DEVELOPMENT OF ENERGY AND TRANSPORT SYSTEMS

N. Nakićenović and A. Grübler

1. Introduction

Economic development and the advancement of technology is presented as a process of substituting old forms of satisfying human needs by new ones or, more precisely, as a sequence of such substitutions. The examples, reconstructed from historical records, describe the quantitative, technological changes in energy and transport systems.

The analysis of historical replacement of old by new technologies has shown that most of these processes can be described by simple rules that are captured in the logistic substitution model (see Marchetti, 1979; Marchetti and Nakićenović, 1979; Nakićenović, 1984; and Grübler, 1988), and that technological substitution, expressed in terms of market shares, follows characteristic S-shaped curves. In order to illustrate and describe the properties of the approach we will first give examples of how new energy forms replaced their predecessors, since technological changes in the energy system constitute one of the first and most complete applications of logistic substitution analysis. To further explore this method we then describe similar substitution processes in transport systems.

The application of the logistic substitution model to the above examples indicates that improvements and growth are achieved through a regular but discontinuous process. Each new technology goes through three distinct substitution phases: growth, saturation and decline. This regular pattern points to a certain schedule and recurrence in structural change of competitive markets.

2. Dynamics of Energy Substitution

2.1. Technological Change

Substituting an old way by a new one to satisfy a given need has been the subject of a large number of studies. Most of these studies analyzed the substitution of products or technological processes. One general finding is that almost all substitutions of an old product or technology by a new one, expressed as fractional market shares, follow characteristic S-shaped curves with increasing substitution rates at the beginning and decreasing rates toward the end of the market takeover. Typical examples include the substitution of soaps by detergents, corn by hybrid corn seeds, black-and-white by color TV sets, and various manufacturing processes for making turpentine, paints, steel and so on.

Most of the studies of technological substitution are based on the use of the logistic curve to represent the substitution process. The logistic function, however, is not the only S-shaped function, but it is perhaps the most suitable one for empirical analysis of growth and substitution processes because both the shape and the form of the function correspond to the theoretical explanation of the S-shaped growth processes. The function is symmetrical around its point of inflection and the *relative* rate of increase (actual growth-rate over the achieved growth) declines linearly with increasing level of growth. In other words, the growth-rate is proportional both to the achieved growth-level and the possible growth left to the asymptotic level; the growth-rate increases up to the inflection point and decreases afterwards. In addition, the parameters of the function represent

meaningful attributes of the growth process - the growth-rate, the asymptotic growth-level and the location of the function in time.

Griliches (1957), in his study of the diffusion of the hybrid corn seed in the United States, was one of the first to use the S-shaped curve to describe technological substitution. He showed that hybrid corn replaced traditional corn seed in different States in a very similar way; the S-shaped substitution only being displaced in time by a few years and lasting for longer or shorter periods in different States.

Following the work of Griliches, Mansfield (1961) developed a model to explain the rate at which firms follow an innovator. He hypothesized that the adoption of an innovation is related positively to the profitability of employing the innovation and negatively to the expected investments associated with its introduction. Mansfield substantiated the theoretical implications of his model by analyzing the diffusion of twelve industrial innovations in four major industries.

Mansfield's findings were further extended by Fisher and Pry (1970), who considered only fractional shares of a market controlled by two competing technologies. They postulated, on the basis of their analysis of many substitution processes, that the rate of fractional adoption of a new technology is proportional to both the fraction of the market penetrated by the new technology and the fraction of the market held by the still-used old technology. They also assumed that this substitution process proceeds to complete market takeover by the new technology once it has progressed as far as a few percent of the market. These two basic assumptions describe a growth process that can be represented by a two-parameter logistic function:

$$\frac{f(t)}{1-f(t)} = \exp(\alpha t+\beta)$$

where t is the independent variable usually representing some unit of time, α and β are constants, f is the fractional market share of the new competitor, and $1-f$ that of the old one. The parameters α and β are sufficient to describe the whole substitution process. They cannot be directly observed; they can, however, be estimated from historical data.

The characteristics of the logistic function describing fractional substitution are illustrated in Figure 1 which shows the seventeen substitution cases studied by Fisher and Pry and the logistic fit of these data as a smooth line. The plot on the right-hand side of the figure shows the S-shaped form of the function and the scatter of the observed data around the trend line. The plot on the left-hand side shows another convenient property of the logistic function for empirical analysis. When a sample of data points is plotted as the transformation $f/(1-f)$ on the logarithmic scale, the scatter of observed fractional shares (f) has a linear secular trend, provided that the data points can be described by the logistic function. In the plot on the left-hand side this is the case, the straight line being the logistic fit of the data. On both curves the time has been normalized at the inflection point ($t=-\alpha/\beta$) of the curve, where the fractional share equals one-half and where the slope of the curve is the steepest.

Two sets of examples are shown here in Figures 2.A and 2.B from the original papers of Fisher and Pry (Fisher and Pry, 1970; Pry, 1973). Figure 2.A shows the substitution of basic oxygen furnace for open-hearth and Bessemer steel production expressed by fractional market share (f) for Japan, the Federal Republic of Germany, the United States and the Union of Soviet Socialist Republics. The triangles and the circles on the middle line represent the Federal Republic of Germany and the United States, respectively. Figure 2.B shows technological substitution, also expressed by fractional market shares (f), in the production of steel, turpentine, and paint. These two sets of examples illustrate that the logistic functions appear to give an excellent description of substitution, not only for very different products and technologies, but also for different types of economies.

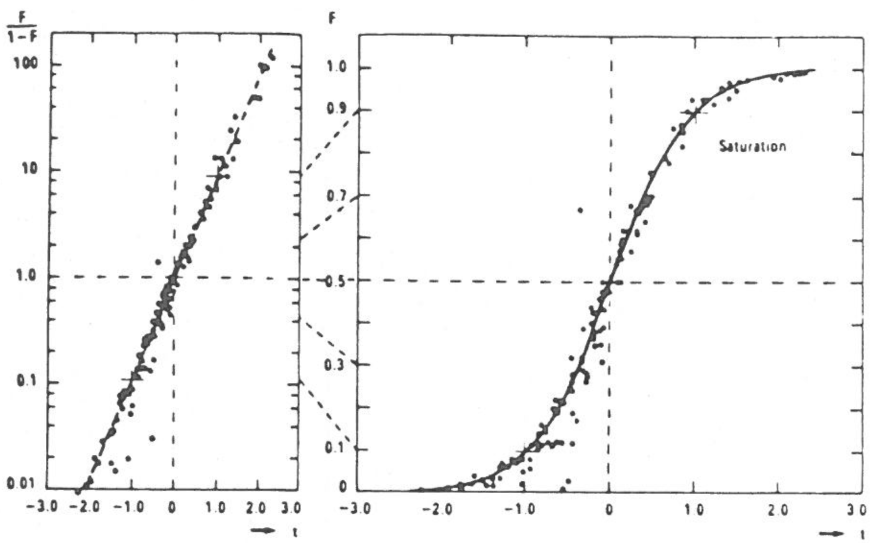

Figure 1 Growth to Limits: Seventeen Substitution Cases, in fractional market share and linear transformation (adapted from Fisher and Pry, 1970).

These examples from the work of Fisher and Pry show the regularity of technological substitution processes. An obvious next step is to investigate whether the regularity of the substitution dynamics can be applied to more broader technological, economic and social changes. Substitution of various forms of energy offers a good example in this context because energy represents one of the most important inputs for overall economic activity, and because it is used universally in everyday life, it also indicates changes in the acceptance of new energy systems. Energy substitution is representative of large changes at the whole infrastructure level, and thus goes beyond single technological improvements. A new energy source requires changes in accepted practice that range from exploiting new resources of energy, converting them to more convenient energy forms such as electricity and fuels, and transporting and distributing them to the point of end-use. At the end-use level new devices and practices are also needed to convert the delivered energy into the services that are required in various sectors of the economy such as motive power, lighting, space conditioning and so on. Substitution of one form of energy for another implies structural changes that go beyond the changes needed within the energy system because it affects and is in turn affected by the overall economic activities and the whole social structure.

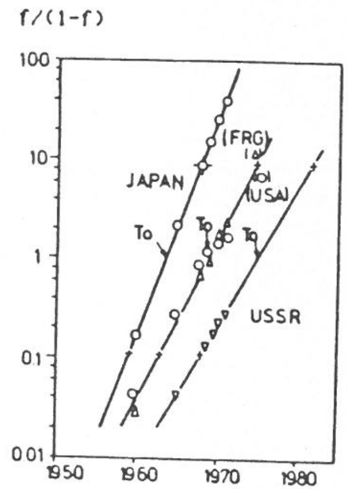

Figure 2.A Substitution in Steel Production, logistic transformation of fractional shares (Pry, 1973).

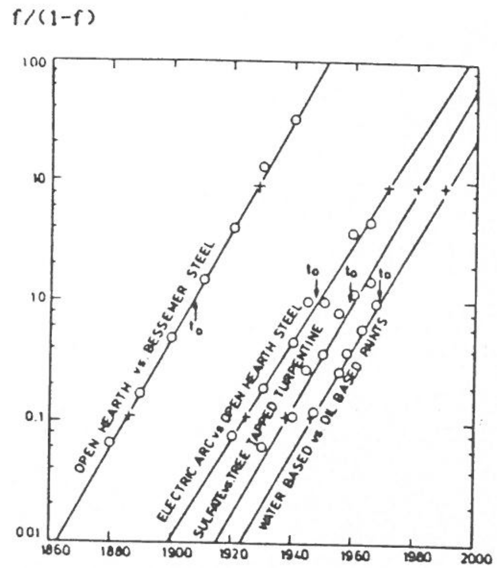

Figure 2.B Steel, Turpentine and Paint Substitution, logistic transformation of fractional shares (Fisher and Pry, 1970).

2.2. Global Energy Consumption

During millenia, dating back to the dawn of human civilization, the main sources of energy were fuelwood, wastes (dung and agricultural wastes), and human and animal muscle power. Major exceptions constitute the use of sailing vessels, river flotation, hydro and wind mills. In a wider sense, solar energy was also used, although not directly, but in transformed forms such as dried food and other artifacts. When compared with contemporary energy consumption all of these traditional energy forms were used at low absolute levels of exploitation and low densities of generation and end-use with hardly any need for transportation or transformation. This practice did not only prevail over thousands of years, but it was also basically similar even in different cultural settings. Major variations in energy use were spatially governed. Different geographic and climatic environments imposed different energy use patterns.

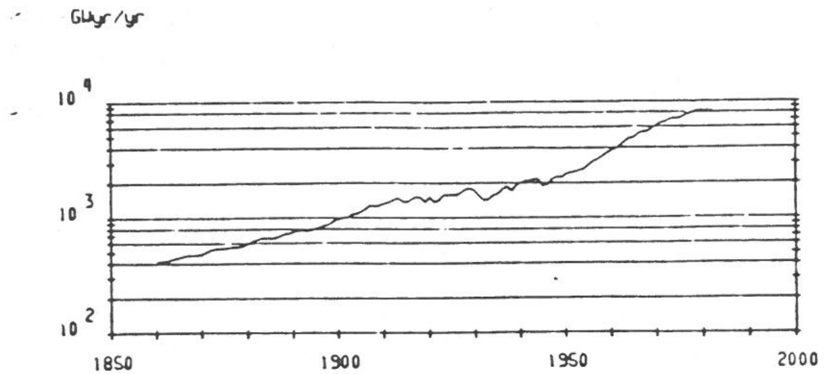

Figure 3 Total Primary Energy Consumption, World, in GWyr/yr.

With the emergence of manufacturing, industrial production, and the recent rapid social and technological changes over the past few centuries, the energy use patterns have also been altering and improving, energy generation and use densities have increased, and total energy consumption has been growing. Figure 3 shows that during the previous 125 years, global primary energy consumption (including the use of fuelwood) increased exponentially at an average growth rate of 2.3 percent per year. Yet during this period energy consumption did not draw equally from all sources, nor did the use of all energy sources increase equally.

