potential of civilization to its initial state, ready for the next cycle of energy consumption. Burning coal at a power station raises an electrical potential or voltage which then allows for a down-voltage electrical flow; the potential energy is dissipated at some point between the power station and the appliance; because what the appliance does is useful, a human demand is sustained for more coal to burn. Similarly, energy is dissipated as cars burn gasoline to propel vehicles to and from desirable destinations. Or, people consume food to maintain the circulations of their internal cardiovascular, respiratory, and nervous systems while dissipating heat and renewing their hunger.

Such cycles are fairly fast; at least the longest might be the annual periodicities that are tied to agriculture. This paper provides a framework for the slower evolution of

civilization over timescales where such rapid cyclical behavior tends to average out. Instead, the perspective is that material growth and decay of civilization networks is driven by a long-run imbalance between energy consumption and dissipation.

The approach that is followed here builds upon a more general treatment for the evolution of natural systems that has been outlined previously in Garrett (2012c), which starts from first thermodynamic principles in order to develop a fairly general expression for the spontaneous emergence of natural systems. From this point, analytical expressions are provided for economic growth that can be expressed in units of currency. These are then presented in a form that can be evaluated against economic statistics for past behavior and be used to provide physically constrained scenarios for the future.

2 Energetic and material flows to systems

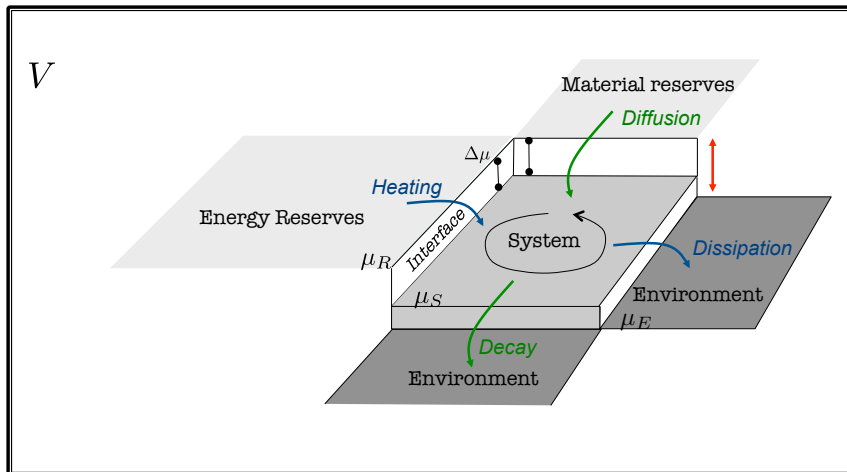


Figure 1: Schematic for the thermodynamics of an open system within a fixed volume V . Energy reserves, the system, and the environment lie along distinct constant potential surfaces μ_R , μ_S , and μ_E . Internal material circulations within the system are sustained by heating and dissipation of energy that is coupled to a material flow of diffusion and decay. The level μ_S is a time-averaged potential. Over shorter time-scales, the legs of a heat-engine cycle would show the system rising up and down between μ_E and μ_R in response to heating and dissipation, as shown by the red arrow, allowing for material diffusion to the system and decay from the system. If flows are in balance then the system is at equilibrium and it does not grow.

The universe is a continuum of matter and potential energy in space. Local gradients drive thermodynamic flows that redistribute matter and energy over time. In the sciences, we invoke the existence of some “system” or “particle” from within this continuum, requiring as a first step that we define some discrete contrast between the system and its surroundings as shown in Fig. 1. This discrete contrast can be approx-

imated as an interfacial jump in potential energy $\Delta\mu$ between the system potential μ_S and some higher level μ_R ; or, $\Delta\mu = \mu_S - \mu_E$ with respect to a lower level μ_E . Matter that lies along the higher potential μ_R has a higher temperature and/or pressure, so it can be viewed as a “reserve” for downhill flows that “pour” into the system. Flows also “drain” to the lower potential environment lying along the potential surface μ_E .

Viewed from a strictly thermodynamic perspective, any system that is defined by a constant potential implicitly lies along a smooth surface within which there is no resolved internal contrast, i.e., one where there is a fixed potential energy per unit matter μ_S and no internal gradients. This specific potential represents the time-integrated quantity of work that has been required to displace each unit of matter within the surface through an arbitrary set of force-fields that point in the opposite direction of the potential vector μ : for example, the gravitational potential per block in a pyramid is determined by the product of the downward gravitational force on each block and its height.

Although internal gradients and circulations are not resolved within a constant potential surface, the presence of the continuum requires that they exist nonetheless. For example, when a bathtub is filled, internal gradients force the water to slosh from side to side. While, the short timescale of these small waves might be of interest to a child, a typical adult cares only about the time-averaged water level of the bathtub as a whole, and that it gradually rises as the water pours in. The definition of what counts as a “system” is only a matter of perspective. It depends on what timescale is of most interest to the observer looking at the system’s variability. As a general rule, however, coarse spatial resolution corresponds with coarse time resolution (e.g., Blois et al., 2013).

The total energy of a system, or its enthalpy H_S , can be expressed as a product of the amount of matter in the system N_S and the specific enthalpy given by

$$e_S^{tot} = \left(\frac{\partial H_S}{\partial N_S} \right)_{\mu_S} \quad (1)$$

The specific enthalpy can be decomposed into the product of the total number of independent degrees of freedom ν in the system and the oscillatory energy per independent degree of freedom e_S ¹

$$e_S^{tot} = \nu e_S \quad (2)$$

The quantity e_S represents the circulatory energy per degree of freedom per unit matter. Thus,

$$H_S(\mu_S) = N_S e_S^{tot} = \nu N_S e_S \quad (3)$$

Conservation of energy considerations dictate that enthalpy is the energetic quantity that rises when there is net heating of the system at a constant pressure (Zemanksky and Dittman, 1997), i.e.

$$\left(\frac{\partial H_S}{\partial t} \right)_p = \left(\frac{\partial Q^{net}}{\partial t} \right)_p \quad (4)$$

¹For example, nitrogen gas at atmospheric temperatures and pressures has a specific enthalpy that is the product of the specific heat at constant pressure c_p and the system temperature T_S , or $e_S^{tot} = c_p T_S$. The specific enthalpy can be decomposed into $\nu = 7$ degrees of freedom. The internal energy has three translational degrees and two rotational degrees. Plus there are an additional two effective degrees that are associated with the pressure energy within a volume. Each degree of freedom has a time-averaged kinetic energy equal to $kT_S/2$ where k is the Boltzmann constant.

and that net heating of the system is a balance between a supply of energy to the system at rate a and a dissipation at rate d

$$\left(\frac{\partial Q^{net}}{\partial t}\right)_p = a - d \quad (5)$$

The Second Law requires that dissipation is to some lower potential and that heating drains some higher potential reserve of enthalpy. Not all enthalpy in the reserve H_R is necessarily *available* to the system. For example, unless the temperature of the system is raised to extremely high levels, the nuclear enthalpy of a reserve $H_R = mc^2$ might normally be inaccessible. Thus, available enthalpy is distinguished here by the symbol ΔH_R .

Heating is coupled to material flows in what can be idealized as a four step cycle termed a “heat engine”, whose circulation is shown by the red arrow in Fig. 1. A system that is initially in equilibrium with the environment at level μ_E is heated, which raises the potential level of the system μ_S an amount $2\Delta\mu$ to level μ_R with a timescale of $\tau_{heat} \sim 2\Delta\mu/a$. It is at this point that the surface μ_S comes into diffusive equilibrium with respect to external sources of raw materials, allowing for a material flow to the system (Kittel and Kroemer, 1980)². There is then cooling through dissipation of heat to the environment with timescale $\tau_{diss} \sim 2\Delta\mu/d$, which brings the system back into diffusive equilibrium with surface μ_E , allowing for material decay.

How the thermodynamics is treated depends on whether the timescale of interest is short or long compared to τ_{heat} .

2.1 Systems in material equilibrium

Over time scales much shorter than τ_{heat} , the legs of the heat engine are resolved, so that the amount of matter in a system N_S would appear to change sufficiently slowly that it could be considered to be fixed. In this case, the response to net heating would be that the specific enthalpy per unit matter rises at rate

$$\left(\frac{\partial e_S^{tot}}{\partial t}\right)_{p,N_S} = \frac{1}{N_S} \left(\frac{\partial Q^{net}}{\partial t}\right)_{p,N_S} \quad (6)$$

For the example that heating is a response to radiative flux convergence, then it may be that the temperature rises according to:

$$c_p \left(\frac{\partial T}{\partial t}\right)_{p,N_S} = \frac{1}{N_S} \left(\frac{\partial Q^{net}}{\partial t}\right)_{p,N_S} \quad (7)$$

where c_p is the specific heat of the substance at constant pressure and $\partial Q^{net}/\partial t$ is the radiative heating. In a materially closed system, the response to net heating is for the temperature to rise.

In the atmospheric sciences, Eq. 7 expresses how radiative heating is the driving force behind weather (Liou, 2002). At timescales longer than τ_{heat} , however, the establishment of a temperature gradient ultimately leads to a material flow that we call the wind.

²A well-known expression of this physics is the Gibbs-Duhem equation (Zemanksy and Dittman, 1997).

2.2 Systems in material disequilibrium

Over timescales much longer than τ_{heat} , the legs of the heat engine are not resolved. Instead, because the heat engine cycles are much faster than the timescales of interest, one only views some average level of μ_S that lies in between the points of maximum and minimum potential energy, μ_R and μ_E (Fig. 1).

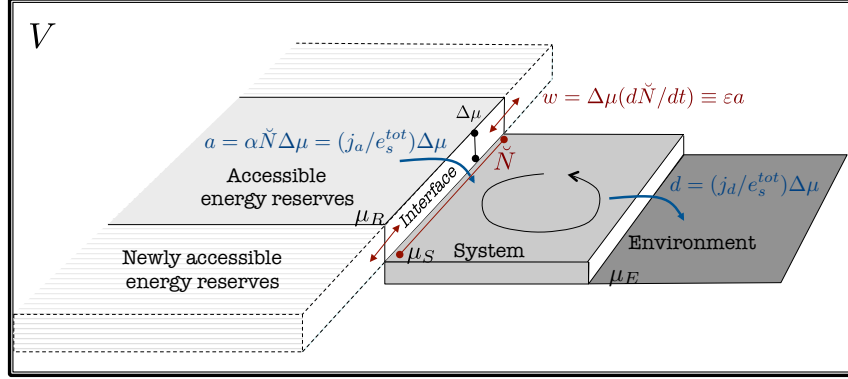


Figure 2: Schematic for the thermodynamic evolution of a system within a constant volume V . Energy reserves, the system, and the environment lie along distinct constant potential surfaces μ_R , μ_S , and μ_E . The size of an interface $\dot{N}\Delta\mu$ between surfaces determines the rate of heating a and the speed of downhill material flow j_a . The system grows or shrinks according to a net material flux convergence $j_a - j_d$ along μ_S . System growth is related to expansion work w that is done to grow the interface, extending the system's access to previously inaccessible energy reserves. The efficiency of work is determined by $\epsilon = w/a$.

In this case, energetic and material flows appear to be coupled. An illustration of this coupling is shown in Fig. 2, which recasts Fig. 1 in terms of a single coordinate. Where there is a disequilibrium, material convergence along a surface of constant potential μ_S corresponds with growth of the system enthalpy at rate

$$\left(\frac{\partial H_S}{\partial t}\right)_{\mu_S} = \left(\frac{\partial Q^{net}}{\partial t}\right)_{\mu_S} = e_S^{tot} \left(\frac{\partial N_S}{\partial t}\right)_{\mu_S} \quad (8)$$

so that from Eq. 5, the bulk grows at rate

$$\begin{aligned} \left(\frac{\partial N_S}{\partial t}\right)_{\mu_S} &= \frac{(\partial Q^{net}/\partial t)_{\mu_S}}{e_S^{tot}} \\ &= \frac{a - d}{e_S^{tot}} \end{aligned} \quad (9)$$

If there is zero time-averaged net heating, then $\langle (\partial Q^{net}/\partial t)_{\mu_S} \rangle = 0$ because $\langle a \rangle = \langle d \rangle$, in which case the size of the system N_S does not change. Like water pouring into

and draining from a bath tub at equal rates, circulations within the system maintain a steady-state³

Material growth occurs for the non-equilibrium condition that energy consumption exceeds dissipation, in which case $\langle (\partial Q^{net}/\partial t)_{\mu_S} \rangle > 0$. In this case, there is a net convergence of matter along the potential surface μ_S at rate j^{net} . Material flows into civilization at rate j_a and out of civilization at the decay rate j_d form a balance defined by

$$j^{net} = \left(\frac{\partial N_S}{\partial t} \right)_{\mu_S} = j_a - j_d \quad (10)$$

so that the timescale for growth of the system is $\tau_{growth} \sim N_S/j^{net}$. Combined with Eq. 9, this implies that

$$j_a = a/e_S^{tot} \quad (11)$$

$$j_d = d/e_S^{tot} \quad (12)$$

$$j^{net} = \frac{a-d}{e_S^{tot}} \quad (13)$$

A straightforward and familiar example of this physics is what happens when we boil a pot of water. Once the water reaches the boiling point, the temperature of the water is maintained at a constant 100° C, and the energy input from the stove goes into turning liquid water into bubbles. Setting aside the energetics of forming the bubble surface, and assuming the pot is well insulated, the energy input that is required to vaporize a single liquid water molecule is $e_S^{tot} = l_v$ where l_v is the latent heat of evaporation at boiling. Thus, vapor molecules contained in the bubbles are created at a rate that is proportional to the rate of energetic input: $j_a = a/e_S^{tot} = a/l_v$.

Heating creates an internal circulation of bubbles that we call a boil. When bubbles rise to the surface, molecules escape the fluid at rate j_d , and there is an associated evaporative cooling of the water at rate $d = j_d e_S^{tot} = j_d l_v$. With a steady simmer, a constant vapor concentration N_S is maintained within the pot because heating equals cooling. In this case, from Eq. 13, $j_a \simeq j_d$ and $j^{net} = 0$.

If the output from the heating element is suddenly raised to high, then there is a non-equilibrium adjustment period of $\tau_{growth} \sim N_S/(j_a - j_d)$ during which heating temporarily exceeds dissipation and bubble production at the bottom of the pot j_a exceeds bubble popping at its top j_d . The size and number of vapor bubbles in the water increases, and a new stasis is attained only when evaporative cooling d rises to come into equilibrium with the element heating a . At this point, the pot has gone from a simmer to a rolling boil.

2.3 Gradients and flows

As shown in Fig. 2, a material flow at rate j can be seen as a diffusion of matter downhill as it flows across a material interface. The interface between the system and its higher

³For the case of zero net heating, there is nonetheless an increase in global entropy even though local entropy production $(\partial Q^{net}/\partial t)_\mu/\mu = 0$. A continuous flow from high to low potential requires increasing global entropy $\sum_\mu (\partial Q^{net}/\partial t)_\mu/\mu$ because there is global redistribution of matter to low values of μ .

potential reservoirs can be defined by a potential step with a rise $\Delta\mu = \mu_R - \mu_S$ and an orthogonal quantity of material that lies along the interface \check{N} . The total energy required to grow the interface is the product of these two quantities: i.e., $\Delta G = \check{N}\Delta\mu$. Because the gradient enables flows, there is a proportional consumption of available potential energy ΔH_R at rate

$$a = \alpha\Delta G = \alpha\check{N}\Delta\mu \quad (14)$$

where α is a rate coefficient with units of inverse time. The quantity $\Delta G = \check{N}\Delta\mu$ in Eq. 14 differs from the available enthalpy $\Delta H_R = N_R\Delta\mu$. The available enthalpy is a reserve of energy, but it is ΔG that is associated with the gradient that drives flows across an interface.

From Eqs. 10 and 11, energy consumption is coupled to a material flux $j_a = (\partial N_S/\partial t)_{\mu_S}$. Thus, from Eq. 14:

$$j_a = \alpha\check{N}\Delta\mu/e_S^{tot} \quad (15)$$

The magnitude of the interface \check{N} reflects the respective sizes of the two components it separates. In general, when there is a diffusive flow to a system, \check{N} is proportional to a product of the available enthalpy within a high potential energy “reservoir” $\Delta H_R = N_R\Delta\mu$ and the size of the system N_S taken to a one third power (Garrett, 2012c), or that

$$\check{N} = kN_S^{1/3}N_R \quad (16)$$

where a dimensionless coefficient k is related to the object shape ⁴.

At first glance, one might guess that the system interface should be proportional $N_S N_R$ instead, since both the size of the system and the size of the reserve are what drive flows between the two. A system’s size is proportional to its volume $V_S = N_S/n_S$, where N_S is the number of elements in the system and n_S is the internal density; V_S and N_S are proportional to a dimension of length cubed, or volume. However, flows to a system are not determined by a volume. Rather, flows are down a linear gradient that lies normal to a surface. The surface area has dimensions of length squared or $N_S^{2/3}$, and the linear gradient has dimensions of inverse length or $N_S^{-1/3}$. Both factors control the flow rate, and their product yields a one third power or a length dimension: $N_S^{2/3} \times N_S^{-1/3} = N_S^{1/3}$.

In any case, if it were assumed that \check{N} is proportional to the product $N_S N_R$, then the implication would be that wholes are interacting with wholes. A perfect mixture of the system and its reserve, even if possible (which it is not), would make it impossible to resolve flows between N_S and N_R : the two components would be indistinguishable. Finally, assuming a unity exponent for N_S removes any element of persistence or memory from rates of system growth, as will be shown below. Unphysically, it would divorce what happens in the present from what has happened in the past.

Since $\Delta H_R = N_R\Delta\mu$, Eqs. 14 and 15 for energy dissipation and material flows can now be expressed as

$$j = \alpha k N_S^{1/3} \Delta H_R / e_S^{tot} \quad (17)$$

$$a = \alpha k N_S^{1/3} \Delta H_R \quad (18)$$

⁴For a system that is spherical with respect to its reserves then $k = (48\pi^2)^{1/3}$ (Garrett, 2012c)

In Garrett (2012c) it was shown that the quantity $\alpha k N_S^{1/3}$ can be expressed in an equivalent fashion in terms of a length density times a diffusivity $\Lambda \mathcal{D}$, where the length density is analogous to the electrostatic capacitance within a volume and the diffusivity has dimensions of area per time⁵. Thus, the flow and dissipation equations can be alternatively expressed as

$$j = \mathcal{D} \Lambda \Delta H_R / e_S^{tot} \quad (19)$$

$$a = \mathcal{D} \Lambda \Delta H_R \quad (20)$$

The rate of material flows is proportional to a rate of energy dissipation a , which in turn is proportional to some measure of the length density within the system Λ or its accumulated size N_S to a one third power, and the number of potential energy units in the reserve $N_R = \Delta H_R / \Delta \mu$. The final component is e_S^{tot} , which expresses the amount of energy that must be dissipated to enable each unit of material flow towards the system.

2.4 Efficiency and growth

As described above, a system grows if there is net heating that drives an imbalance between diffusive material flows (Eqs. 10 and 13), so that the size of the system N_S and interface with energy reserves $\Delta G = \check{N} \Delta \mu$ evolve over time.

Taking the approach that the resolved rise of the interface $\Delta \mu$ is fixed, then flows evolve as the magnitude of the “step” $\check{N} \Delta \mu$ grows laterally (Fig. 2). Here, this material expansion or “stretching” of the interface \check{N} and the potential difference ΔG is termed “work” w , where:

$$w = \left(\frac{\partial \Delta G}{\partial t} \right)_{\mu_R, \mu_S} = \left(\frac{\partial \check{N}}{\partial t} \right)_{\mu_R, \mu_S} \Delta \mu \quad (21)$$

The efficiency of converting heating to a rate of doing work is normally defined by the ratio

$$\varepsilon = \frac{w}{a} \quad (22)$$

Here, efficiency can be either positive or negative depending on whether the interface is shrinking or growing in response to heating, and therefore on the sign of w (Eq. 21).