Figure 4 shows global primary energy consumption since 1860, according to the five major primary energy sources - fuelwood, coal, crude oil, natural gas and nuclear energy. It is evident from the plot that the initial growth in the use of every energy source except fuelwood is exponential (i.e., the secular trend is linear on the logarithmic scale), and that many features, apparently related to economic and political events, influenced energy consumption. During the time when coal held the largest share of the market its consumption was subject to great fluctuations that coincide with the two world wars and the intervening period of worldwide economic depression. The consumption of fuelwood, once the most important source of energy, has decreased since the beginning of the century, although its use is still widespread, especially in the developing parts of the world. In 1965, oil surpassed coal, and natural gas will probably close up in a few years. In fact, oil and natural gas curves have the same shape and almost identical growth rates, they are just shifted in time by about ten to fifteen years. Nuclear energy is still in its early phase

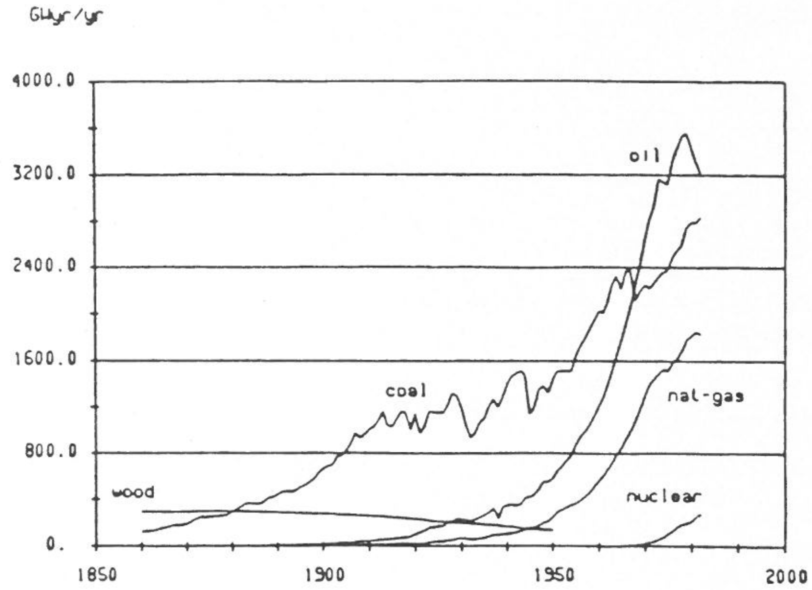

Figure 4 Major Primary Energy Sources, World, in GWyr/yr.

of development, therefore the steep growth prevailing over the last decade may not be indicative of its future role.

Other forms of primary energy have only marginal shares of the total energy consumption today and are not included. Even hydropower never exceeded the two percent share it reached in 1955. The importance of hydropower would slightly increase to about a six-percent share in the same year provided that its contribution is converted to primary energy equivalent in terms of the fossil energy needed to generate the same amount of electricity. Dung is not included because of the lack of reliable consumption estimates, although its share in primary energy consumption was considerable in the past. Putnam (1953) suggested that dung, as a fuel source, has had a fairly constant share of 16 percent since the 1860s. This is only a rough estimate, the actual consumption of dung and other agricultural wastes was probably much higher during the last century and may be even higher than ten percent of all commercial energy today. Even the fuelwood consumption time series cast some doubt on their accuracy due to the lack of any fluctuations present in the use of all other primary energy sources.

Global primary energy consumption, according to the five most important sources of energy given in Figure 5, indicated the dynamic evolution of the world primary energy system. It is evident that the older forms of energy have been substituted by the newer ones. These dynamic changes are more clearly seen in Figure 5 which shows the fractional shares of the total primary energy market taken by these five energy sources. In terms of fractional market shares, fuelwood was already substituted by coal during the last half of the nineteenth century. In 1860, fuelwood supplied about 70 percent of consumed energy, but by the 1900s its share had dwindled to little more than 20 percent. Due to the insignificant use of crude oil and natural gas during the last century, most of

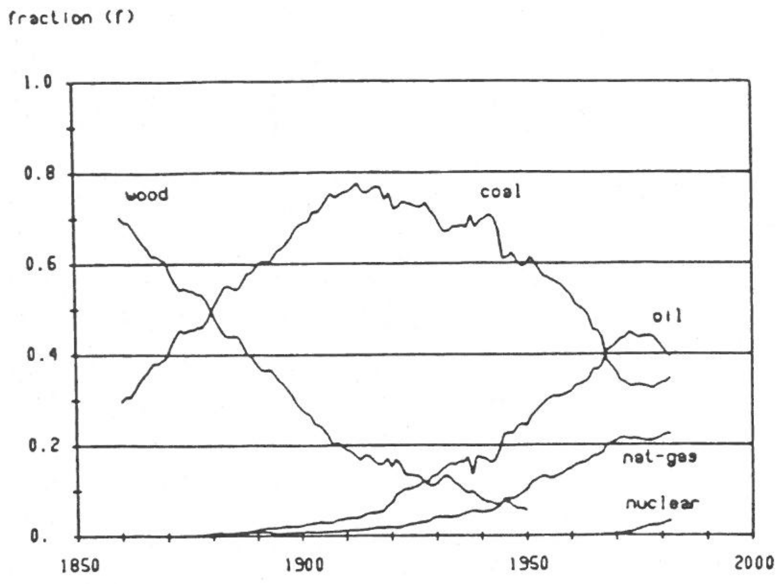

Figure 5 Fractional Shares of Major Primary Energy Sources, World.

the market losses incurred by fuelwood were taken by coal. Consequently, coal's share in total energy supply increased from 30 percent in 1860 to almost 80 percent by the 1900s. By 1910 the rapid increase in coal use ceased, and during the 1920s a phase of decline set in. This decline in the relative share of coal use resembles the market losses of fuelwood fifty years earlier. The replication of this pattern is almost symmetrical because after the 1920s both fuelwood and coal are substituted by still newer sources of energy - crude oil and natural gas.

2.3. Global Energy Substitution

Substitution of the old ways of meeting energy demand by the new ones resembles similar substitution processes on the level of single technologies, such as various steel, paint or turpentine production methods. Consequently, we made a heuristic assumption that energy systems, like other goods and products, can be viewed as technologies competing for a market and that they should behave accordingly (see Marchetti and Nakićenović, 1979).

The evolution in primary energy use, seen as a technological substitution process, is shown in Figure 6 on a logarithmic plot of the fractional market shares of the five primary energy sources. The fractional shares (f) are not plotted directly but as the linear transformation of the logistic curve, i.e., $f/(1-f)$ - the ratio of the controlled over the uncontrolled market share on logarithmic scale. This is the same form of showing a substitution process as that used by Fisher and Pry in their analysis of technological processes. This form of presentation reveals the logistic substitution path as an almost linear secular trend with small annual perturbations. Thus, the presence of some linear

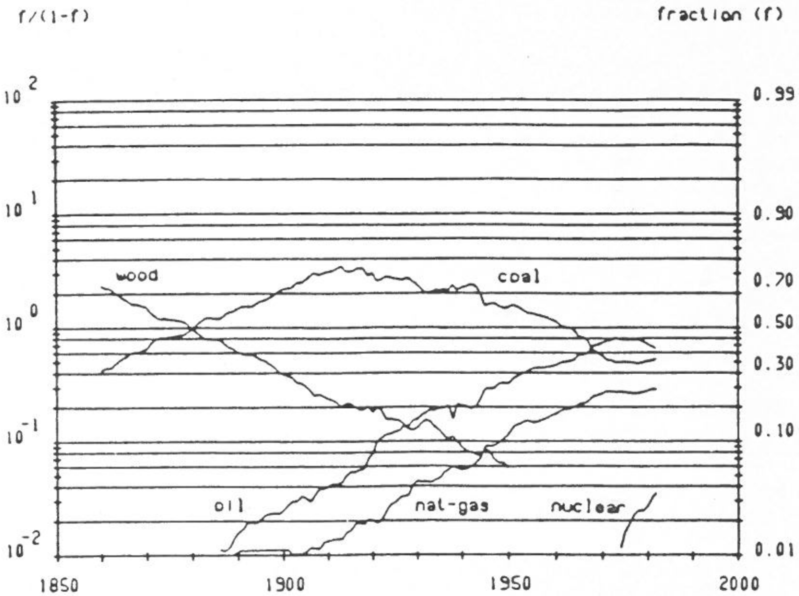

Figure 6 Primary Energy Substitution, World, logistic transformation of fractional shares.

trends in Figure 6 indicates where the fractional substitution of energy sources follow a logistic curve. It also indicates that energy substitution is similar, at least in this respect, to the technological substitution patterns observed by Fisher and Pry. Figure 6 lends strong support to our approach of treating energy sources with the logistic substitution method. However, Figure 6 also shows that in the case of multiple competition, logistic substitution is not preserved in all phases of the process. This point is well illustrated by the substitution path of coal, which curves through a maximum from increase to decline in its market shares.

As mentioned above, when dealing with more than two competing technologies, we had to generalize the Fisher and Pry model since the logistic paths are not preserved in all phases of the substitution process. As was illustrated in the substitution path of coal, every technology undergoes three distinct substitution phases: growth, saturation, and decline. The growth phase is similar to the logistic substitution by a new for an old technology in the Fisher and Pry model, but it usually terminates before full takeover of the market is reached. This is followed by the saturation phase which encompasses the slowing of growth and the beginning of decline. During the saturation phase, the secular trend is not logistic. After the saturation phase, the market shares proceed to decline logistically.

In order to describe this more complex substitution process of more than two competing technologies, we have assumed that only one technology is in the saturation phase at any given time, that declining technologies fade away steadily at logistic rates uninfluenced by competition from new technologies, and that new technologies, after entering the market grow at logistic rates. The technology in the saturation phase is left with the residual market share and is forced to follow a nonlogistic path that joins its period of growth to its subsequent period of decline. In other words, the market share of the oldest still-growing technology - for example, coal - can be defined as a complement to one of the sums of the other fractional shares following logistic substitution paths. Thus, during the saturation phase the oldest still-growing technology takes the residual of the market, but eventually its share starts to decline and becomes logistic again. Subsequently, the next newer technology - oil in this example - would undergo the transition from growth to decline. In effect, the technologies that have already entered their period of market phase-out are not influenced by the introduction of new ones. Deadly competition exists between the saturating technology and all other, newer technologies.

This short description of the generalization of the Fisher and Pry model illustrates the most important features of multiple competition that are captured by the logistic substitution model. A more formal presentation of the model is given in Nakićenović, 1979 which translates our assumptions into mathematical language.

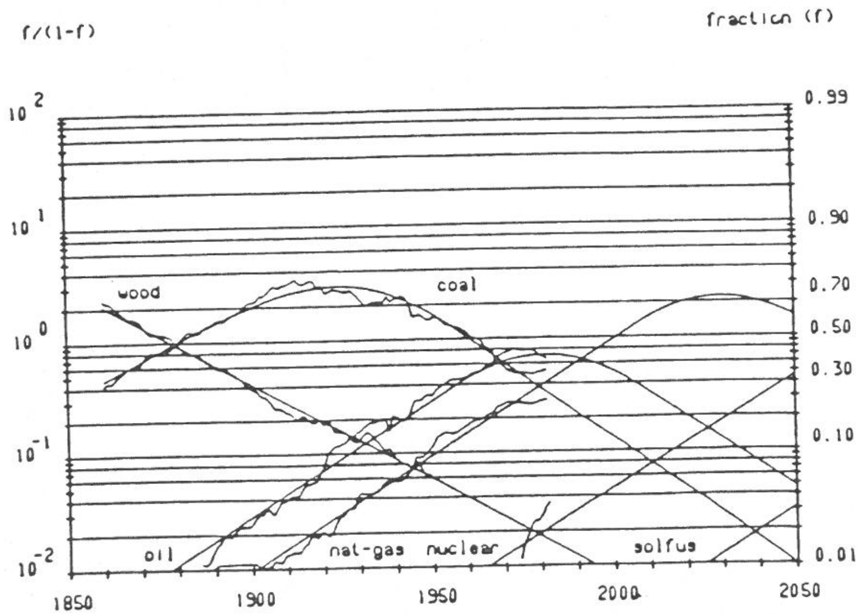

Figure 7 Primary Energy Substitution, World, logistic transformation of fractional shares with projections.

Figure 7 shows the application of the model to our example of global primary energy substitution. The fractional market shares of the five primary energy sources are plotted on the logarithmic scale as the linear transformation of the logistic function introduced above. The smooth secular trends are the model estimates of the historical data; the straight lines show where energy sources follow logistic substitution paths. The result is encouraging, since it indicates that this kind of model is capable of reproducing the evolution of energy substitution with very high accuracy. The scatter of the historical data is in good agreement with the smooth trends generated by the model. Significant deviations from the estimated trend can be observed only during relatively short time periods.

The model estimates of the substitution process are extended beyond the historical period up to the year 2050. For such an explorative "look" into the future, additional assumptions are required because potential new competitors such as nuclear and solar energy have not captured sufficient market shares in the past to allow estimation of their penetration rates. Nuclear energy controls slightly more than a two-percent share in all primary energy supply and is therefore visible from the historical data on the graph. Solar energy contributes much less than one percent so it is not visible on the graph. Thus, the starting point for market penetration of nuclear energy can be determined from the historical data and dated back to the early 1970s when nuclear power acquired a commercially significant scale with slightly less than one percent share in global primary energy consumption, but its future market penetration rate cannot be determined from the historical data. In order to explore and test the behavior of the logistic substitution model when the competition between the primary energy sources is extended into the future, we made explicit assumptions about the penetration rate of nuclear energy and possible entry of solar or some other advanced source of energy into the market at the one-percent level, as well as its future penetration rate.

In 1986, the nuclear energy share in total primary energy consumption was about 4.9 percent (B.P., 1987). In assessing the future nuclear energy development, we have chosen a more modest nuclear shares for the early 1990s to account for possible delays in the construction of the planned power plants so that our nuclear scenario prescribes a one-percent share in 1965 and a three-percent share twenty-five years later. For the next energy source, which we symbolically call "solfus" in order to indicate the potential use of both solar and fusion energy, we have made an equivalent scenario with a one-percent share in the year 2025 and three-percent in 2050. In other words, we have assumed that solfus energy will reach a level of commercial utilization by the year 2025 that is comparable to that of nuclear energy in the 1970s.

Figure 7 shows these two scenarios and the resulting dynamics of energy substitution throughout the first half of the next century. Note that the penetration trends have almost the same slope for all energy sources including our two scenarios. The extreme regularity and slowness of the substitution process is also conserved in our exploratory use of the model to project the future competition between primary energy sources. We have assumed in the two scenarios that about twenty-five years would be necessary before the new energy sources, nuclear and solfus power, could increase their market shares from one to three percent. Similar penetration rates can be observed for coal, oil and natural gas during the last 120 years. It takes about one hundred years to go from one to fifty percent of the market share, or about fifty years from a ten to fifty percent share. We call this length of time the time constant of the system.

The regularity of the substitution process refers not only to the fact that the penetration rate of various energy sources remains constant over periods of about one hundred years for a given energy source and is almost the same for all energy sources, but also to the fact that most perturbations are reabsorbed after a few years without affecting the long-term trend. Only during the initial phases of market penetration, do the market shares not immediately stabilize to long-term substitution trends. Oil penetrated somewhat faster until it reached a two-percent market share, while natural gas controlled almost a constant one percent of the market for over a decade before both energy sources stabilized to their long-term substitution trends. Our scenario of the future penetration rates for nuclear energy indicates a similar departure of the long-term behavior from the relatively rapid growth before the 1970s.

It is also interesting to note that the saturation levels achieved by all energy sources are much lower than the full market takeover. The introduction of new energy sources and the long time constant lead to maximum market penetrations of between 50 and 70 percent. New energy sources are introduced before the most dominant ones have even reached a 50 percent share. For example, crude oil and natural gas were introduced during the 1880s whereas coal, then the dominant source of energy, reached its maximum market share 30 years later. Figure 7 shows that coal reached its maximum market share in 1925, oil in the 1980s and that natural gas is projected to achieve its maximal market penetration in the year 2030. This indicates that the maxima are roughly spaced at intervals of about 50 years, which corresponds to the time constant of about 50 years for market share increases from 10 to 50 percent. Thus, the dynamics of energy substitution during the last 125 years and over the projected 70 years in the future indicate that the global energy system behaves as though it had a schedule that governs its evolution. In order to further explore the regularity of this process we have used the logistic substitution model to test the robustness of this behavior with respect to different assumptions about the availability and quality of historical data and with respect to different scenarios for the competition of new energy sources.

2.4. Energy Substitution in the United States

The records of energy use and production in the United States are very sparse for the Colonial Times. Historical statistics for the period after the Revolution, however, are fortunately almost complete, so that it is possible to reconstruct annual time series for the consumption major primary energy sources with a high degree of precision. This is particularly interesting because it offers the possibility to analyze the use of older, now practically extinct forms of energy. Even during the early nineteenth century the United States remained basically an agrarian and rural society. In 1870, farming still contributed almost forty percent of the gross domestic product of the United States (U.S. Department of Commerce, 1970), so that local energy sources constituted an important component of the overall energy supply.

The consumption of mineral energy and fuelwood can be traced back to 1800 when timber certainly represented the most important source of energy. Wood was the principal fuel for domestic as well as industrial purposes. Estimates of the fuelwood use during the last century indicate that it was used lavishly, guaranteeing adequate energy supply. In contrast to Europe and especially the United Kingdom, the United States had vast territories at their disposal, and timber production was usually not limited through resource availability, but rather by the logistics of harvesting and transport opportunities. In practice this means that there was ample supply for local uses. Larger cities in the East were supplied mostly through a growing network of turnpikes. However, water and wind power and work animals also provided substantial inputs to the overall energy supply. As in the Medieval Europe, water and wind power supplied the greater part of inanimate mechanical energy. The rest was provided by animal muscle power and human labor. Especially horses and mules represented a very important source of the motive power needed in agriculture and transport. Even as late as 1920, work animals provided larger aggregate horsepower in farming than tractors and all other agricultural machinery (more than 22 million compared to 21.5 million horsepower, see U.S. Department of Commerce, 1970).

Figure 8 shows the primary energy inputs for the United States from 1850 up to the present in five-year intervals. We have included all major energy sources including fuelwood, the energy content of work-animal feed, mechanical wind and water power, coal, crude oil, natural gas, hydropower and nuclear energy. In 1850, fuelwood, animal feed and wind and water mills supplied ninety percent of all primary energy, and coal the other ten percent. Thus, the transition from traditional to commercial energy use was initiated in the United States during the middle of the last century, marking the beginning of the industrial age. In absolute terms, the use of all traditional sources of energy has been declining since 1850 with the exception of slight increases in fuelwood consumption up to the 1870s. This decline continued so that today basically all primary energy use originates from commercial sources and the marginal use of traditional energy forms disappears in the "noise" of the statistical records.

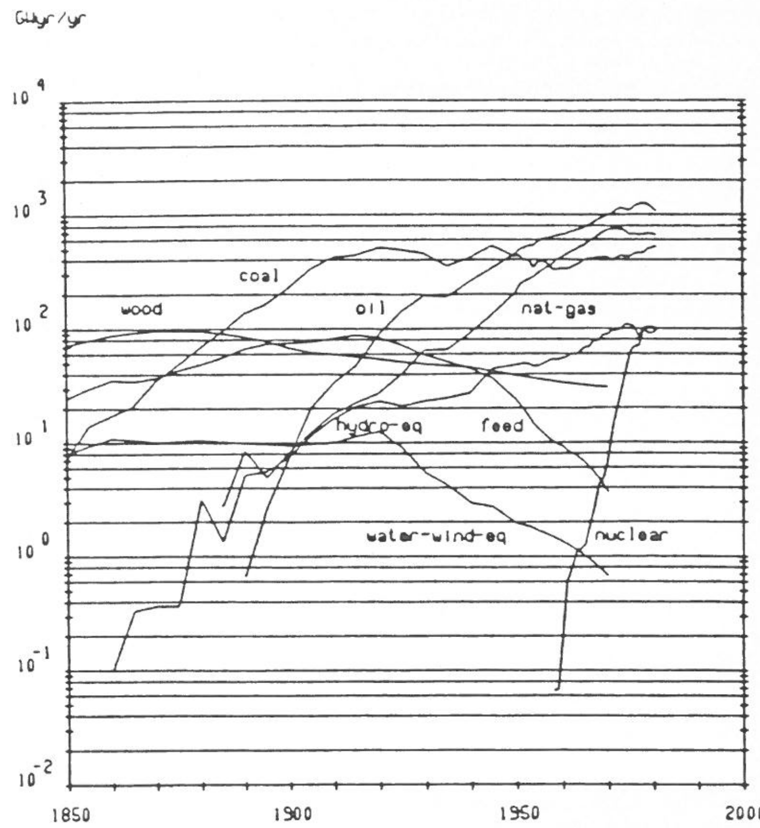

Figure 8 Primary Energy (Equivalent) Consumption, USA, in GWyr/yr.

During the second half of the last century, coal not only accounted for all of the increases in total energy use, but it also substituted for the losses incurred by traditional energy forms. This explains the very rapid growth of coal during these fifty years with an average annual growth rate of 6.7 percent. During 1870, both crude oil and natural gas entered the energy market, but the enormous expansion of energy use was primarily due to coal until the end of the century. Between 1900 and 1950 the consumption of crude oil and natural gas increased by an average annual growth rate of 8.5 and 6.6 percent, respectively. Overall energy consumption increased from about 100 GWyr in 1850 to over 2400 GWyr in 1981 which corresponds to an average annual growth rate of 2.4 percent.

Among all traditional energy sources, fuelwood was by far the most important suppling almost seventy percent of all primary energy inputs in 1850. Unfortunately, data on the use of all traditional energy forms during the last century must be taken to represent orders of magnitude rather than precise quantities since they are all based on fragmentary information. Reynolds and Pierson (1942) based their estimates of fuelwood consumption on the population size and distribution, climate, housing conditions, and the availability of wood in the various regions of the United States. Historical records on actual wood use are sparse. As an explanation of this fact, Reynolds and Pierson remarked that there was probably no need to "write about firewood, or even record statistics about it" since "cordwood was about as plentiful as air" in the United States and "nobody wrote about air" use either. Fisher (1974) estimated the feed energy content of work animals by multiplying the number of farm and nonfarm horses, mules and oxen in use with the annual average energy content of the feed required by them. This calculation is based on an annual energy requirement of 3 kWyr per animal, derived from an average animal weight of about 750 kilograms, an average daily consumption of 1 kilogram of feed per 50 kilograms of animal weight, and an average energy content of about 0.5 Wyr per kilogram of feed. Computed in this way, the energy content of work animal feed was the second most important traditional energy source in the United States. It is interesting to note here that although oxen constituted almost 30 percent of all work animals in 1850, horses and mules had displaced them all by the 1900s. Thus, we have here another example of technological change during the early period of American economic development.

Wind and water mills and sailing vessels constitute the last traditional energy source accounted for in our statistical records. An appropriate treatment of this traditional energy source is the most intricate of all energy forms. The difficulty arises from the fact that it can be accounted for either in terms of the direct mechanical energy (inputs) provided, or in terms of equivalent amounts of other energy forms required to produce the same mechanical work. During the last century only animal and human work could have been substituted for the mechanical power provided by wind and water flow. These two possible accounting methods are equivalent to the two alternatives of calculating hydropower consumption either in terms of energy inputs (i.e., amount of electricity generated, sometimes called primary electricity) or in terms of fossil energy requirements to produce the same amount of electricity. In Figure 8 both mechanical water and wind power and hydropower are given in terms of energy equivalents. Water and wind consumption are calculated in terms of the feed equivalent energy that would have been required by work animals to generate the same mechanical energy. Hydropower is given in terms of the fossil energy needed to generate the equivalent amount of electricity at the prevailing average efficiency of power plants in corresponding years. This calculation method has the disadvantage of overstating the importance of these two energy forms, especially the contribution of wind and water power. The average efficiency of work animals in converting the energy content of feed into mechanical work is very low and does not exceed four percent. This means that the mechanical energy of wind and water power has to be multiplied by at least a factor of 25 to obtain the feed equivalent energy. In addition, it is simply unrealistic to calculate total energy inputs to the American economy by implicitly assuming that all wind and water mills could have been substituted by horses and mules had it been required. To replace wind and water power in 1860, for example, would imply the feasibility of increasing the 12 million work animals by 30 percent.

Furthermore, with the onset of the coal age wind and water power were not replaced by work animals but rather by steam. Figure 9 shows the primary energy inputs of the United States with the difference that the direct mechanical energy of wind, water and hydropower is given in terms of electricity inputs. This does not affect the overall pattern of energy use, since both of these energy sources provided marginal energy inputs that even in Figure 9 never exceeded more than a ten-percent share.

In the case of the United States we have the unique opportunity to analyze the substitution of traditional energy forms by commercial energy sources as done in Figure 10. This substitution process can almost be considered as the "proxy" indicator for the pace of industrialization and the economic structural change from agricultural to industrial production. The application of the logistic model to describe this process is shown together with the historical shares of the two broad classes of energy sources. We have grouped fuelwood, animal work, and wind and water mills under the umbrella of traditional energy sources. They all represent renewable energy forms that were basically only

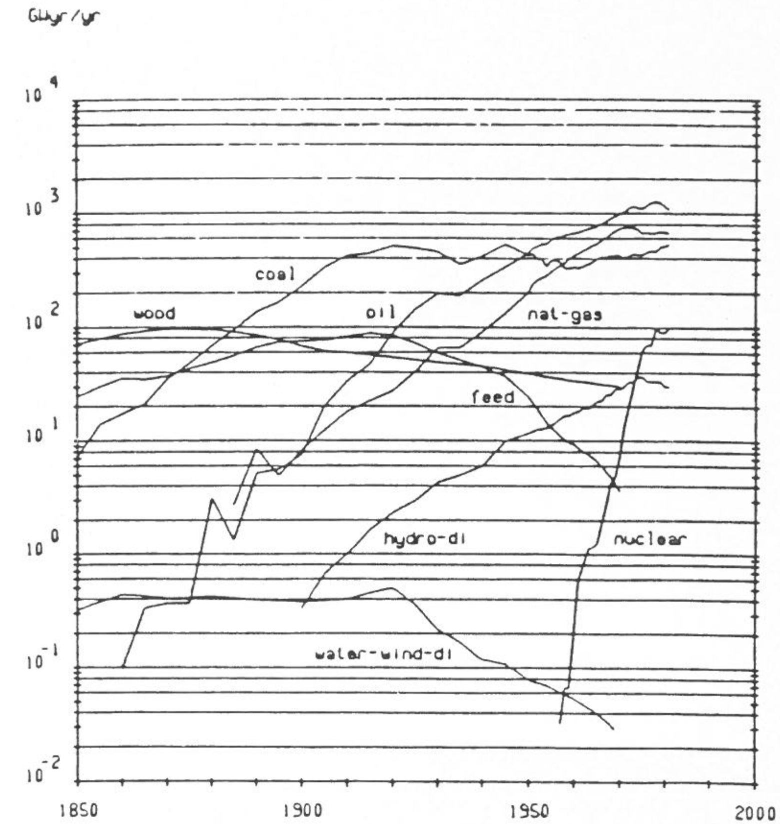

Figure 9 Primary Energy Consumption, USA, in GWyr/yr.

suitable for local use by a rural society with small concentrations of industrial production in a few urban areas. The traditional energy forms are shown in competition with commercial energy forms including coal, crude oil and natural gas. Until the 1900s almost all fossil energy consumption was based on coal use.