From Eq. 21, the relative growth rate of the interface can be defined by

$$\eta = \frac{w}{\Delta G} = \frac{d \ln \Delta G}{dt} = \frac{d \ln \check{N}}{dt} \quad (23)$$

where η has units of inverse time. In other words, $1/\eta$ is the characteristic time for exponential growth of ΔG and \check{N} .

⁵A very simple example of this physics is the diffusional growth of a spherical cloud droplet of radius r through the condensation of water vapor, where $j = 4\pi r \mathcal{D} N_R / V$ and N_R / V is equivalent to the excess vapor density relative to saturation. In this case $\alpha k N_S^{1/3} \mathcal{D} = \Lambda \mathcal{D} = 4\pi r \mathcal{D} / V$. Note that a length dimension is what determines flows, insofar as it is coupled to available reserves of potential energy. For more dendritic structures like snowflakes, there is no clearly definable “radius”, yet it is still a length dimension Λ or “capacitance” that drives diffusive growth (Pruppacher and Klett, 1997).

Since, from Eqs. 21 and 22, $w = d\Delta G/dt = \varepsilon a$ and from Eq. 14, $a = \alpha\Delta G$, it follows that the relationship between the growth rate η and efficiency ε and heating a is given by

$$\eta = \alpha\varepsilon \quad (24)$$

$$= \frac{d\ln a}{dt} \quad (25)$$

which has the advantage of expressing η in terms of a measurable flux a . So, efficient systems grow faster to consume more. For the special case of pure exponential growth where η is a constant, then $a = a_0 \exp(\eta t)$, but, more generally, nothing is ever fixed in time: η constantly changes as the interface evolves, and it can even change sign if it shrinks. The growth rate η is positive if the efficiency ε is greater than zero meaning that the system is able to do net work on its surroundings in response to heating (i.e. $d\ln\check{N}/dt > 0$). Otherwise, the growth rate is negative and the system collapses (i.e. $\varepsilon < 0$ and $d\ln\check{N}/dt < 0$).

2.5 Emergence, diminishing returns, and decay

A pot of boiling pot of water has an external agency with its hand on the energetic flow. “Emergent systems” might be characterized by a spontaneous development of a structure. A way to view emergence is through Fig. 2, where heating and dissipation sustain internal circulations. If heating exceeds dissipation then a net incorporation of matter into the system allows it to expand into newly accessible energy reserves. The thermodynamic recipe for emergence is that sufficient energy reserves exist to be “discovered” that the disequilibrium that drives growth can be sustained.

While emergent phenomena are ubiquitous in nature, they might be most evident in living organisms who survive by eating, drinking and inhaling a matrix of matter and potential energy, which is then diffused through a linear of network of vascular structures. Consumption of the potential energy in carbohydrates, proteins and fats sustains the organism and facilitates an incorporation of water, chemicals, vitamins, minerals. Meanwhile, heat is dissipated, and matter is lost, through radiation, perspiration, exhalation, and excretion. The flow of raw materials and the dissipation of potential energy are coupled within cardiovascular, respiratory, gastro-intestinal and nervous networks. Over short timescales, dissipation simply allows for further consumption. In the long-run though, where consumption is in excess of dissipation, flows are out of equilibrium, and the organism networks grow. The demand for energy by the organism increases out of a requirement to sustain its growing network length and the associated internal circulations.

For a given availability of available energy supplies $\Delta H = N_R\Delta\mu$, then from Eqs. 16 and 23 the instantaneous growth rate is related to the system size N_S or its network

length density Λ through

$$\eta = \left(\frac{\partial \ln \check{N}}{\partial t} \right)_{N_R} \quad (26)$$

$$= \left(\frac{\partial \ln N_S^{1/3}}{\partial t} \right)_{N_R} \quad (27)$$

$$= \left(\frac{\partial \ln \Lambda}{\partial t} \right)_{N_R} \quad (28)$$

If the rate of emergent growth η is positive then a positive feedback loop dominates and this length dimension grows exponentially (i.e. $\Lambda = \Lambda_0 \exp(\eta t)$). Negative values of η correspond with decay.

From Eq. 10 and 27, the rate of emergent growth can be related to rates of material consumption j_a and decay j_d through:

$$\eta = \frac{1}{3N_S} \left(\frac{\partial N_S}{\partial t} \right)_{N_R} \quad (29)$$

$$= \frac{1}{3} \frac{j_a - j_d}{\int_0^t (j_a - j_d) dt'} \quad (30)$$

$$= \frac{1}{3} \frac{j^{net}}{\int_0^t j^{net} dt'} \quad (31)$$

Note that the timescale for growth of the system discussed earlier τ_{growth} is related to the growth rate of flows through $\eta = 3/\tau_{growth}$.

A ‘‘decay parameter’’ δ can be defined as the rate of material decay relative to the rate of material consumption:

$$\delta = \frac{j_d}{j_a} \quad (32)$$

and, since the current system size is the time integral of past net material flows, $N_S = \int_0^t j^{net} dt'$, it follows that the rate of emergent growth is given by:

$$\eta = \frac{1}{3} (1 - \delta) \frac{j_a}{N_S} \quad (33)$$

$$= \frac{1}{3} \frac{(1 - \delta) j_a}{\int_0^t (1 - \delta) j_a dt'} \quad (34)$$

The final step is to account for the motive force for current flows to the system, which is obtained by substituting Eq. 17 into Eq. 33 to yield

$$\eta = \alpha k (1 - \delta) \frac{N_S^{1/3} N_R \Delta \mu}{N_S e_S^{tot}} \quad (35)$$

$$= \alpha k (1 - \delta) \frac{\Delta H_R}{N_S^{2/3} e_S^{tot}} \quad (36)$$

Eq. 35 for emergent growth has seven parameters. Three – α , k and $\Delta\mu$ – are considered as constants in this treatment. So, current growth rates η are determined by the quantity of energy $\Delta H_R = N_R \Delta\mu$ that is available to drive material flows to the system; the amount of energy e_S^{tot} that must be dissipated to incorporate each unit of matter into the system; the fraction $1 - \delta$ of this new matter whose addition is not offset by decay; and, crucially, past flows leading to the current system size N_S : as a system gets bigger, there is a natural propensity for its growth rate to slow with time.

This last element leads to a “law of diminishing returns” and introduces memory to emergent growth. Note that, had it been assumed that flows were proportional to $N_S N_R \Delta\mu$ rather than $N_S^{1/3} N_R \Delta\mu$ in Eq. 17, then this dependence of current growth rates on past flows $\int_0^t j^{net} dt'$ would not be present – the N_S terms would have canceled in Eq. 35. Clearly, this would be inconsistent with our observations of emergent systems. Expressed logarithmically, large objects tend to grow more slowly than small objects. And, the growth of all emergent systems is somehow tied to their past. Systems are built from matter that was accumulated during prior growth. “Great oaks from little acorns grow”.

3 Thermodynamics of the growth of wealth

Taken as a whole, civilization might be viewed as another example of an emergent system that, like other living organisms, consumes “food” in the form of a matrix of matter and energy. The raw materials include water, wood, cement, copper and steel. The potential energy that is consumed is contained in fossil fuels, nuclear fuels, and renewables. The linear networks within civilization are our roads, shipping lanes, communication links, and interpersonal relationships.

Energy consumption at rate a enables civilization to raise raw materials across a potential energy barrier so that they can be incorporated through diffusion into civilization’s bulk at rate j_a . The amount of energy that is required to turn raw materials into the stuff of civilization is an enthalpy for rearranging matter into a new form. Section 2.2 included a discussion of how heating transforms liquid into vapor within a pot of boiling water. A similar “phase transition” might be seen when we burn oil to extract such things as iron ore and trees from the ground. Energy consumption continues as we reconfigure raw materials from their low potential, natural state into carefully arranged steel girders and houses where they becomes part of civilization’s structure.

In what we might call the economy, this energy consumption or heating sustains all of civilization’s existing internal circulations against the continuous dissipation of heat at rate d and material decay at rate j_d . Civilization radiates heat to space while we and our physical infrastructure fall apart. If civilization consumes energy at rate a , largely through the exothermic reaction of primary energy reserves (e.g. through combustion and nuclear reactions), and it dissipates energy at an equivalent rate d , then the size of civilization stays fixed. But if there is a disequilibrium where consumption exceeds dissipation, then a remnant of power is able to go towards incorporating new raw materials at rate $j_a - j_d$.

Civilization falls under a class of “emergent systems” because the disequilibrium allows civilization to expand into new reserves of raw materials and energy, leading to

a positive feedback that accelerates growth. From Eq. 26, growth rates are equivalent to the expansion of a length density Λ that is tied to the system's accumulated bulk to a one third power $N_S^{1/3}$. This suggests that growth rates can be thought of as a lengthening and concentration of the networks that form civilization's fabric. From Eq. 35, we can infer that civilization growth is promoted by the following factors: that civilization has access to large reserves of available energy $\Delta H_R = N_R \Delta \mu$; that the amount of energy e_S^{tot} that is required to incorporate raw materials into civilization's structure is low; and that civilization does not fray too quickly, such that the decay parameter $\delta = j_a / j_a$ expressing relative rates of decay is small⁶.

In what follows, these concepts are extended to provide specific formulations for the long-term evolution of civilization, expressible in such purely fiscal terms as rates of return on wealth, economic production, innovation, and technological change.