Figure 10 shows the two classes of energy use in terms of their respective fractional market shares (f) of total primary energy use as the linear transform of the logistic function, $f/(1-f)$. The substitution process is remarkably regular over the whole time period of over 130 years. It is interesting to note that the fifty-percent mark in the substitution of commercial for traditional energy was reached shortly before the turn of the century. If we extrapolate this energy substitution process back into the past, the emergence of coal (commercial energy) dates back to the 1820s. In fact, we will show later that this is a very accurate estimate, indicating once more the remarkable regularity of the substitution process. The time constant is quite long - more than 80 years were required before commercial energy sources could capture fifty percent of the market. The corollary of this observation is that traditional energy sources also sustained their decline for over 80 years

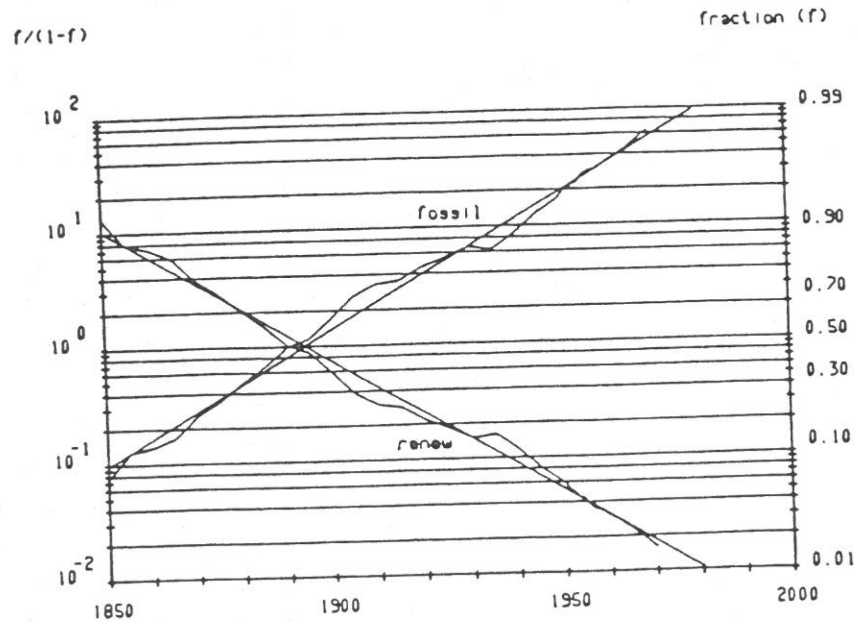

Figure 10 Commercial-Fossil for Traditional-Renewable Energy Substitution, USA, logistic transformation of fractional shares.

from the fifty-percent share to the one-percent mark in 1980.

Figures 11.A and B show this substitution process from the perspective of individual energy sources. The difference between the two examples is that Figure 11.A gives water and wind power in terms of feed energy equivalent, and hydropower in terms of fossil energy equivalent and that Figure 11.B gives the actual energy inputs of these two sources. Thus, Figure 11.A overemphasizes the role of these two energy sources. This does not affect the structure of the primary energy substitution process however, except for small shifts in the market shares of other energy sources. In both examples the logistic substitution model describes the substitution paths with remarkable accuracy. The departures of historical market shares from their long term paths last for over two decades only to return to the trend after the prolonged perturbation. This is the case with the market shares of coal and oil during the 1940s and 1950s, and fuelwood and animal feed during the 1860s and 1870s. It may also indicate a possible reabsorption of the departure of coal and natural gas market shares from their long term paths during the last ten years. The substitution process clearly indicates the dominance of coal as the major energy source between the 1870s and 1950s. The last phases of railroad expansion during the first 30 years of this period and the growth of steel, steam ships and many other sectors are also associated with and based on the technological opportunities offered by the mature coal economy. After the 1940s, oil assumed the dominant role in parallel with the maturity of the automobile, petrochemicals and many other modern industries.

The evolution of commercial energy use in the United States has a longer recorded history than the use of traditional energy sources. Figure 12 gives the annual consumption of all commercial energy sources and fuelwood starting in 1800. Here again we have

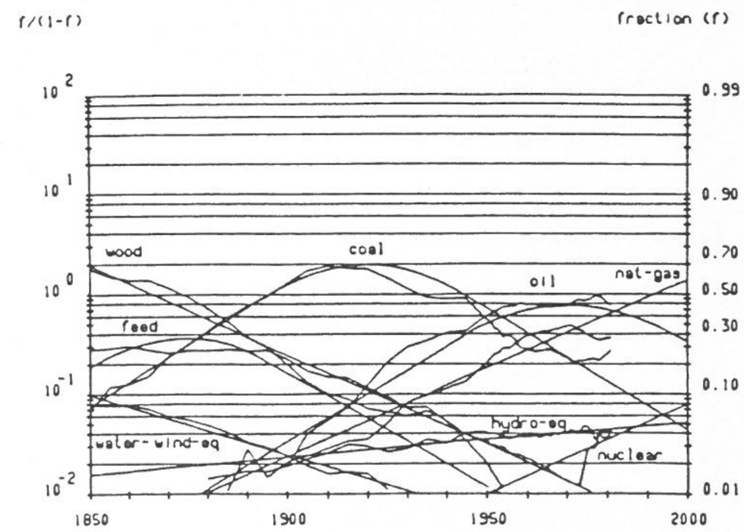

Figure 11.A Primary Energy (Equivalent) Substitution, USA, logistic transformation of fractional shares.

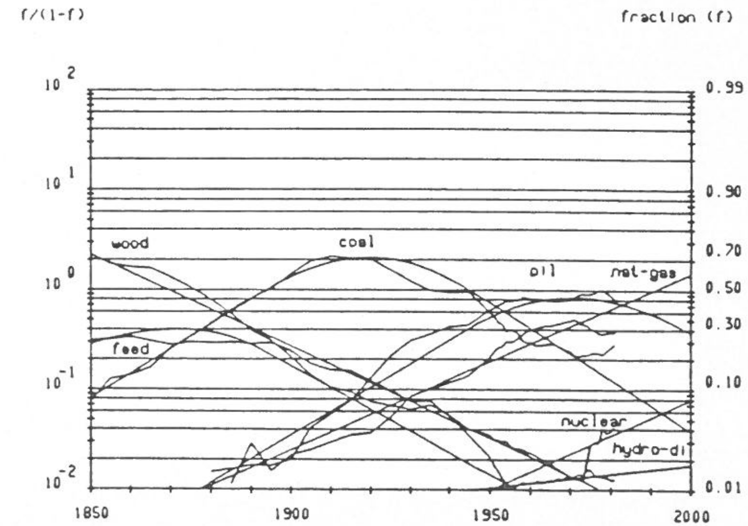

Figure 11.B Primary Energy Substitution, USA, logistic transformation of fractional shares.

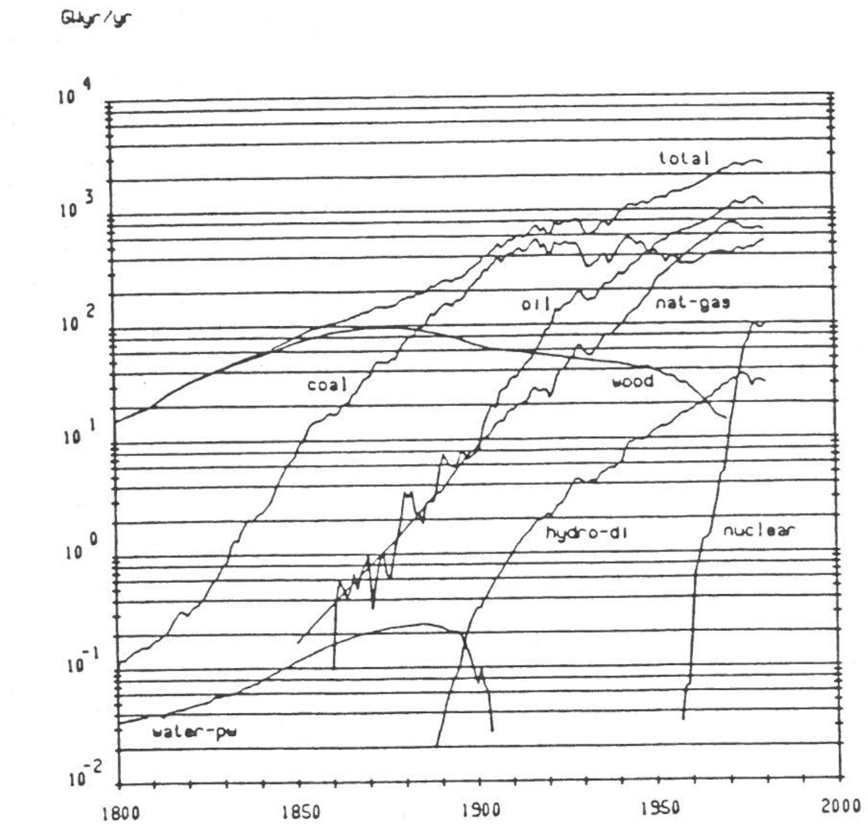

Figure 12 Primary Energy Consumption, USA, in GWyr/yr.

two possible representations of hydropower and they are both given in Figure 12, the direct energy inputs and the fossil energy equivalent, respectively. The fossil-equivalent of hydropower tends to overemphasize the actual contribution especially during the first few decades of the twentieth century because the prevailing efficiencies of coal to electricity conversion were very low at that time. For example, in 1920 the average efficiency of installed power plants did not exceed ten percent compared with over thirty percent in 1980. In any case, hydropower shares of the primary energy inputs were not very large, reaching slightly more than four percent in terms of fossil equivalent or little more than one percent in terms of direct energy inputs during the 1970s.

Figure 13 shows the substitution of the five most important commercial sources of energy and fuelwood. The logistic substitution model in this example also describes with high precision the evolution of primary energy consumption in the United States. Due to the dominance of fuelwood as the major source of energy during most of the last century, the information loss associated with the lack of adequate annual estimates for animate mechanical inputs to the energy system is not very large. Direct wind and water power

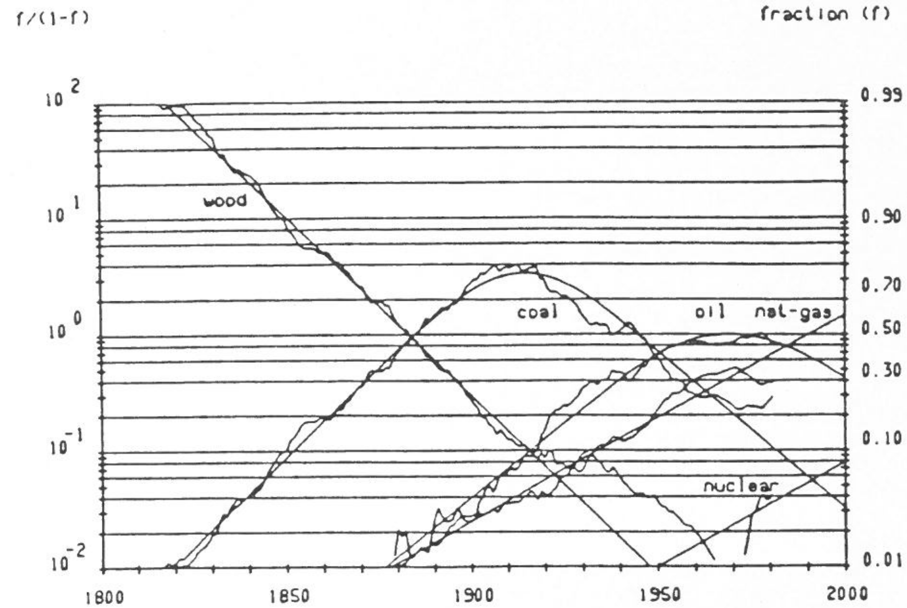

Figure 13 Primary Energy Substitution, USA, logistic transformation of fractional shares.

are included in the data set, but due to their low contribution to total energy supply, when expressed in terms of their actual energy input, they are not observable at the one-percent level. Thus, before the 1820s fuelwood provided for virtually all the energy needs of the United States. Coal entered the competition process in 1817, which corresponds almost exactly to our extrapolation based on the previous example of primary energy substitution. This example illustrates the senescence of fuelwood and the rise of coal very clearly. Up to the late 1880s it was essentially a two technology market - whatever gains coal made were translated into losses for fuelwood. The initially slow introduction of crude oil and natural gas during this decade translated into market dominance 80 years later. The dominance of fuelwood and coal show an interesting symmetry, each period of dominance lasting slightly over 60 years.

The prolonged use of fuelwood in the United States compared with Europe and especially Great Britain, shows that wood was an important source of heat and power for early industrial purposes well into the second half of the nineteenth century. In the United States the steam age already began in the economy based on wood use. The first steam-boats and locomotives were fired with wood, which remained the principal fuel used by railroads until about 1870 (Schurr and Netschert, 1960). The only other large use of wood was found in the iron industry. Around 1850, more than half of all the iron produced was still smelted with charcoal (see, Figure 13). Nevertheless, during this early period of industrialization, the United States was still basically a rural society, so that the total amount of fuelwood consumed in manufacturing and transportation was small compared to the huge quantities used in households. In 1880, the domestic use of fuelwood still accounted for more than 96 percent of fuelwood consumed (Schurr and Netschert, 1960). At the same time, however, coal already supplied almost one-half of all energy needs, most of it being used by emerging industries. In 1880, coal supplied almost ninety percent of the fuel used for smelting iron (see Figure 13). Thus, the end of the last century marks the beginning of the industrial development period in the United States.

In the United States the first use of crude oil and natural gas dates back to 1859. During the 1880s, both of these two energy sources reached the one-percent market share. From this point on the use of crude oil expanded somewhat faster as time progressed. In 1950, crude oil consumption surpassed that of coal, and natural gas use surpassed coal nine years later. It should be noted that as late as the 1920s the use of crude oil was not much larger than the consumption of fuelwood. It is remarkable that the structure of energy consumption changed more during the period of oil dominance when compared with earlier periods. The 1950s, when oil became the dominant source of energy, represent the beginning of more intense competition between various energy sources. All the way up to the 1950s, the energy source that dominated the energy supply at the time also contributed more than one-half of all primary energy consumption. After the 1950s in both countries each primary energy source contributed less than one half of primary energy. Crude oil was close to achieving a fifty percent share during the 1970s, but before actually surpassing this mark proceeded to decline. Thus, during the last three decades three important sources of energy shared the market with no single source having a pronounced dominance, which is contrary to the pattern observed during earlier periods.

The logistic substitution model indicates that it is possible to describe the broad features of the evolution of the energy system at the global level as well as in the United States over very long periods of time by rater simple mechanisms, in spite of so many turbulent and profound changes since the beginning of the industrial revolution. We also applied the model to describe energy substitution in some three hundred examples ranging from primary to final energy use, electricity generation, etc. (see, Marchetti and Nakićenović, 1979; Nakićenović, 1984). Thus, it is evident that the logistic substitution model is a powerful tool when applied to technological change within the energy system. The changes within the energy system are easier to record in terms of long time series since a natural common denominator is available for measuring the contribution of each important component of the system - the contributions of all energy sources can be measured in common energy units. Unfortunately, such a relatively simple and common unit is not available for describing the evolution of other systems. To describe changes in other sectors of the economy, only one obvious common unit is available - the monetary value of the various technologies and activities of the sector. This is not however an appropriate measurement unit for very long time periods, since the price system itself changes with the structural changes of the whole economy. The energy content of a ton of coal depends only on the quality of coal and it is independent of the time period when the coal is mined or used, but one Pound Sterling in 1700 represents a different monetary system than the same unit of value two hundred years later. If a commodity is essentially free because of its abundance (as was the case with fuelwood in the rural areas of the United States during the last century), then that commodity also has no real economic price in the same sense as the air that we breathe. This of course does not mean that a cord of wood had no value at the point of collection. Wood was actually critical for survival but since it was in abundance, its price was insignificant representing basically the cost of timer cutting and processing.

2.5. Efficiency of Energy Use

There are many ways of determining the efficiency of energy use. The most obvious indicators are the efficiencies of primary energy conversion to secondary and final energy forms. Another possibility is to estimate the efficiency of energy end-use. Examples include the amount of fuel needed for travel, or for space conditioning. All of these efficiences have improved radically since the beginning of the industrial revolution along

with the introduction of more efficient technologies. In some cases the improvements span almost an order of magnitude. For example, in 1920 the average efficiency of natural gas power plants in the United States was nine percent, whereas today the best combined cycle gas turbine power plants can operate with efficiences of over 50 percent. This improvement spans a period of about fifty years.

All of these efficiency improvements of individual technologies are translated into more effective use of energy and other materials at the level of the overall economic activity. Some efficiency increases result from improved technologies and others from substitution of the old by new technologies. The extent of these changes and improvements can be expressed at an aggregate level by the amount of primary energy consumed per unit of gross national product in a given year.

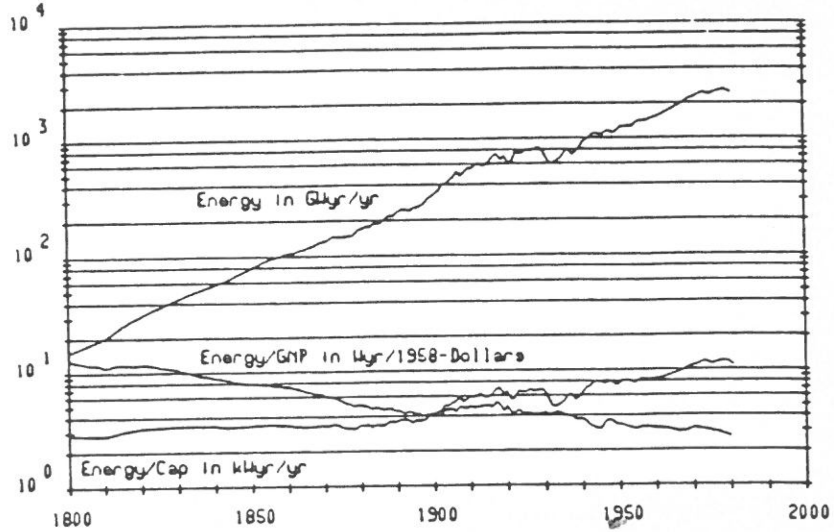

Figure 14 Primary Energy, Gross National Product and Energy Intensity, USA, in Wyr and constant 1958 Dollars.

Figure 14 shows the ratio of energy consumption over gross national product (energy intensity) for the United States, including non-commercial forms of energy. The average reduction in energy consumed to generate one dollar of gross national product decreased by about 0.9 percent per year during the last 180 years. The ratio decreased from more than ten kilowatt-years per dollar in 1800 to slightly more than two kilowatt-years per dollar in 1982. Thus, a regular decline in energy intensity of the whole economy prevailed over a long historical period indicating that energy conservation is a historical process that was discovered as a concept only during the last decade.

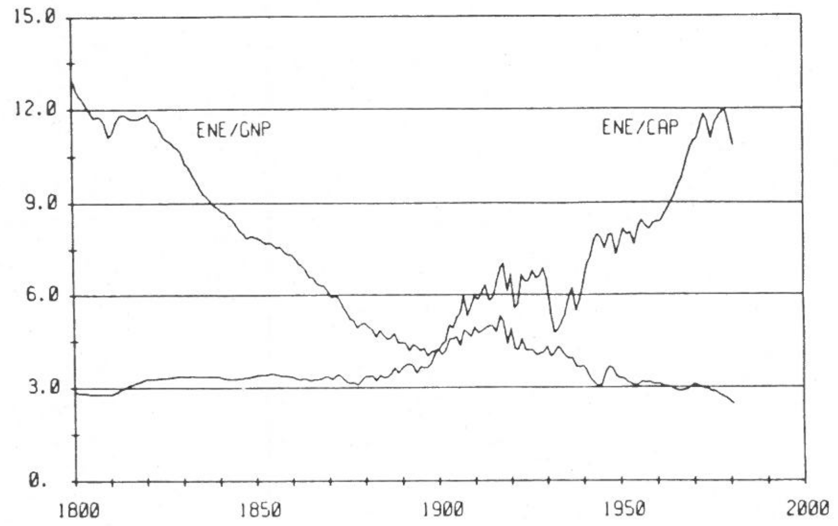

Figure 15 Energy per GNP and per Capita Consumption, USA, in kWyr per capita and Wyr per constant 1958 Dollar.

This decrease in the energy intensity fluctuated considerably around the decreasing secular trend of 0.9 percent per year. In fact, there are clearly visible periods when the amount of energy needed per unit value added increased, while in other periods the rate of decrease appears to have accelerated. Much of this variation is not clearly visible in Figures 14 due to the logarithmic "compression" of the data. Figure 15 shows the same data on a linear scale together with the increases in per capita primary energy consumption. It is interesting to note that the major periods without longer-term improvements in energy intensity occurred in the 1820s, 1900 to 1920 and most recently from about 1945 to the early 1970s. After the OPEC oil embargo another phase in improving energy efficiencies has been initiated.

During the last 180 years, the cumulative improvement in energy intensity was about a three-fold decrease from approximately 12 Wyr per constant 1958 Dollars) in the 1800s to less than 3 Wyr per 1958 Dollar during the late 1970s and early 1980s. This means one fourth of primary energy inputs are needed today to generate a Dollar of value added in the American economy compared with the 1800s. Figure 15 also shows that during the same period the per capita primary energy consumption also increased three-fold from less than 3 kWyr per capita (which is almost 50 percent higher than the current global average, indicating how abundant fuelwood was in the United States) to about 12 kWyr per capita today. Thus, effective energy services in terms of embodied aggregate input of energy into the economy have increased by a factor of six due equally to both reductions in energy intensiveness and increases in per capita energy consumption. In

other words, a half of the increase in effective energy availability is due to economic structural change and improvements in energy efficiences while the other half is due to increases in energy inputs in excess to population growth (about 2.1 percent per year since 1800 compared to 2.7 percent per year primary energy consumption growth). We will use this historical trend in the change of energy intensity in the American economy as a yardstick for assessing the improvements achieved since the so-called energy crisis of 1973 in a few selected OECD countries.

Table 1 Primary Energy Intensity of Selected OECD Countries.

Country	1973	1979	1983	1985	Change 1973-85
	(Megajoules per 1980 Dollar of GNP)				(percent)
Australia	21.6	23.0	22.1	20.3	-6
Canada	38.3	38.8	36.5	36.0	-6
Greece[1]	17.1	18.5	18.9	19.8	+16
Italy	18.5	17.1	15.3	14.9	-19
Japan	18.9	16.7	13.5	13.1	-31
Netherlands	19.8	18.9	15.8	16.2	-18
Turkey	28.4	24.2	25.7	25.2	-11
United Kingdom	19.8	18.0	15.8	15.8	-20
United States	35.6	32.9	28.8	27.5	-23
West Germany	17.1	16.2	14.0	14.0	-18

1 Increase is a result of a move toward energy-intensive industries such as metal processing.

Table 1 shows that the highest improvement in reducing the energy intensity (of GNP expressed in constant 1980 Dollars) since 1973 occurred in Japan, followed by the United States and the United Kingdom. Comparable improvements were also realized in Italy, the Netherlands and FRG. Australia and Canada as countries that are rich in domestic energy sources achieved much lower reductions. In contrast, the energy intensity of the Greek economy increased by 16 percent during this period, primarily due to the development of energy intensive industries such as metal processing. A similar trend can be observed in Turkey since 1979, although it is not as pronounced. This last example, illustrates that all savings in energy were not achieved through rationalization and introduction of more efficient technologies and organization, but also by decreasing the energy-intensive economic activities. Most of the countries with relatively high energy savings reduced the output of basic industries with a simultaneous shift from manufacturing to service sector. This aspect of structural change in the economy had a fundamental impact on reducing specific energy needs during the last 15 years.

It is interesting to note that an average decrease of 0.9 percent per year observed in the United States over the last 180 years corresponds to a change of -11 percent over the period from 1973 to 1985, the reduction in energy intensity observed for Turkey. Thus, the energy intensity decreased twice as fast since 1973 in the United States than its long-term historical rate and in Japan almost three times as fast. While these are very impressive achievements, they are not as large as could be expected from many programs and investments directed toward using energy more effectively. The increases in the international competitiveness would also suggest additional attempts to reduce the specific energy requirements. Apparently this was not the case, or at least the effect was not as large as could be expected on the basis of the long-term historical trends.

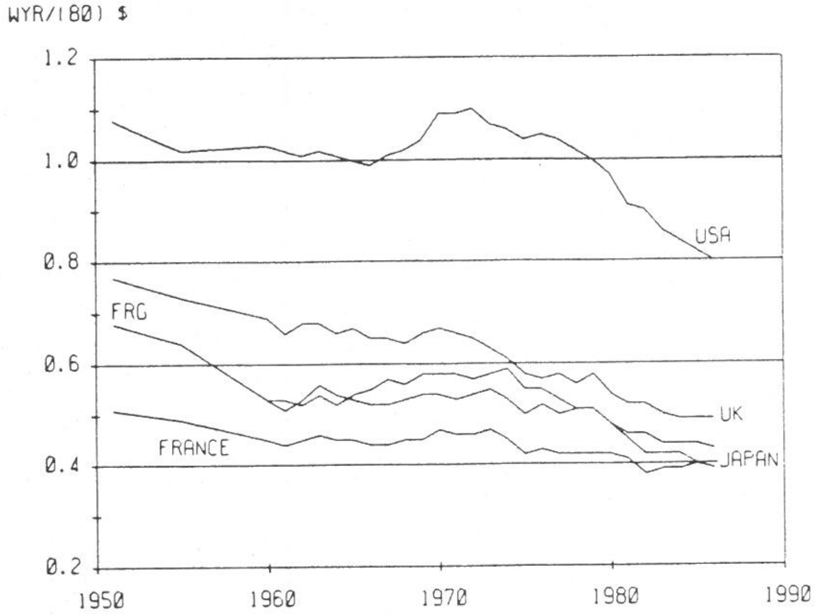

Figure 16 Primary Energy Intensity in Selected OECD Countries.