3.1 Expression of fiscal quantities in thermodynamic terms

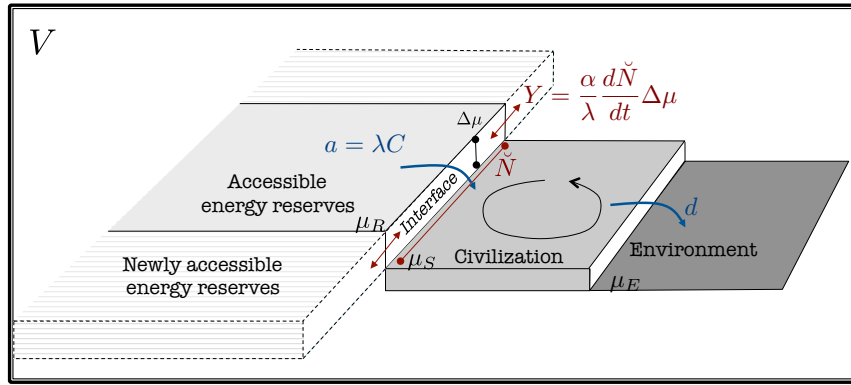


Figure 3: Representation of Fig. 2 in terms of global fiscal wealth C and economic production Y , as linked to rates of primary energy consumption a and the size of an interface with respect to energy reserves \check{N} . Economic production Y is tied to interface growth, representing an expansion of the capacity of civilization to draw from newly accessible energy reserves. Energy consumption sustains civilization circulations against dissipation to the environment at rate d .

In Garrett (2011), it was hypothesized that global rates of energy consumption a are linked to a very general metric of global economic wealth C through a constant λ :

$$a = \lambda C \quad (37)$$

where current wealth is viewed as the time integral of past inflation-adjusted economic production

$$C = \int_0^t Y(t') dt' \quad (38)$$

⁶As a practical matter, N_R might be expressed by civilization in units of millions of barrels of oil equivalent (mmboc), where the potential energy of combustion contained in one barrel is equivalent to $\Delta \mu$.

The motivation for these expressions was that global energy consumption at rate a sustains the internal circulations of civilization against an associated power dissipation d . If the capacity to sustain the global economy's circulations is what we implicitly value, then primary energy consumption should be fundamentally tied to a general representation of economic wealth (Fig. 3).

The hypothesis that λ is a constant is falsifiable. Since Gross Domestic Product GDP is the total productivity within a period of one year, Eq. 38 can be calculated from

$$C_i = \sum_i \text{GDP}_i \quad (39)$$

where i is a time index starting from the beginnings of civilization. Historical estimates of world GDP are available from such sources as Maddison (2003) and can be used to calculate C as outlined in Appendix C of Garrett (2011). Combined with available statistics for global primary energy consumption, Eq. 37 was shown to be supported by the data. Expressed in inflation-adjusted 2005 US dollars, available statistics indicate that λ has maintained a steady value for the past few decades for which global statistics for a are available. Effectively, what sustains the purchasing power embodied in each one thousand dollar bill, and distinguishes it from a mere piece of paper, is a continuous 7.1 ± 0.1 Watts of primary energy consumption.

Alternatively, in the year 2009, a global wealth of 2290 trillion U.S dollars was supported by 16.1 terawatts of primary energy consumption. In 1980, 1303 trillion 2005 dollars was sustained by 9.6 terawatts. In the interim, the ratio of these two quantities was essentially unchanged (Garrett, 2011, 2012a). Thus, it appears that fiscal wealth can be considered to be a human representation of the magnitude of the associated circulations that power consumption can support.

While wealth, as defined by Eq. 37, has units of currency and therefore might appear to be much like the term "capital" used in traditional economic treatments (e.g., Solow, 1956), there is a key difference. The term capital is normally reserved for the additive value of fixed "physical" structures such as buildings and roads. Economic output Y is not considered to be directly additive to physical capital because a portion is "consumed" by people rather than "saved" for the future. The motivation for this approach is that it seems logical to focus on people apart from non-living structures given that, after all, the economy is human; human labor uses physical capital to enable future consumption and certainly not the reverse.

This traditional approach offers a self-consistent way to track financial accounts, but it approaches economic growth as if it does not need to directly acknowledge universal physical laws. A lack of appeal to resource constraints has been pointed out by many others (e.g., Georgescu-Roegen, 1993; Costanza, 1980; Warr and Ayres, 2006). However, one point that has been missed is that the Second Law of Thermodynamics forbids the existence of isolated systems, either in space or time, and this places constraints on what an economic growth model should look like (Garrett, 2012a).

For example, where the human and physical components of civilization are interconnected, they cannot be mathematically treated as being independent. This means that labor cannot be easily separated from physical capital, and physical capital cannot be treated as being purely additive. All aspects of civilization are intertwined through

their networks. People need houses as much as houses need people in order to maintain their respective worth; removing one affects the worth of the other. In the same vein, human consumption cannot disappear to the past because the past is intertwined with the present. Even if someone is only “consuming” a hamburger, a hamburger is nourishing and satisfying in a way that both enables human interactions with the rest of civilization and carries a memory of the pleasures of hamburger consumption into the future. Even for ourselves, the thoughts in our brain cannot be meaningfully separated from our cardiovascular system and stomachs; each has no independent value since as each needs the others to work.

So there is no embodied value within any object itself, but only within its ties to other elements of civilization. A brick of solid gold is worth nothing if it is forgotten and lost in the middle of the desert, but much more if it facilitates financial flows through its integration within an economic network. Wealth includes people, their knowledge, their buildings, and their roads, but only insofar as they are interconnected through networks to the rest of the whole. The elements of networks cannot easily be treated as being mathematically additive, as in traditional economic treatments. Rather their value is only in how their relationships facilitate the internal circulations that demand civilization-scale thermodynamic flows.

So, here the approach is to treat civilization as a system with constant potential μ_S (Fig. 3) whose collective wealth is a fiscal expression of how its elements are intertwined in a way that mutually supports global scale diffusive and dissipative flows. From Eqs. 14 and 37,

$$C = \frac{\alpha}{\lambda} \tilde{N} \Delta \mu \quad (40)$$

where through Eq. 16, \tilde{N} is related to the system size through $N_S^{1/3}$ and a quantity of potential energy $N_R \Delta \mu$. Or, from Eq. 20,

$$C = \frac{\mathcal{D}}{\lambda} \Lambda \Delta H_R \quad (41)$$

The financial value of civilization lies in the total length density of a global network Λ , with the caveat that the total network must be coupled to reserves of potential energy ΔH_R that enable diffusive flows with diffusivity \mathcal{D} . Expressing the diffusion of knowledge and goods within human systems in terms of a network length density and a proximity to resources is in fact a common approach to human systems, albeit one that is normally discussed in less strictly thermodynamic terms (e.g., Barabási and Albert, 1999; Jackson, 2010; Bahar et al., 2012).

The complexity of civilization is extraordinary, and it would be extremely challenging if not impossible to model all possible interactions within the network. While nothing forbids looking at civilization’s internal components alone, as a first step, thermodynamic principles offer simplification of lowering resolution so that human and physical capital are regarded at global scales. With this approach, the trade-off is that nothing can be said about the internal details of civilization, except perhaps in a statistical sense (e.g., Ferrero, 2004). The advantage is that it enables a straight-forward link between physical and fiscal quantities. Stepping back to view civilization as a whole simplifies the relevant economic growth equations by removing the complexities of internal communications and trade.

3.2 Thermodynamics of nominal and inflation-adjusted economic production

Where fiscal wealth is defined holistically by $C = \int_0^t Y(t') dt'$ (Eq. 38), there appears to be a fixed relationship to thermodynamic flows through $a = \lambda C$ (Eq. 37). Thus, the very general physical principles derived in Section 2 can now be applied to derive economic production functions that are expressible in units of currency.

The simplest expression of the production function is that it adds to economic wealth as it has been defined above:

$$Y = \frac{dC}{dt} \quad (42)$$

where Y is inflation-adjusted (or real) economic output or productivity, with units of currency per time. However, since $a = \lambda C$, $w = \Delta\mu d\check{N}/dt$, and from Eqs. 40 and 41, any of the following expressions also apply, where α , \mathcal{D} and λ are constants:

$$Y = \frac{1}{\lambda} \frac{da}{dt} \quad (43)$$

$$= \frac{\alpha}{\lambda} w \quad (44)$$

$$= \frac{\alpha}{\lambda} \frac{d\check{N}}{dt} \Delta\mu \quad (45)$$

$$= \frac{\mathcal{D}}{\lambda} \frac{d}{dt} (\Lambda \Delta H_R) \quad (46)$$

Perhaps rather intuitively, economic production is directly tied to the amount of physical work w that is done to expand the capacity to consume energy through an increase in network density Λ and an expansion of available energy reserves ΔH_R . Real production is valuable to the extent that it accelerates the energetic flows a that sustain civilization's circulations. From Eq. 38 current global wealth is a consequence of past net work $C = (\alpha/\lambda) \int_0^t w dt'$. From Eq. 21, net work w expands the material interface \check{N} between civilization and the primary energy reserves that sustain it (Fig. 3). Net work is done where there is an imbalance between consumption and dissipation, allowing civilization to incorporate matter into its structure faster than it decays.

From Eqs. 24 and 37, a more purely fiscal expression of the production function is one that is related to wealth and rates of energy consumption through

$$Y = \frac{dC}{dt} = \eta C \quad (47)$$

where, η is the rate of emergent growth for thermodynamic systems. For economic systems, the rate of emergent growth η can be termed more fiscally as the "rate of return" since, like money in the bank, the rate of return expresses the growth rate of global wealth through

$$\eta = \frac{d \ln C}{dt} \quad (48)$$

In Garrett (2012b), it was argued that this rate of return η can be expressed in terms of two components $\eta = \beta - \gamma$, expressing a source and a sink, in which case production

is related to wealth through

$$\begin{aligned} Y &= (\beta - \gamma)C \\ &= \hat{Y} - \gamma C \end{aligned} \quad (49)$$

where β is a coefficient of nominal production, $\hat{Y} = \beta C$ is the nominal economic output, and γC is the magnitude of any correction to nominal production that is required to yield inflation-adjusted real production. From Eq. 29, this implies a link to the rates of material consumption and decay through,

$$\beta = \frac{j_a}{3N_S} \quad (50)$$

and

$$\gamma = \frac{j_d}{3N_S} \quad (51)$$

or, from Eq. 32

$$\gamma = \delta \frac{j_a}{3N_S} \quad (52)$$

Expressed thermodynamically, β can be viewed as a rate coefficient for growth and γ as a rate coefficient for decay, each with units of inverse time.