Figure 16 shows the improvements of primary energy intensity of Gross National Product (GDP)* expressed in constant 1980 Dollars for the United States, the United Kingdom, FRG, Japan and France. It illustrates the developments since 1950 in contrast to the reduction achieved since the oil shock of 1973 given in Table 1. It clearly shows that the United States not only consume much more energy per unit value added than the other leading OECD countries, but also that during the 1960s there was a substantial increase in the energy intensity in the United States that is hardly perceptible in the development trajectories of other four countries. The average improvement in the energy intensity for the United States since 1950 is thus reduced exactly to the long-term average of 0.9 percent per year despite the large improvements since 1973! Apparently the rapid economic growth and expansion of the 1950s and 1960s has led to retardation in improvement of the energy efficiencies at the level of the whole economy, and the accelerated energy savings and efficiency programs launched after 1973 can be seen as a correction and a reversal that brought the trajectory back to the historical average. While the other four countries use energy much more efficiently than the United States, the relative

* In contrast to the Gross National Product, the Gross Domestic Product does not include the factor transfers form abroad.

improvements since the 1950s have been on the order of 1.3 percent in the United Kingdom and FRG, about 1.1 percent (since 1960) in Japan and about 0.6 percent in France. It should be also noted that the evolution of energy intensity in the other four countries appears to be convergent within a rather narrow range of between 0.4 and 0.5 Wyr per (constant 1980) Dollar, while the United States with the current level of about 0.8 Wyr per Dollar is not only twice a high but is also higher than the intensity the United Kingdom had in 1950.

3. Transportation Systems

The evolutionary character of the long-term structural changes in the primary energy system discussed in the previous section will be corroborated by a similar analysis of the transportation sector. In absence of any data at the global level, we will analyze below the evolution of transport systems in the USA and the USSR to consider whether the model of structural technological change introduced above can also be applied in a different sector and in a comparative framework between countries with different market clearing mechanism. In a first step we consider the evolution of the transport infrastructure in terms of the respective length of individual infrastructure grids: canals for inland navigation, railways, surfaced roads and finally also airways. In a second step, we use another indicator for the relative importance of different transport systems: the intensity of usage expressed in terms on passenger-km performed by different transport systems. Finally, we will analyze also the long-term relationship between economic activity and transportation requirements in considering the passenger- and ton-km transported per constant unit of GDP in the case of France since 1800. More details on the analysis summarized in the following section may be found in Nakićenović, 1988 and in Grübler, 1987 and 1988.

3.1. Infrastructure Substitution

Figure 17 summarizes the historical evolution of individual transport infrastructures of the USA, expressed in miles of the respective transport infrastructures in operation. By analogy, we have complemented this quantification of the evolution of *physical* infrastructures also with the length of air transport infrastructure using the length of the federal (i.e. common carrier) airways network.

Since 1800 the length of transport infrastructures increased over more than five orders of magnitude. In analyzing the growth of individual transport infrastructures (Figure 18) we conclude, that each individual infrastructure proceeds through a characteristic life cycle pattern, with slow initial expansion, followed by rapid growth and finally onsetting saturation and eventual decline.

The saturation of the expansion of canals occurred in the 1860s, that of railways in the mid 1920s and the expansion of road infrastructures appears presently to approach saturation. In order to compare the expansion of different transport infrastructures independent from their different maximum network size (i.e. their saturation level), we show in Figure 18 the growth of the physical transport infrastructures normalized as percentage of the maximum network size, to demonstrate the recurring pattern of transport infrastructure life cycles in the USA. Figure 18 illustrates that time periods involved in the expansion of large infrastructure grids are indeed very long: the Δts* range from 31 years in the case of the expansion of the canal network to 55 and 64 years in the case of the growth of the railway and surfaced roads network respectively.

* time in years to grow from 10 to 90 percent of the ultimate saturation level

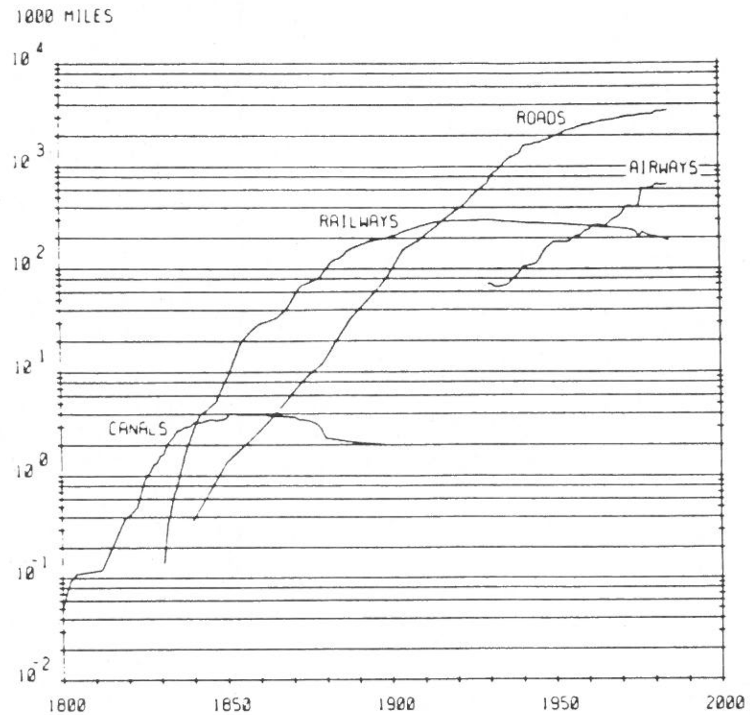

Figure 17 Evolution of the Length of Transport Infrastructures (Canals, Railways, Surfaced Roads and Federal Airways) in the USA, in 1000 miles.

If we consider the evolution of the length of individual transport infrastructures *relative* to each other, it is interesting to note that, measured by length, new infrastructures overtook existing ones only, at the time the latter started saturating. This was apparently the case with canals and railways in the 1840s, with railways and surfaced roads prior to the 1920s and based on this historical pattern, one would expect that airways could become as transport infrastructure more important (in terms of their length) only, once the expansion of the road network becomes completed (say shortly after the year 2000).

This sequence of replacements in the relative importance of individual infrastructures in the total length of the transport system of a country, points already to a possible analogy to market substitution. Seen from such a perspective, we may interpret the successive sequence of life cycles of individual transport infrastructures as changing gradually the morphology or structure of the transport infrastructure network of the USA. Previously dominant infrastructures, loose out on importance to newer systems, which in turn tend to saturate themselves and become replaced by newer infrastructures, developed in order to correspond better to evolving transportation requirements.

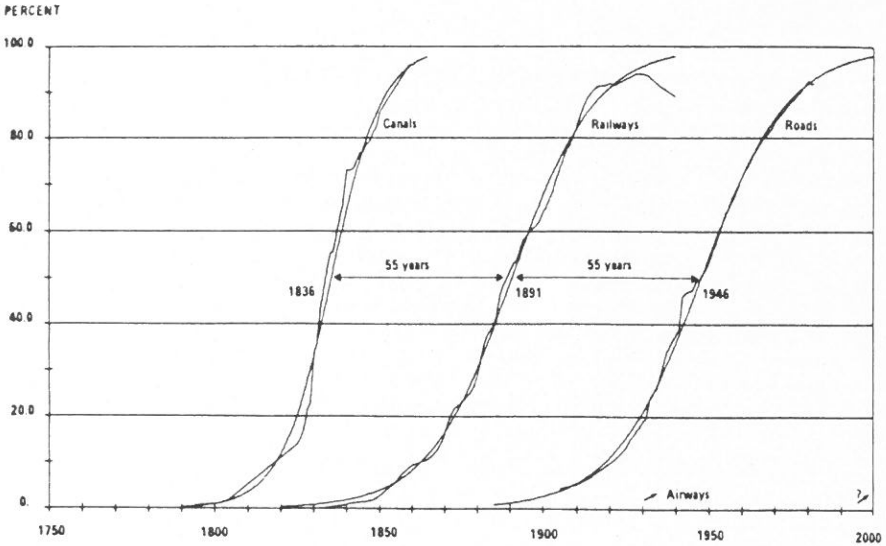

Figure 18 Growth of Physical Transport Infrastructures (Canals, Railways and Surfaced Roads) in the USA, in percent of saturation level.

Figure 19 reports an analysis of the share of individual transport infrastructures in the total length of the transport system of the USA, organized with the help of a multiple logistic substitution model.

Despite that the model fails to describe the rapid early phase of the introduction of the railways (i.e. prior to 1835), and the growth of air transport infrastructure (route mileage in operation) proceeds rather turbulent,* we can conclude that the model performs reasonably well in describing the long-term sequence of replacements in the importance of individual transport infrastructures in the USA.

Based on Figure 19 we can distinguish three main periods in the structural evolution of the transport infrastructure of the USA: 1) Growth of the railway network (main tracks operated) and resulting decline in the importance of canals as transport infrastructure, with a Δt of around 45 years. Note here, that around 1860 (maximum size of canal network), canals accounted for only more than 10 percent of the total transport infrastructure (canals, railways and surfaced roads) of the USA. 2) Saturation of the importance of railways (in the length of the transport infrastructures) around 1870 (i.e. around

* The turbulent trajectory of air transport infrastructure, should not come to a surprise, as we deal here with a conceptual and functional analogon, compared to *physical* infrastructures. Thus, whereas canals, railroads or surfaced roads evolve through a stable trajectory as the result of the inherent (long) lead times (and costs) in building up and decommissioning physical systems, air routes may easily be opened and closed down following particular market constraints (see e.g. the turbulent period during WW II in Figure 19

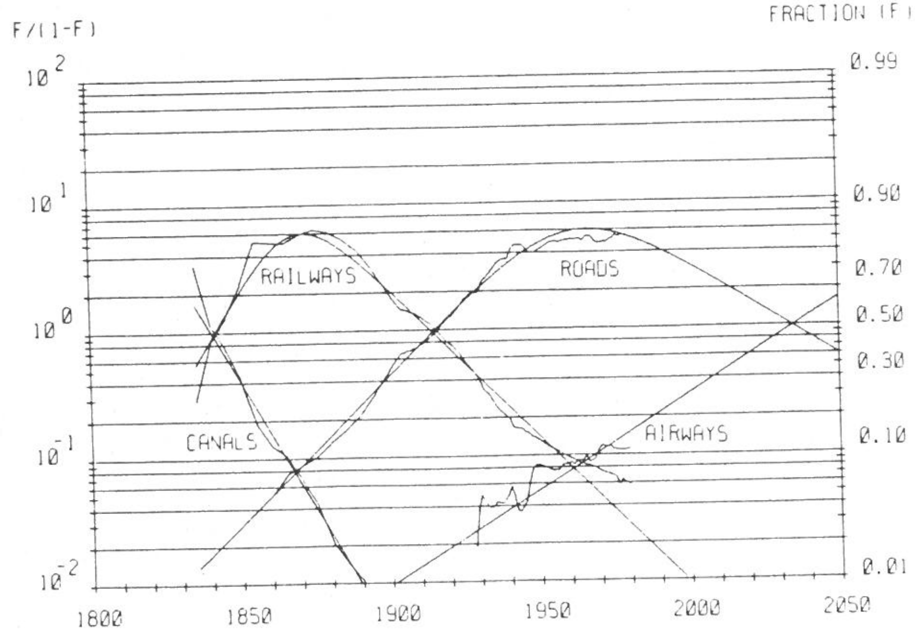

Figure 19 Infrastructure Substitution in the USA, logistic transformation of fractional shares.

50 years prior to the time when railways reached their maximum network size) and subsequent decline in importance, due to the growth of the surfaced roads share in total infrastructure, which proceeded with a Δt of around 80 years. 3) Long saturation phase (from around 1940 to 1980), during which road infrastructure constituted the most important transport infrastructure in the USA, to loose however this dominant position to the growing importance of air transport infrastructure, growing with a Δt of around 130 years.

Several invariant features in the long-term structural evolution of the US transport system appear from Figure 19 The dominant (longest) transport infrastructure is always longer than at least half of the total length. Consequently, the second and third longest infrastructures account for less than half of total length. This symmetry and the dominating role of individual infrastructures (canals prior to 1840, railways between 1840 to 1920 and roads since the 1920s and possibly until well into the 21st century) is complemented by another regular feature during the last 180 years. The time constants Δt in the growth (decline) of importance in the total transport infrastructure length increase from around 45 years (decline of canals and growth of railways) to 80 years (decline of railways and growth of surfaced roads) and appear to increase further to some 130 years (decline of road and growth of air transport infrastructure). This could be the result that each newer infrastructure system grows to considerably larger levels of network extension (although the Δts of the individual expansion pulses increase only from 31 to 55 and 64 years for canals, railways and roads respectively). A corollary of this observation is, that it takes apparently longer and longer time constants for the phasing out of transport

infrastructures. From such perspective, we appear in the long-term heading towards a situation of keeping infrastructure systems in parallel in achieving increasing levels of specialization in the services rendered by different transport infrastructures.

Finally, we note that the distance between the maxima in share of total infrastructure length between railways and surfaced roads is about 100 years, indicating the considerable time span involved in the transition from the dominance of one infrastructure system to the next one. Based on this conjecture one could expect the period of maximum dominance for the airways not to occur before the year 2050, before possibly loosing out dominance to new transport infrastructures such as for instance high speed Maglevs (magnetic levitation trains).

The difference in the dynamics (Δts) of the growth of individual infrastructures and their relative shares in total infrastructure length may appear at first sight as a contradiction. However, this difference is the result of the complex coupled dynamics of total infrastructure growth, and the growth and decline rates of individual transport infrastructures. As the total length of infrastructures increases, even rapid growth of individual infrastructures, such as airways will translate only into slower growth rates in their *relative shares*. Once the growth rates of an individual transport infrastructure fall behind the growth of the total system, their relative share starts to decline. Thus, the share of railways in total infrastructure length decreased since 1870, whereas railways continued to expand until the end of the 1920s. Similarly, the length of the surfaced road network still continues to increase (although, as close to apparent saturation, at relatively low rates), however their relative share started already do decrease in the 1960s. Thus, the total length of an individual infrastructure (in our case canals, railways and surfaced roads) can still be growing, even decades away from ultimate saturation and subsequent senescence in absolute network size, while the shares of its length in that of the whole transport system are already declining. The saturation and decline in relative market shares precedes thus saturation in *absolute* growth in a growing market (an expanding "niche"). This implies, that eventual saturation of any competing technology may be anticipated in the substitution dynamics in a growing market, such as for railways as early as 1870 and for roads as of 1960. The infrastructure substitution model presented above, may thus be considered as an "precursor" indicator model, on the long-term evolution of individual infrastructures.

Let us conclude this discussion on the long-term evolution of transport infrastructures in the USA by pointing out to a final fact, which is the regularity in the rise and fall of the importance of individual transport infrastructures. This regularity appears affected neither by the discontinuities in the development of individual transport infrastructures as part of their proper life cycle (growth, saturation and subsequent decline) nor by external events like the depression of the 1930s or the effects of major wars. We conjecture, that this stable behaviour may be the result of an invariant pattern in the societal preferences with respect to individual transport infrastructures, resulting from differences in the performance levels (seen as a complex vector rather than represented by a single measure) inherent to different transport infrastructures and technologies.

In the following we will show that the pattern we observed in case of the USA appears invariant, in terms that a similar structural evolution can be observed in case of a planned economy (the USSR), pursuing an entirely different transport infrastructure development policy. Our analysis of the structural change in the transport infrastructure network will then be further corroborated by using as a more accurate measurement of the importance of different transport infrastructures their *utilization (throughput)*, i.e. the modal split (passenger-km transported).

Figure 20 reports a similar analysis of the long-term infrastructure substitution for the case of the USSR, as we have analyzed in Figure 19 above for the USA.
As in the case of the USA, we complement our analysis by including as analogon the evolution of airways in the long-term structural evolution of the composition of the transport

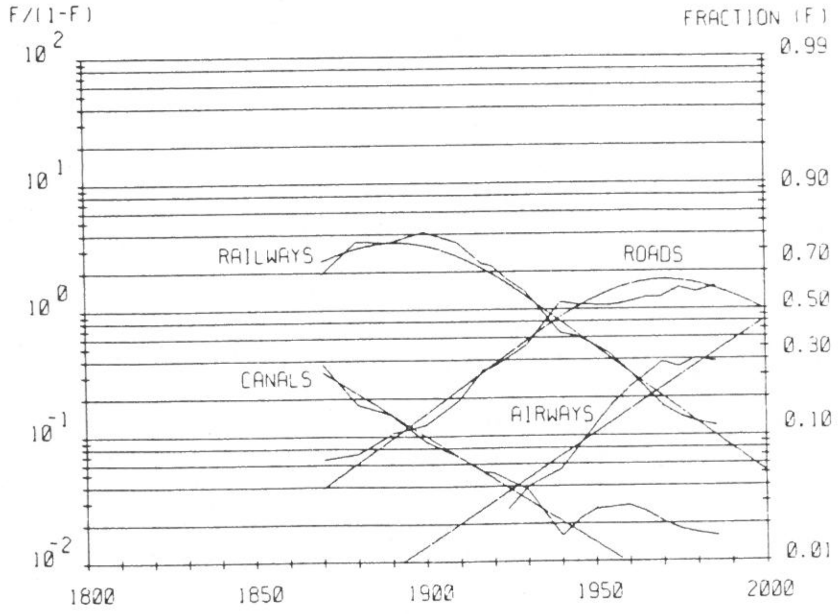

Figure 20 Infrastructure Substitution in the USSR, logistic transformation of fractional shares.

infrastructure of the USSR. Despite the fact that there may be some uncertainty related to our estimate of the early (pre 1965) extent of air transport infrastructure, we observe a similar dynamic picture with respect to the structural evolution of the transport infrastructure of the USSR as was the case for the USA.

Canals appear to have reached the end of their technological life cycle, in finally accounting only above one percent of the total transport infrastructure network. Railways reached their maximum importance at the beginning of this century and are loosing in importance to the growing surfaced road network. The latter (similar to the situation in the USA) is presently saturating in terms of its relative importance in the total length of transport infrastructures, to become in turn replaced in the long-term by a growing importance of airways as transport infrastructure. The time constants of this long-term structural change process are very similar. The Δts of the replacement of canals and railways and of the growth (and long-term replacement) of roads and airways are all around 100 years, as is the spacing between the (estimated) peaks in importance of railways and roads in the total length of the transport infrastructure of the USSR (again similar to the USA). Thus, the time period involved in the change from the dominance of one transport infrastructure (railways) to another one (roads) is again about 100 years.

This result is noteworthy for several reasons. 1) It demonstrates a continuity in the structural evolution of the transport system, which does not appear affected by major societal discontinuities such as the October Revolution. 2) The long-term dynamics and structural trends in the transport infrastructure evolution show similar results for a

planned economy as for a market economy. 3) The long-term structural change in the importance of different infrastructures in the USSR does not appear affected by the entirely different infrastructure development policy pursued by the USSR, compared to the USA. Recall here, that contrary to the USA where the length of railways and canals are declining or stagnating at a low mileage, the USSR (as practically only industrialized country) follows an ambitious canal and railway construction program.

Private car ownership in the USSR is at such low per capita levels (i.e. around 50 passenger cars per 1000 population) compared to the USA (562 cars per 1000 population), that the dominance of road infrastructure in the transport system of the USSR may come to a surprise. We will see below, that in terms of transport infrastructures, surfaced roads are of similar *functional* importance in the USSR, although road transport is not assured by private individual vehicles (cars) but instead by collective ones (buses).

Figure 21 present a summary comparison of the infrastructure substitution pattern between the two countries.

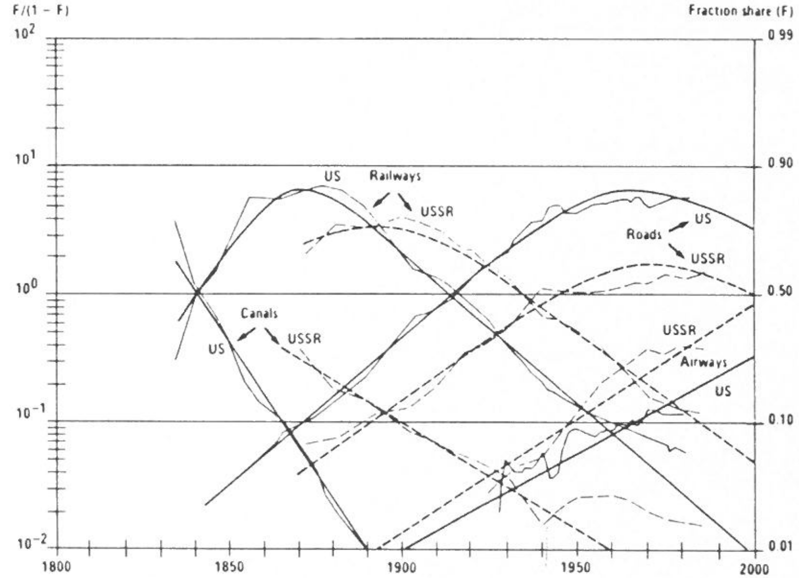

Figure 21 Infrastructure Substitution in the USA and the USSR, logistic transformation of fractional shares.

The pattern of temporal changes in both countries is marked by a high degree of regularity and quest for infrastructures allowing higher speed and productivity. The phase transitions, however, are lagged by a few decades when compared with the United States. This is best illustrated with respect to the relative share of canals in the transport infrastructures in the two countries. Canals accounted already before 1900 for less than one percent of the total transport infrastructure in the USA, whereas their logistic decline trajectory proceeds much slower and is in addition lagged by several decades in case of the USSR.

As a function of a continuation of an ambitious canal construction program since the 1940s, the decline trend of the share of canals in the total transport system has stopped and remained at a level of around 2 percent during this time period. The dominance of railways lasts also longer than in case of the USA, i.e. until the 1940s while in the United States it ended two decades earlier, their logistic decline trajectory (in relative market share terms) however, proceeds at a similar rate (see similar slopes of their logistic decline trajectory in Figure 21) in both countries in symmetry to the similar rate of the increasing importance of road infrastructures. Particularly noteworthy is also the very similar (in case of the USSR slightly faster) penetration trajectory of air transport infrastructures

Thus, during the last decades, the development of the transport infrastructures in the two countries has been converging. Also the rate of relative growth in the importance of road infrastructure and their saturation appear synchronized in the two countries. Thus, there is an increasing congruence and similarity in the structural and functional evolution of the transport system in the two countries, especially in the longer term. The transport infrastructure system in both countries after the year 2000, should the structural change pattern continue to unfold as in the past, would see railways accounting for only a few percent of the infrastructure network, with a dominance of roads and airways accounting for approximately equal shares each.

To a large extent this convergence and similarity in the structure of the transport system between the USA and the USSR is due to the fact that both countries have relatively low population densities and vast territories that modern transport systems must bridge in a matter of hours. The infrastructural evolution in both countries appear thus to transgress differences prevailing in transport cost structures and market clearing (centrally planned versus market economy) mechanism, pointing at a deeper causality responsible for the observed pattern. This is seen primarily related to the performance (in particular travel speed) differentials between individual transport systems and their correspondence to individual (travel) time allocation and resulting decision and preference criteria favoring faster transportation means.

3.2. Evolution of Performance of Infrastructures (modal split)

As a second step in our analysis of the importance of various transport infrastructures we analyze their relative contribution to the transportation output of the USA and the USSR. This will be done in analyzing the modal split in terms of the passenger-km performed by various transport modes for long-distance (intercity) travel, where comparable data exist between the two countries. Our analysis covers some shorter time series (since 1950) for the USA. In the case of the USSR available data permit to go back in our analysis as far as 1920.

As corresponding measure of the output (performance) of individual transport infrastructures we consider in the following cases the long-distance passenger modal split. This measure is a more accurate indicator for assessing the relative importance of different transport modes as the relative share in total infrastructure length discussed above. Long-distance passenger transport is a good precursor indicator, in terms that it constitutes the highest value transport market niche, in which competition between different transport modes is largest.

Figure 22 shows the substitution of different transportation modes in intercity (long-distance) passenger travel in the United States and the Soviet Union. By excluding urban and metropolitan transport, the competition for intercity passenger traffic is reduced to four major transport modes in the United States: railways, buses, cars and airways. The major competitors over this period in the Soviet Union were boats, railways, buses and airways; the automobile never gaining any significant importance. Today, most of intercity travel is by car and plane in the United States and by in the rail, bus and plane in the Soviet Union. Comparison of the two countries shows that in the past

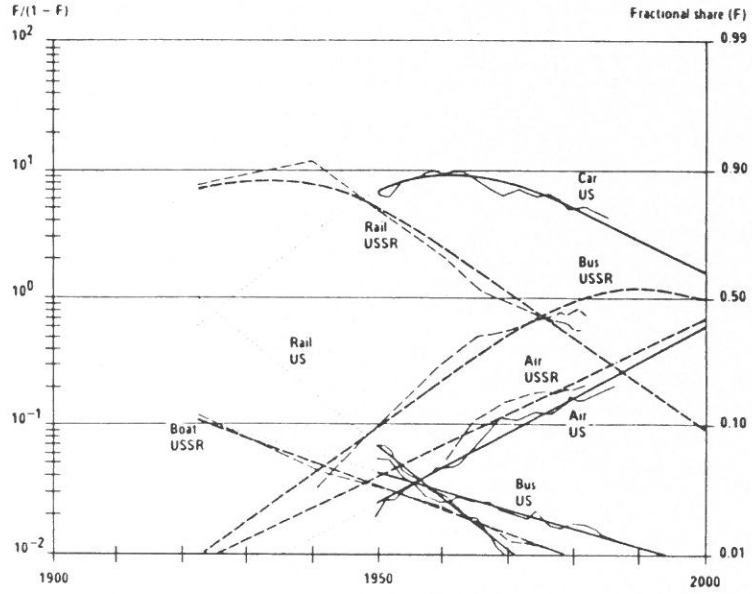

Figure 22 Intercity Passenger Modal Split in the USA and the USSR, logistic transformation of fractional shares of passenger-km.

the intercity passenger transport development portrayed a phase-shift in the two countries. Rail and bus are virtually extinct in the United

Thus the structural change of the long-distance passenger modal split in the USSR compared to the USA is, combines in fact a structural evolution of different technological life cycles. Whereas inland navigation for long-distance passenger transport has disappeared only by the middle of the 1970s, railways still account presently for some 37 percent of all intercity passenger-km of the USSR. Their displacement process appears however, while lagged by some 50 years (t_0 1972 compared to 1921 in case of the USA), to proceed at a similar rate (Δt of 55 years compared to 48 years in the USA). The apparent saturation in the market share of road transport (buses) appears to occur somewhere in the 1990s (i.e. some 30 years after the saturation of the market share of cars in the USA). The most striking similarity however occurs in the area of the market share of air transport. Air transport in the USSR presently accounts for some 18.5 percent of all intercity passenger-km, compared to 17.6 percent in the case of the USA. The dynamics of market share gains of air transport in the case of the USSR (t_0 estimated to occur in 2006, Δt 77 years) approach very closely the market dynamics of air transport in the case of the USA (estimated t_0 in 2008 and Δt of some 70 years).