Normally, the the GDP deflator is what is used to represent the degree of any revisions to calculations of nominal output, i.e., the nominal GDP is revised downward by a factor \hat{Y}/Y . The GDP deflator is linked to inflation insofar that it is estimated from price changes in a very broad, moving basket of goods. For inter-annual calculations, the factor by which the nominal GDP must be adjusted to be compared to the nominal GDP in a prior year is:

$$\text{GDPDeflator} = \frac{\hat{Y}}{Y} \simeq 1 + \langle i \rangle \quad (53)$$

where $\langle i \rangle$ is the calculated average inflation rate for the year. Assuming the inflation rate is much less than 100% per year, it follows that

$$\langle i \rangle = \frac{\text{G}\hat{\text{D}}\text{P} - \text{GDP}}{\text{G}\hat{\text{D}}\text{P}} \simeq \frac{\hat{Y} - Y}{\hat{Y}} = \frac{\langle \gamma \rangle}{\langle \beta \rangle} \quad (54)$$

From Eqs. 50 and 52, this leads to the very simple result that global-scale inflation rates can be viewed as a fiscal expression of the decay parameter $\delta = j_d/j_a$:

$$\begin{aligned} \langle i \rangle &= \frac{\langle \gamma \rangle}{\langle \beta \rangle} \\ &\simeq \langle \delta \rangle = \frac{\langle j_d \rangle}{\langle j_a \rangle} \end{aligned} \quad (55)$$

The interpretation might be that civilization decay is an inflationary pressure on economic production because it “devalues” the productive capacity of existing assets by taking away that which has previously been built, learned, or born. This fraying

of networks occurs because people die or forget, buildings crumble, and machines oxidize. For example, it has been estimated that 10% of our twentieth century accumulation of steel has been lost to rust and war (Smil, 2006). Where human and physical networks fall apart, there is a diminished capacity to enable the thermodynamic flows that sustain civilization wealth. Any monetary assets that were previously created to support human and physical wealth no longer possess the same real purchasing power.⁷

To see the sources of inflationary trends, since $j_a = a/e_S^{tot}$ (Eq. 11) and $a \propto \Delta H_R$ (Eq. 17), then assuming that e_S^{tot} changes slowly:

$$\begin{aligned} \frac{d \ln \langle i \rangle}{dt} &= \frac{d \ln \langle j_d \rangle}{dt} - \frac{d \ln \langle j_a \rangle}{dt} \\ &\simeq \frac{d \ln \langle j_d \rangle}{dt} - \frac{d \ln \langle \Delta H_R \rangle}{dt} \end{aligned} \quad (56)$$

So, rising inflation might occur if material decay j_d accelerates, perhaps from the types of global scale natural disasters that might be associated with climate change (Zhang et al., 2007; Lobell et al., 2011). Alternatively, inflation might be driven by a declining availability of energy reserves ΔH_R (Bernanke et al., 1997).

As a caution, traditional interpretations of price inflation (e.g., Parkin, 2008) may not be a perfect match for the treatment described here. Pure price inflation is a form of devaluation that arises because existing monetary wealth has a lower purchasing power, so it is often viewed as being simply a matter for control by central banks.

However, the very general expression of wealth C that has been discussed here extends beyond money and physical assets to comprise our physical and human relationships. In this case, devaluation might arise because previously acquired skills might no longer be needed by others because our capacity for work goes idle for lack of an energetic impetus. Car production might decline if oil becomes scarce and expensive. The workers and their factories remain but the external demand for petroleum driven transportation declines and this leads to car manufacturer layoffs (Lee and Ni, 2002). Unemployment is just another side of a more general inflationary coin⁸.

Of course, governments might rebuild human networks through financial investments that bring workers back into paying jobs. But to have a sustained effect on economic output, the hope would need to be that these investments lead to a commensurate increase in energetic consumption (Eq. 43). Real civilization wealth and energy consumption are intertwined through $a = \lambda C$ (Eq. 37), where wealth is tied to a capacity to access resources (Eq. 40). Simply printing money does not add to real wealth; being able to access new energy reservoirs does. Stimulating the economy by loosening the availability of money may be associated with nominal production in the short term (Eq. 49); but, if it fails to ultimately create or be associated with a sustained increase in energy consumption, the thermodynamics suggests that there will be an offset to nominal wealth production through some combination of unemployment and price inflation (Eq. 54).

⁷Deflation (or negative inflation) is associated with $\langle i \rangle \simeq \langle \delta \rangle < 0$, which can be satisfied provided that $j_a < 0$. Negative raw material consumption might arise where raw materials are sourced from within rather than without.

⁸In fact, and apparent short-term trade-off between unemployment and price inflation is well known in the field of Economics and has been termed the ‘‘Phillips Curve’’ (Phillips, 1958).

4 Thermodynamics of technological change, innovation, and growth

Thus far, it has been shown that an economic growth model can be defined by the coupled equations for the production function for real output Y , and the growth of real wealth C given by $dC/dt = Y$ and $Y = \eta C$, where η is a variable rate of return on wealth. As described in Garrett (2011), these equations can be viewed as being a more thermodynamically based (and dimensionally self-consistent) form of the Solow-Swan neo-classical economic growth model (Solow, 1956), where C is a generalized form of physical capital (K) that encompasses labor (L), and η is analogous to the total factor productivity (A), whose changes relate to technological change ($d \ln A/dt$).

Technological change is often seen as a primary driver of long-run economic growth (Solow, 1957) but the source of technological change remains somewhat of a puzzle. Sometimes it is regarded as having endogenous origins, perhaps due to government investments in research and development (Romer, 1994). However, the forces behind technological change can also be seen in light of a more strictly thermodynamic context. The rate of return η evolves according to a deterministic expression obtained by taking the derivative of the logarithm of Eq. 35:

$$\frac{d \ln \eta}{dt} = -2 \frac{dN_S/dt}{3N_S} + \frac{d \ln(1-\delta)}{dt} + \frac{d \ln \Delta H_R}{dt} - \frac{d \ln e_S^{tot}}{dt} \quad (57)$$

$$\begin{aligned} &= -2\eta + \eta_\delta + \eta_R^{net} - \eta_e \\ &= -2\eta + \eta_{tech} \end{aligned} \quad (58)$$

Here, the term $d \ln \eta/dt$ is referred to as *economic innovation* because positive values of Eq. 57 represent an acceleration of existing rates of return η . Innovations are what are required for rates of return on wealth to rise. Defining $\tau_\eta = 1/(d \ln \eta/dt)$ as the characteristic time for innovation, then wealth grows from an initial value C_0 as

$$C = C_0 e^{\eta \tau_\eta (e^{t/\tau_\eta} - 1)} \quad (59)$$

If innovation is positive, then wealth grows explosively or super-exponentially. In the limit of no innovation and $\tau_\eta \rightarrow \infty$, the growth of wealth reduces to the simple exponential form $C = C_0 \exp(\eta t)$ (Garrett, 2011).

The sum $\eta_{tech} = \eta_\delta + \eta_R^{net} - \eta_e$ is termed here as the *rate of technological change* η_{tech} because it is the driving force behind innovation $d \ln \eta/dt$. It represents the sum of reductions to net decay (η_δ), rates of net energy reserve expansion (η_R^{net}), and reductions in the amount of energy required to access raw materials ($-\eta_e$). The following examines each component of technological change in more detail.

4.1 Innovation through increased longevity

The first component of technological change is η_δ , which relates to reductions in the decay parameter δ (Eq. 32). From Eq. 55, and assuming that the global inflation rate is

much less than 100%, the decay parameter is approximately equal to the inflation rate through $\langle i \rangle \simeq \langle \delta \rangle$. In this case, the first order expansion in η_δ yields

$$\eta_\delta = \frac{d \ln(1 - \delta)}{dt} \simeq - \frac{d \langle \delta \rangle}{dt} \quad (60)$$

Since $\delta = j_d / j_a$ (Eq. 32), one way of interpreting η_δ is through

$$\eta_\delta = - \frac{1}{j_a} \left(\frac{\partial j_d}{\partial t} \right)_{j_a} \quad (61)$$

or, for a given rate of material consumption j_a , innovation is favored by decreasing decay rates j_d . If people are enabled to live longer through advancements in health (Casasnovas et al., 2005), or their structures are built so that they last longer (Kalaitzidakis and Kalyvitis, 2004), then this is a form of positive technological change that contributes to faster growth. Since the decay parameter is related to the inflation rate, it follows that this innovative force would show up in global scale economic statistics as declining inflation. In other words:

$$\eta_\delta \simeq \frac{d \langle \delta \rangle}{dt} \simeq - \frac{d \langle i \rangle}{dt} \quad (62)$$

4.2 Innovation through discovery of energy reserves

The expression η_R^{net} refers to the net rate of expansion of available energy reserves ΔH_R . Having a newly plentiful supply of energy accelerates economic innovation and growth (Smil, 2006; Ayres and Warr, 2009).

There are two forces here. One is for energy reserves to decline due to potential energy consumption at rate a . The second is that civilization discovers new reserves of energy at rate D . The balance of these two forces is given by

$$\begin{aligned} \frac{d \ln \Delta H_R}{dt} &= \frac{\text{Discovery} - \text{Depletion}}{\text{Existing}} \\ &= \frac{D - a}{\Delta H_R} \\ &= \eta_D - \eta_R \end{aligned} \quad (63)$$

Net reserve expansion occurs when rates of reserve discovery η_D exceed rates of reserve depletion η_R , requiring that $\eta_D / \eta_R > 1$.

As illustrated in Fig. 3, civilization consumes energy as it grows, and it grows into surroundings that may or may not contain new reserves of fuel (Murphy and Hall, 2010). If $\eta_D / \eta_R > 1$, then civilization discovers new reserves faster than it depletes previously discovered reserves. In some global sense, energy becomes “cheaper” relative to the existing quantity of wealth C , allowing the rate of return on wealth η to be higher than it would be otherwise.