The substitution dynamics indicate thus that by the end of the century airways may become the dominant form of intercity travel in both countries. The evolutionary paths appear to be convergent. Equally important is that the average choice of different modes of passenger travel changes consistently in both countries and favors faster and more productive systems. This points at very similar *comparative advantages* of transport modes in the long distance passenger modal split between the two countries (as reflected in their similar rates of change Δts), whether we consider negative comparative advantage (as in the case of the decline of importance of railways) or positive comparative advantage (as in the case of the growing air transport). This similarity in the dynamics of structural change in modal split is the more surprising, considering the differences between a market and centrally planned economy and especially with respect to the policies pursued with respect to alternative transport modes between the two countries (promotion and further railway construction in case of the USSR compared to effective disappearance of railways as passenger transport mode in the case of the USA, road transport by public buses as opposed to private car ownership, etc.).

We consider, that the influence of air transportation in altering the human space-time activity framework, i.e. the increased travel range resulting from the higher transport speeds of air travel, is responsible for the market share gains of air transport. This comparative advantage (as resulting from fundamental human time allocation mechanisms) is influencing in a similar way individual time budgets and resulting long-distance transport mode preferences of the people in the two countries.

We conclude, that the historical evolution of transport infrastructures is characterized by a regular sequence of replacements, where newer systems, corresponding in their performance levels better to societal preferences with respect to transport infrastructures, replace existing ones. This evolutionary process appears invariant between different economic systems, and in addition of a highly homeostatic nature. We conjecture, that the main driving force responsible for this invariant evolutionary pattern relates to the human time allocation mechanism (the "law of constant travel time" as formulated by Zahavi, 1979 and 1981) favoring faster transport systems. From this perspective, the impacts of the growing importance of air transport and its infrastructure and especially the optimal integration of of the various transportation modes providing feeder functions to and from hubs constitutes an important area of concern for transport policy and planners.

3.3. Long-term Passenger and Goods Transport Intensity of an Economy

As a conclusion of our discussion on the long-term structural transformations in the transport sector we will consider similar as in the case of the energy intensity discussed above, the case of the long-term evolution of the transport intensity of an economy. Similarly as in the case of energy, where a comprehensive analysis of the long-run primary energy to GNP relationship requires the inclusion of noncommercial energy sources, which were the dominant energy source throughout much of the 19th century, an analysis of the transport intensity has to go beyond today's commercial transport carriers. We have thus to include in addition to long-distance (intercity) traffic presented above in our analysis of the USA and the USSR, also short distance transport operations and include transport modes like horses which have entirely vanished as transport technologies in industrialized countries.

Figure 23 presents a long-term analysis in the case of France, where we have been able to reconstruct long historical time series since 1800 based on estimates of Toutain, 1967 and 1987, which were complemented by additional data obtained from a variety of sources (for details see Grübler, 1988).

The increasing transport intensity as shown in Figure 23 is in contrast to the long-term declining energy intensity shown in our analysis of the case of the USA above. Another interesting finding of Figure 23 is, that the passenger transport intensity (passenger-km performed per unit of constant GDP) increases much stronger than the goods transport intensity, particularly since the 1930s and especially after the war, as a

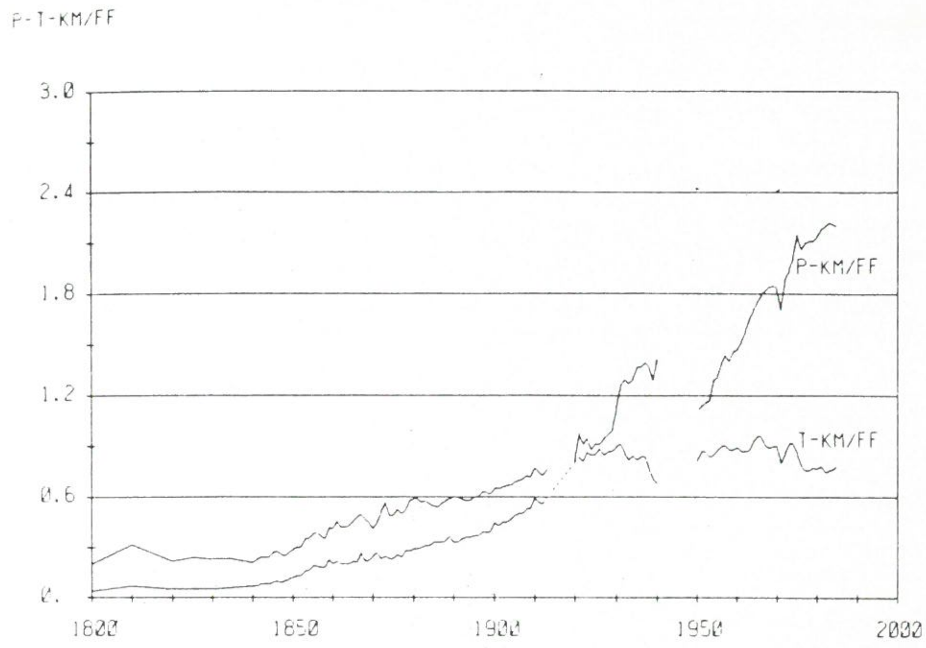

Figure 23 Transport Intensity of French Economy Since 1800, in ton- and passenger-km transported (all transport modes) per constant 1913 French Franc GDP.

result of the availability of new transport modes such as private automobiles and aircrafts. Both variables show however that from a long-term perspective the *transport intensity* of the French (and we would conjecture similar tendencies also in other countries) economy has been steadily rising. The goods transport intensity increased by a factor of 4 between 1800 and 1985 (i.e. from 0.197 to 0.79 ton-km per 1913 Franc GDP) corresponding to an average growth rate of 0.75 percent/year. The passenger transport intensity has on the other hand side increased by a factor over 53 (!) during the same time period (i.e. from 0.04 to 2.2 passenger-km per 1913 Franc GDP), which corresponds to an average growth rate of 2.2 percent/year.

One should not be biased by the apparently high figure of transport output per unit of constant GDP, as the GDP data are expressed in constant 1913 money. In terms of current Francs the transport intensity in 1985 corresponded to 0.136 passenger-km and 0.049 ton-km per current Franc of the French GDP. The increasing transport intensity of the French economy observed in Figure 23 does however not imply that a ever higher share of the GDP is spent in the transport sector, at least not at a similar fashion, as the increase of transport intensity per constant unit of GDP discussed above would suggest. This because of the drastic decrease in the real-term transport costs over the whole time horizon.

On basis of Figure 23, we can conclude that industrialized countries appear to be amidst of a transition to increasing "dematerialization" (i.e. higher information and value and lower material content of the output mix of their economies) as reflected in the decreasing ton-km intensity per unit of economic activity in France since around 1970. However it is also important to note that for the demand for passenger travel no similar decoupling tendencies can be observed, which indicates the likelihood of further significant increases in personal travel along with economic growth.

4. Conclusion

In a number of examples analyzed in this paper, we have shown that the long-term structural changes in energy and transport systems can be described as a regular process with logistic secular trends. The changing morphology of the energy and transport system can thus be described as a regular sequence of replacements, proceeding along structured evolutionary paths. The time constants of change in the area of energy and transport systems thus identified are very long: ranging from several decades to even a century. A general conclusion of a large number of studies of technological change is that substitution processes at the level of existing capital or rolling stock such of road vehicles or locomotives, last a few decades, while changes in the energy and transport system and their respective infrastructures are much longer processes. We almost can speak of a temporal hierarchy in the pattern of technological change: the more one moves upwards in a hierarchy from the micro level to the level of large pervasive systems and infrastructures the slower the process of change.

The regular pattern and the long time spans involved in long-term structural change processes at the level of transportation or energy infrastructures may be considered as one of the most stringent constraints for shorter term adjustment processes. This means that infrastructures themselves change only slowly, but concurrent (and much faster) technological changes in end-use devices and equipment can improve both the performance, efficiency and quality of the transport services provided by the current systems.

Both transport and energy systems must meet more stringent economic and environmental requirements over *longer distances and in a shorter time*. Our results of transport infrastructure evolution and energy substitution indicate that aircraft and natural gas represent evolutionary rather than revolutionary technologies that could meet these more stringent requirements. Both aircraft and natural gas technologies could meet these requirements through refinements and improvements in current designs and practices during the next two decades, but new solutions and corresponding new infrastructure systems must be developed during this period for the decades thereafter.

5. References

B.P. (British Petroleum), 1987, BP Statistical Review of World Energy, London, B.P.

Fisher, J., 1974, Energy Crises in Perspective, New York, John Wiley and Sons.

Fisher, J. C., and R. H. Pry, 1970, A Simple Substitution Model of Technological Change, Report 70-C-215, Technical Information Series, Schenectady, New York: General Electric Company, Research and Development Center.

Griliches, Z., 1957, Hybrid Corn: An Exploration in the Economics of Technical Change, Econometrica 25:501-22.

Grübler, A., 1987, Aufstieg und Fall von Infrastrukturen, Der öffentliche Sektor 3/4:52-83, Vienna.

Grübler, A., 1988, The Rise and Fall of Infrastructures, Dynamics of Evolution and Technological Change in Transport, Vienna, Technical University.

Mansfield, E., 1961, Technical change and the Rate of Imitation, Econometrica 29:741-766.

Marchetti, C., 1979, Energy Systems - The Broader Context, Technological Forecasting and Social Change, 14:191-203.

Marchetti, C., and N. Nakićenović, 1979, The Dynamics of Energy Systems and the Logistic Substitution Model, RR-79-13, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Nakićenović, N., 1979, Software Package for the Logistic Substitution Model, RR-79-13, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Nakićenović, N., 1984, Growth to Limits, Long Waves and the Dynamics of Technology, Vienna, University of Vienna.

Nakićenović, N., 1988, Dynamics and Replacement of U.S. Transport Infrastructures, in: J.H. Ausubel and R. Herman (Eds.), Cities and their Vital Systems, Infrastructure Past, Present and Future, National Academy Press, Washington, D.C.

Pry, R. H., 1973, Forecasting the Diffusion of Technology, Report 73CRD220, Technical Information Series, Schenectady, New York: General Electric Company, Corporate Research and Development.

Putnam, P. C., 1953, Energy in the Future, New York: D. Van Nostrand Co.

Reynolds, R. and A. Pierson, 1942, Fuel Wood Used in the United States, 1630-1930, USDA Forest Service, Circular No. 641.

Schurr, S. and B. Netschert, 1960, Energy in the American Economy, 1850-1975, An Economic Study of its History and Prospects, Baltimore, Published for Resources for the Future, Inc. by The Johns Hopkins Press.

Toutain, J.C., 1967, Les Transports en France de 1830 à 1965, Cahiers de l'ISEA, Série Histoire Quantitative de l'Économie Francaise No. 8, Presses Universitaires de France, Paris.

Toutain, J.C., 1987. Le Produit Intérieur Brut de la France de 1789 a 1982, Cahiers de l'ISMEA, Série Histoire Quantitative de l'Économie Francaise No. 15, Presses Universitaires de Grenoble, Grenoble, France.

U. S. Department of Commerce, 1970, Historical Statistics of the United States, Colonial Times to 1970, Part 1 and 2, Bureau of the Census, Washington D. C.

Zahavi, Y., 1979, The "UMOT" Project, Report DOT-RSPA-DPB-20-79-3, prepared for U.S. Department of Transportation and Ministry of Transport, FRG, NTIS, Springfield, Va.

Zahavi, Y., 1981. The UMOT-urban Interactions, DOT-RSPA-DPB 10/7, U.S. Department of Transportation, Washington, D.C.