4.3 Innovation through increased efficiency of raw material extraction

The expression $-\eta_e$ in Eq. 57 refers to changes in the specific enthalpy of civilization e_S^{tot} . Since $e_S^{tot} = a/j_a$ (Eq. 11), a decline in e_S^{tot} would appear as a decrease in the amount of power a that is required for civilization to extract raw materials and incorporate them into civilization at rate j_a .

Comparing Eqs. 28 and 29, civilization networks grow through raw material consumption. If growing civilization requires less energy per unit matter, then civilization can grow faster for any given rate of global energy consumption a . Since $a = \lambda C$, the implication is that raw materials have become cheaper relative to total global wealth C . This is an innovative force because the material growth of civilization increases access to the resources that sustain it.

4.4 Innovation through a common culture

I suggest that a second interpretation of the expression $-\eta_e$ in Eq. 57 is that economic innovation can be derived from seeking a common culture. The specific enthalpy of civilization can be expressed as $e_S^{tot} = \nu e_S$ (Eq. 2), where e_S is the specific energy of each independent mode and ν represents the number of orthogonal (or independent) modes within a mechanical system. The energy associated with each mode e_S can be assumed to be equal through the equipartition principle, provided sufficiently long timescales are considered.

For the purpose of facilitating the thermodynamics in this treatment, civilization is considered to lie along a surface of constant potential, in which case e_S does not change. However, the number of degrees of freedom ν in the system remains a free parameter. If innovations occur when e_S^{tot} declines, this is due to a decrease in ν .

For guidance on what this rather abstract thermodynamic result actually means, it might help to first look at the behavior of a “stiff” molecule such as gaseous molecular nitrogen (N_2). Nitrogen can be idealized as two nitrogen atoms connected by a stiff spring. At room temperatures, N_2 has five orthogonal degrees of freedom that determine its specific enthalpy e_S^{tot} . Three of these come from molecular translational motions within the three dimensions of space; two come from orthogonal rotational motions. Two additional degrees are added to account for molecular pressure to yield $\nu = 7$. If the specific energy (or temperature) increases ten-fold, vibrational transitions within N_2 no longer stay “frozen out”, and ν increases from seven to nine.

So, molecules that are more internally “stiff” have a smaller number of independent degrees of freedom for molecular motion. At room temperatures, N_2 has relatively low values of $e_S = kT/2$ that prevent the individual atoms from oscillating independently. The stiffness of the bond requires that the two nitrogen atoms rotate and translate *together*, as if they were connected.

With regards to civilization, we have witnessed an extraordinary increase in internal connectivity through ever improving transport and communications networks (van Dijk, 2012). A way to interpret this growth in network density is that it corresponds to a reduction in the effective number of degrees of freedom in society. Technological change gives us a more collective experience and global culture.

For example, international trade allows us to consume very similar products; transportation in the form of petroleum fueled cars is now ubiquitous; and, we have now accepted English as our global *lingua franca*. Through communications, travel, international markets and shipping, our world has become more interconnected and “stiff”.

There are some obvious tradeoffs to cultural similarity, setting aside that a more homogeneous world is less interesting. Economic growth and volatility is now sensed more globally. On one hand, civilization might be fragile if it becomes too uniformly reliant on the same things. A potato monoculture became susceptible to a blight that trade had brought in from the New World, leading to a catastrophic population collapse known as the “Irish Famine” (Donnelly Jr., 2001). A modern parallel is the proposition that our reliance on oil will lead to a dramatic slowing of global economic growth should it rapidly become scarce (Lee and Ni, 2002; Bardi and Lavacchi, 2009; Sorrell et al., 2010; Murray and King, 2012).

On the other hand, increasing global common modes of transportation, communication, and language facilitate innovation and growth. We build roads for a reason, because they bring us together. By lowering the specific enthalpy $e_S^{tot} = ve_S = a/j_a$ and reducing the effective number of degrees of freedom v , less energy is required to diffuse an equivalent quantity of matter through the structure when new resources are uncovered.

4.5 Diminishing returns as a drag on innovation

The final term in Eq. 57, -2η , expresses a drag on how fast rates of return can grow. Innovation naturally slows due to a *law of diminishing returns*. In the absence of innovation, wealth converges on a steady-state where rates of return equal zero (Romer, 1986). Diminishing returns exists as a force because current growth unavoidably becomes diluted within the accumulated bulk that was built from past growth (Eq. 35). Each incremental incorporation of raw materials into civilization j^{net} has a decreasing impact relative to the summation of previously incorporated matter $\int_0^t j^{net}(t') dt'$. Diminishing returns makes innovation rates increasingly negative and can only be overcome if technological changes are sufficiently rapid. From Eq. 57, what is required is that $\eta_{tech} > 2\eta$.

5 Modes of growth in economic systems

5.1 Technological change and rates of return

Because the above expressions are prognostic, the implication is that there exist deterministic solutions for how rates of return on civilization wealth and energy consumption change with time. Previous work has identified characteristic sigmoidal or logistic behavior in the effects of technological change on economic growth: after overcoming a period of initial resistance, technological changes rapidly accelerate growth, followed ultimately by saturation (Landes, 2003; Smil, 2006; Marchetti and Ausubel, 2012). In-

deed, Eq. 57 can be expressed in the form of the logistic equation

$$\frac{d\eta}{dt} = \eta_{tech}\eta - 2\eta^2 \quad (64)$$

If rates of technological change η_{tech} are constant, then the solution has the sigmoidal or ‘‘S-curve’’ form

$$\eta(t) = \frac{G\eta_0}{1 + (G-1)\exp(-\eta_{tech}t)} \quad (65)$$

where η_0 is the initial value for the rate of return, and

$$G = \frac{\eta_{tech}}{2\eta_0} \quad (66)$$

represents a ‘‘Growth Number’’ (Garrett, 2012c) that partitions solutions for $\eta(t)$ into varying modes of growth summarized in Table 1.

Table 1: Modes of growth in economic systems. Diminishing returns (DR), Technological Change (TC), Technological Decline (TD).

| | Innovation | DR and TC | DR and TD | Decay | Collapse |
|-------------------------------------|-----------------|-----------------|--------------|--------------|--------------|
| Growth number $\eta_{tech}/2\eta_0$ | $G > 1$ | $0 < G < 1$ | $G < 0$ | $G > 1$ | $G < 1$ |
| Initial rate of return | $\eta_0 > 0$ | $\eta_0 > 0$ | $\eta_0 > 0$ | $\eta_0 < 0$ | $\eta_0 < 0$ |
| Limiting rate of return | $\eta_{tech}/2$ | $\eta_{tech}/2$ | 0 | 0 | $-\infty$ |

The four modes of growth that are available to civilization are innovation, diminishing returns, decay, and collapse, depending on the value of G and the initial value of η . Innovation is characterized by growing rates of return; diminishing returns is associated with declining rates of return, either to a limit of $\eta_{tech}/2$ or to zero. Where rates of return are initially negative, decay either slows with time or it accelerates in a mode of collapse.

Fig. 4 carves these modes within a space of η_{tech} and η , along with associated trajectories for any given value of η_{tech} . For example, for values of $G > 1$, civilization is in a mode of innovation because technological innovation is sufficiently rapid to overcome diminishing returns. At first, rates of return increase exponentially but then they saturate to approach a value of $G\eta_0 = \eta_{tech}/2$. If η is initially 1 % per year and rates of technological change η_{tech} are sustained at a nominal 4% per year, then one would expect rates of return η to grow sigmoidally towards 2% per year. The exponential phase of the sigmoidal growth would have a characteristic time of $1/\eta_{tech}$, or 25 years.

5.2 Technological change and GDP growth

Innovation rates have a direct impact on rates of GDP growth. Since $Y = \eta C$ (Eq. 72), and the rate of return is given by $\eta = d \ln C / dt$ (Eq. 48), it follows that:

$$\frac{d \ln Y}{dt} = \eta + \frac{d \ln \eta}{dt} \quad (67)$$

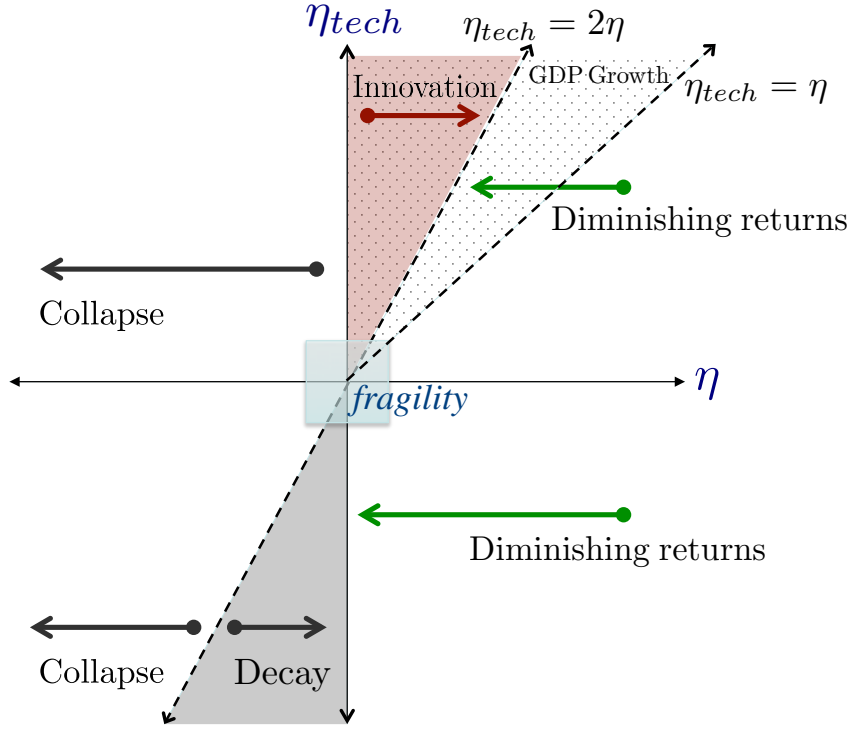


Figure 4: Modes of growth in economic systems, partitioned within a space of rates of technological change η_{tech} and rates of return on wealth η . Arrows represent trajectories for rates of return, assuming that η_{tech} is a constant. The dotted region shows the domain of parameter space associated with GDP growth. See text for details.

GDP growth rates are a simple sum of the current rate of return η and the innovation rate $d \ln \eta / dt$. GDP growth increases when there is innovation.

From Eq. 57, the rate of return itself evolves at rate $d \ln \eta / dt = -2\eta + \eta_{tech}$, where rates of technological change $\eta_{tech} = \eta_{\delta} + \eta_R^{net} - \eta_e$ are a summation of reductions to net decay, net energy reserve expansion, and improvements to the efficiency of raw material extraction and incorporation into civilization. Since GDP growth is related to the sum of the innovation rate and the rate of return in Eq. 67, it follows that:

$$\frac{d \ln Y}{dt} = -\eta + \eta_{tech} \quad (68)$$

The GDP growth rate is buoyed by positive technological change. However, as illustrated in Fig. 4, sustaining a growing GDP in the long-term requires that:

$$\eta_{tech} > \eta \quad (69)$$

or that technological change must be more rapid than the current rate at which energy consumption is growing $\eta = d \ln a / dt$.

In fact, this poses an interesting quandary for economic growth. Supposing that technological change is driven by net discovery of new energy reserves (Eq. 63), Eq. 69 implies that sustaining GDP growth requires energy consumption to continue to grow sufficiently fast that

$$\frac{da}{dt} > \Delta H_R (Da - a^2) \quad (70)$$

Eq. 70 is a logistic equation for energy consumption (Bardi and Lavacchi, 2009; Höök et al., 2010). Where energy consumption rates a are buoyed in the present by discovery of new reserves at rate D , this acts as a drag on growth further down the road. Other forms of technological change staying constant, the GDP approaches a steady-state when consumption equals discovery and $a = D$.

5.3 Fragility and growth

How do civilizations ultimately decay and collapse? Obviously, rates of return for civilization wealth must initially be positive for civilization to have emerged in the first place. But positive growth cannot be sustained forever because civilization networks are always falling apart to some degree. And, on a world with finite resources, we will eventually lose the capacity to keep fixing them.

However, there is no spontaneous mathematical transition between modes of growth that is implied by Eq. 65; in the limit of $t \rightarrow \infty$, rates of return η either asymptotically approach a constant value, or they tend towards collapse. So transitions between modes must be forced by some external impetus. In our case, this might come from a rapid increase in global scale natural disasters, perhaps due to climate change.

For example, in Eq. 35, if the decay parameter $\delta = j_d/j_a$ is greater than unity, then η must be negative. If material decay exceeds material consumption, then civilization transitions from positive to negative rates of return. Of course, things can go both ways and Eq. 35 also allows for conditions that might suddenly favor growth, including decreased decay or significant discoveries of new energy reserves that increase ΔH_R . These might permit civilization to transition from a mode of diminishing returns into one of innovation and super-exponential growth.

To account for both the good and the bad in the future, stochastic and largely unpredictable external events might be represented by introducing noise to Eq. 35. An example of how this might play out is illustrated in Figs. 5 and 6. If there is no noise, then trajectories follow the logistic solutions provided by Eq. 65. But if random Gaussian noise is added to η , then the range of possible trajectories broadens. Notably, there are “unlucky” trajectories that could be associated with frequent and persistent global scale natural disasters. Disasters might push civilization into a transition towards a mode of irreversible decay or collapse. Most notably, a transition is particularly likely when rates of return approach zero.

It has been pointed out that the existence of “tipping points” where there has been a “slowing down” is as a feature of ecological and climate systems (Dakos et al., 2008, 2011). What is interesting in the simulations above is that the most dramatic rates of collapse are associated with noisy trajectories that would otherwise be associated with innovation and accelerating rates of growth. The same conditions that allow for the human system to respond especially quickly to favorable conditions are the some ones

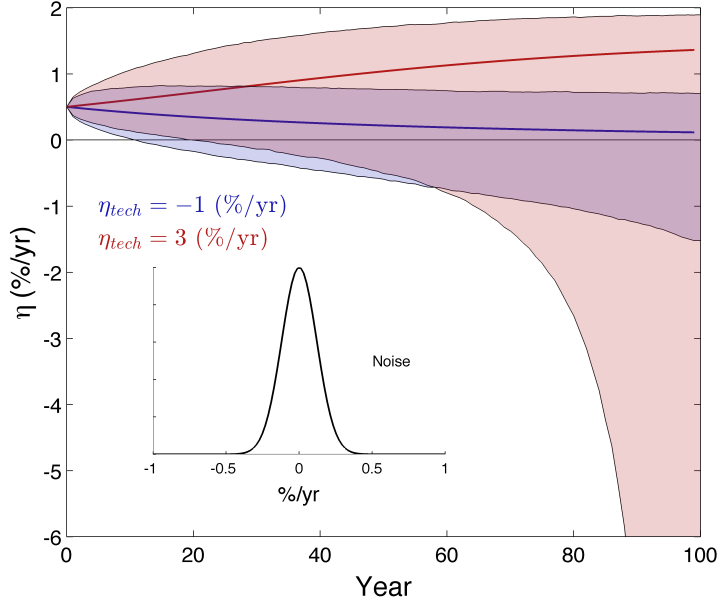


Figure 5: For an initial value for the rate of return η_0 of 0.5% per year, lines are trajectories of the evolution of $\eta(t)$ for scenarios with rates of technological change η_{tech} of 3% per year and -1% per year, as given by Eq. 65. The shaded region is derived from the upper and lower 5% bounds in an ensemble of 10,000 simulations where noise (inset) has been introduced that has a standard deviation of 0.1% per year for η .

that allow the system to rapidly decay quickly when conditions become unfavorable. Having a common culture is a good example (Sec. 4.4). It allows for exceptionally rapid diffusion of matter into civilization’s structure while also lending a fragility that permits co-ordinated decline.

As illustrated in Fig. 4, an innovative economy that enjoys relatively rapid technological change with a growth number $G > 1$ might alternatively be viewed as a “bubble economy” that lacks long-term resilience. Whether collapse comes sooner or later depends on the quantity of energy reserves available to support continued growth and the accumulated magnitude of externally imposed decay. By contrast, an economy that is less innovative, with lower rates of return η , has a lower risk of rapid rates of decline. In the space shown in Fig. 4, it lies “farther away” from modes of collapse.

6 Summary

This paper has presented a physical basis for interpreting and forecasting global civilization growth, by treating it as a thermodynamic system that grows in response to

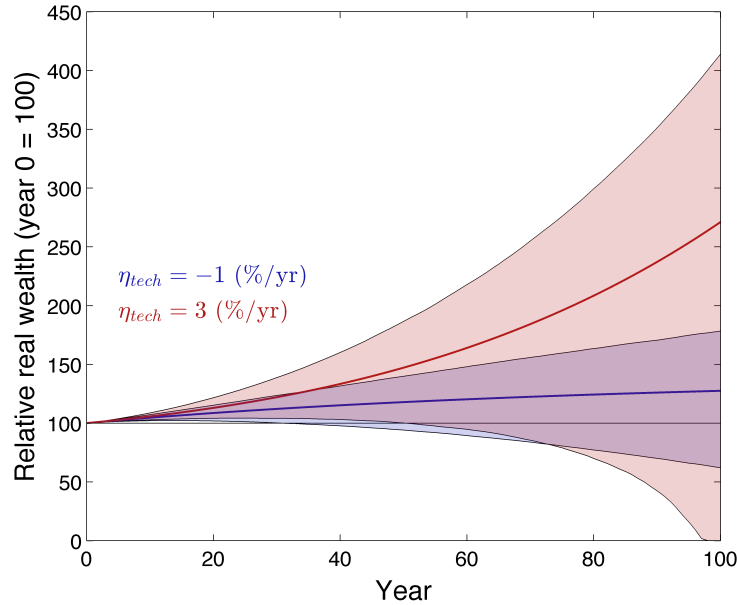


Figure 6: For the scenarios shown in Fig. 5, corresponding values of global inflation-adjusted wealth, referenced to 100 in year 0.

interactions with its environment (Garrett, 2012c). Like other living organisms (Vermeij, 2008), civilization displays spontaneous emergent behavior. Energy dissipation drives material flows to civilization. If there is a net convergence of matter within civilization, then civilization grows. Growth increases the availability of new reserves and this leads to a positive feedback loop that allows growth to persist.

The negative feedback on growth is that civilization carries with it a memory of its past. This slows growth through a “law of diminishing returns” that is common to growing systems: current additions of matter become increasingly diluted within an accumulation of past additions. Diminishing returns can be overcome, but only if there is sufficiently rapid technological change. Technological change has three broad categories: improved material longevity (or reducing decay), the discovery of new reserves of energy, and increased energy efficiency. One manifestation of higher energy efficiency might be a common global culture with fewer independent degrees of freedom, because this decreases the amount of energy that is required to diffuse raw materials throughout civilization’s structure.

These thermodynamic results can be expressed in purely fiscal terms because there appears to be a fixed link between global rates of primary energy consumption and a very general expression of human wealth: $\lambda = 7.1 \pm 0.1$ Watts of primary energy consumption is required to sustain each one thousand dollars of civilization value, adjusting for inflation to the year 2005 (Garrett, 2012a). Wealth does not rest in inert “physical capital”, as in traditional treatments, but rather in the density of connections

between civilization elements, insofar as this network contributes to a global scale consumption and dissipation of energy (Eq. 41).

The economic growth model for wealth C and economic production Y is very simple:

$$\frac{dC}{dt} = Y \quad (71)$$

$$Y = \eta C \quad (72)$$

where η is a variable real rate of return on wealth, somewhat analogous to the total factor productivity in traditional models. The rate of return can be related to basic thermodynamic quantities through

$$\eta = \alpha k (1 - \delta) \frac{\Delta H_R}{N_S^{2/3} e_S^{tot}} \quad (73)$$

where δ relates civilization decay to how fast it incorporates new raw materials, ΔH_R represents the quantity of available energy reserves, e_S^{tot} expresses the amount of energy required to incorporate raw materials into civilization's structure, and $N_S = \int_0^t j^{net} dt'$ represents the accumulated size of civilization due to past raw material flux convergence j^{net} . The constants α and k are unknown rate and shape coefficients; however, values of the rate of return η can be inferred from Eq. 72. For example, current global rates of return are about 2.2 % per year (Garrett, 2012a). What Eq. 73 shows is that trends in η can be forecast based on estimates of future decay and rates of raw material and energy reserve discovery.

Thus, Eqs. 71 through 73, combined with the constant λ , offer a complete set of prognostic expressions for civilization growth. The implications that have been described are summarized as follows:

- Civilization inflation-adjusted wealth grows only as fast as rates of global energy consumption.
- Low inflation is maintained by high civilization longevity.
- Rates of return on wealth decline when decay accelerates, or reserves of raw materials and energy become increasingly scarce.
- Through a law of diminishing returns, high current rates of return imply a stronger drag on future growth. The mathematical form for the evolution of rates of return is sigmoidal, as determined from the logistic equation.
- Rates of return grow when there is "innovation". As it is defined, innovation is driven by technological change, but it must be sufficiently fast to outweigh the law of diminishing returns.
- Global GDP growth requires energy consumption to grow super-exponentially, or at an accelerating rate. GDP growth is sustainable for as long as energy reserve discovery exceeds depletion.

- When growth rates slow and rates of return approach zero, civilization becomes fragile with respect to externally forced decay. It lies along a tipping point that might easily lead to a mode of accelerating decay or collapse.
- Innovation and collapse are two sides of the same coin. Increased internal connectivity allows for explosive growth when times are good, but also for exceptionally fast decline when times turn bad.

Many of these conclusions might seem intuitive, or as if they have been expressed already by others from a more traditional economic perspective. What is novel in this study is the expression of the economic system within a deterministic thermodynamic framework where a very wide variety of economic behaviors are derived from only a bare minimum of ingredients. A sufficient set of statistics exists for global economic productivity, inflation, energy consumption, raw material extraction and energy reserve discovery that the model presented here can be evaluated, and with no requirement for *a priori* tuning or fitting to historical data. If the analytical expressions are consistent with past behavior, then this offers the possibility of providing a range of physically constrained forecasts for future economic innovation and growth. If not, then the model should be re-examined or discarded.

A follow-on paper will compare these prognostic formulations against historical data. Civilization has enjoyed explosive growth since the industrial revolution, but it is unclear how long this can be sustained when it is facing ongoing resource depletion, pollution, and climate change. The prognostic expressions that have been derived here will be used to guide a physically plausible range of future timelines for civilization growth and decay.

Acknowledgments

This work was supported by the Kauffman Foundation, whose views it does not claim to represent.

References

- Ayres, R. U. and Warr, B.: The economic growth engine, Edward Elgar, Cheltenham, UK, 2009.
- Bahar, D., Hausmann, R., and Hidalgo, C.: International Knowledge Diffusion and the Comparative Advantage of Nations, 2012.
- Barabási, A.-L. and Albert, R.: Emergence of scaling in random networks, *Science*, 286, 509–512, 1999.
- Bardi, U. and Lavacchi, A.: A simple interpretation of Hubbert’s model of resource exploitation, *Energies*, 2, 646–661, doi:10.3390/en20300646, 2009.
- Bernanke, B. S., Gertler, M., Watson, M., C., S., and Friedman, B. M.: Monetary policy and the effects of oil price shocks, *Brookings Papers on Economic Activity*, 1, 91–157, 1997.

- Blois, J. L., Williams, J. W., Fitzpatrick, M. C., Jackson, S. T., and Ferrier, S.: Space can substitute for time in predicting climate-change effects on biodiversity, *Proc. Nat. Acad. Sci.*, 110, 9374–9379, doi:10.1073/pnas.1220228110, URL <http://www.pnas.org/content/110/23/9374.abstract>, 2013.
- Casasnovas, G. L., Rivera, B., and Currais, L.: *Health and Economic Growth*, MIT Press, Cambridge, USA, 2005.
- Costanza, R.: Embodied energy and economic valuation, *Science*, 210, 1219–1224, 1980.
- Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., and Held, H.: Slowing down as an early warning signal for abrupt climate change, *Proc. Nat. Acad. Sci.*, 105, 14 308–14 312, doi:10.1073/pnas.0802430105, 2008.
- Dakos, V., Kfi, S., Rietkerk, M., van Nes, E. H., and Scheffer, M.: Slowing Down in Spatially Patterned Ecosystems at the Brink of Collapse, *The American Naturalist*, 177, E153–E166, 2011.
- Donnelly Jr., J. S.: *The Great Irish Potato Famine*, Sutton Publishing, 2001.
- Ferrero, J. C.: The statistical distribution of money and the rate of money transference, *Physica A: Statistical Mechanics and its Applications*, 341, 575 – 585, doi:10.1016/j.physa.2004.05.029, 2004.
- Garrett, T. J.: Are there basic physical constraints on future anthropogenic emissions of carbon dioxide?, *Clim. Change*, 3, 437–455, doi:10.1007/s10584-009-9717-9, 2011.
- Garrett, T. J.: Can we predict long-run economic growth?, *Retirement Management Journal*, 2, 53–61, 2012a.
- Garrett, T. J.: No way out? The double-bind in seeking global prosperity alongside mitigated climate change, *Earth Sys. Dynam.*, 3, 1–17, doi:10.5194/esd-3-1-2012, 2012b.
- Garrett, T. J.: Modes of growth in dynamic systems, *Proc. Roy. Soc. A*, 468, 2532–2549, doi:10.1098/rspa.2012.0039, 2012c.
- Georgescu-Roegen, N.: *Valuing the Earth: Economics, Ecology, Ethics*, chap. The entropy law and the economic problem, pp. 75–88, MIT Press, 1993.
- Höök, M., Zittel, W., Schindler, J., and Aleklett, A.: Global coal production outlooks based on a logistic model, *Fuel*, 89, 3456–3558, doi:10.1016/j.fuel.2010.06.013, 2010.
- Jackson, M. O.: *Social and economic networks*, Princeton University Press, 2010.
- Kalaitzidakis, P. and Kalyvitis, S.: On the macroeconomic implications of maintenance in public capital, *Journal of Public Economics*, 88, 695–712, 2004.

- Kittel, C. and Kroemer, H.: Thermal physics, WH Freeman & Co, 1980.
- Landes, D. S.: The Unbound Prometheus: Technological Change and Industrial Development in Western Europe from 1750 to the Present, Cambridge University Press, 2003.
- Lee, K. and Ni, S.: On the dynamic effects of oil price shocks: a study using industry level data, *Journal of Monetary Economics*, 49, 823–852, 2002.
- Liou, K.: An Introduction to Atmospheric Radiation, International Geophysics Series, Academic Press, 2002.
- Lobell, D. B., Schlenker, W., and Costa-Roberts, J.: Climate trends and global crop production since 1980, *Science*, doi:10.1126/science.1204531, 2011.
- Maddison, A.: The World Economy: Historical Statistics, OECD, 2003.
- Marchetti, C. and Ausubel, J. H.: Quantitative Dynamics of Human Empires, *Int. J. Anth.*, 27, 1–62, 2012.
- Murphy, D. J. and Hall, C. A. S.: Year in review—EROI or energy return on (energy) invested, *Ann. New York Acad. Sci.*, 1185, 102–118, doi:10.1111/j.1749-6632.2009.05282.x, 2010.
- Murray, J. and King, D.: Climate policy: Oil’s tipping point has passed, *Nature*, 481, 433–435, doi:10.1038/481433a, 2012.
- Parkin, M.: The New Palgrave Dictionary of Economics, 2nd Ed., chap. Inflation, Palgrave Macmillan, doi:10.1057/9780230226203.0791, 2008.
- Phillips, A. W.: The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957, *Economica*, 25, 283–299, doi:10.1111/j.1468-0335.1958.tb00003.x, URL <http://dx.doi.org/10.1111/j.1468-0335.1958.tb00003.x>, 1958.
- Pruppacher, H. R. and Klett, J. D.: Microphysics of Clouds and Precipitation, 2nd Rev. Edn., Kluwer Academic Publishing, Dordrecht, 1997.
- Romer, P. M.: Increasing returns and long-run growth, *The Journal of Political Economy*, pp. 1002–1037, 1986.
- Romer, P. M.: The origins of endogenous growth, *J. Econ. Perspect.*, 8, 3–22, 1994.
- Smil, V.: Technical Innovations and Their Consequences, Oxford University Press, 2006.
- Solow, R. M.: A contribution to the theory of economic growth, *Q. J. Econ.*, 1970, 65–94, 1956.
- Solow, R. M.: Technical change and the aggregate production function, *Rev. Econ. Stat.*, 39, 312–320, 1957.

- Sorrell, S., Speirs, J., Bentley, R., Brandt, A., and Miller, R.: Global oil depletion: A review of the evidence, *Energy Pol.*, 38, 5290–5295, doi:10.1016/j.enpol.2010.04.046, 2010.
- van Dijk, J.: *The network society*, 3rd Edition, Sage Publications Ltd, 2012.
- Vermeij, G. J.: *Comparative economics: Evolution and the modern economy*, *J. Bioecon.*, in press, 2008.
- Warr, B. and Ayres, R.: REXS: A forecasting model for assessing the impact of natural resource consumption and technological change on economic growth, *Struct. Change Econ. Dyn.*, 17, 329 – 378, doi:DOI:10.1016/j.strueco.2005.04.004, 2006.
- Zemankys, M. W. and Dittman, R. H.: *Heat and Thermodynamics*, McGraw-Hill, 7th edn., 1997.
- Zhang, D. D., Brecke, P., Lee, H. F., He, Y.-Q., and Zhang, J.: Global climate change, war, and population decline in recent human history, *Proceedings of the National Academy of Sciences*, 104, 19 214–19 219, doi:10.1073/pnas.0703073104, 2007.