Measuring Energy Substitution
An Introduction

by Bengt-Christer Ysander*

For a small open country like Sweden the ability of the manufacturing industry to adjust rapidly and smoothly to changes in relative world prices is of crucial importance. The oil price hikes in the 70s tested this ability in a dramatic fashion, creating at the same time a particularly good opportunity for studying the mechanisms and the adjustment problems involved in industrial factor substitution. Economists all over the world have hastened to exploit this opportunity and, as a result, our knowledge of industrial production structure and factor adjustment has increased considerably over the last few years.

The papers assembled in this volume, focusing on energy use in Swedish manufacturing, all share this common aim of mapping and measuring industrial adjustment to price changes. However, as appropriate for a still developing research area, they try alternative approaches, using different models and analytical techniques.

Mechanization and energy use - the postwar experience

From the end of the war and up to the first oil price hike the ongoing mechanization had a dominant influence on energy use in Swedish manufacturing. This is one of the main lessons to be learned from the first paper — by Joyce Dargay — tracing energy prices and energy use in the postwar period.

---

* The Industrial Institute for Economic and Social Research (IUI), Stockholm.

Mechanization meant continuous substitution of both capital and energy for labor in industrial production and it usually also meant electrification. The specific use of electricity (i.e., electricity input per unit of output) in manufacturing thus increased steadily during the 50s in spite of sharply rising electricity prices, stagnated in the 60s when prices were falling both in absolute terms and relative to other energy prices, but started mounting again in the latter part of the 70s, when electricity prices were catching up with other energy prices.

The other dominant postwar trend in energy use, viz. the switch from solid fuels to oil, can also partly be interpreted as a way of saving labor, whose wages, up to 1973, grew steadily faster than both capital and energy costs. Although the coal price paid by manufacturing industry tended to keep pace with the oil price, although with a certain time-lag, the labor costs involved in handling coal made oil advantageous to use, at least up to the first oil price hike. In terms of specific usage the main switch to oil occurred already in the early 50s. Although prices for heavy oil were completely stagnant up to the middle 60s, the specific oil use in manufacturing did not increase further during this period. The first oil price hike led to a considerable drop in specific oil usage, which then further declined at the end of the decade.

The continued rapid decrease in the use of solid fuels together with the curtailed oil use during the 70s, resulted in a steady reduction in total specific energy use from the middle of the 50s. About a quarter of this total energy saving was due to the changing branch structure. With a few exceptions — printing, chemicals and shipbuilding — the energy/output ratio fell in all manufacturing branches.

The dominant influence of mechanization and labor saving on industrial energy demand underlines the importance of analyzing energy use within the framework of production models, incorporating all the substitution possibilities between different inputs.

There are wellknown reasons for distinguishing also between manufacturing branches with different technologies, different capital/output and energy/output ratios. The price data collected by Dargay indicate another and less discussed reason for disaggregation. They show i.a. that in 1968 the energy intensive branches — pulp and paper, chemicals, primary metals and non-metallic mineral products — paid on the average 1/7 less for heavy oil and only half for electricity compared to other branches. After 1973 there is, however, a considerable convergence of energy prices between branches — in the case of petroleum products partly due no doubt to a shortening of contract lengths and a reduced significance of rebates to large consumers. The energy intensive branches thus faced a rise in energy prices during the first half of the 70s that was on the average 40 % larger than that experienced by the other sectors.

Alternative ways of measuring factor substitution

Any attempt to explain and measure the possibilities of factor substitution in industry must at the start make two kinds of basic choices: a choice of aggregation level and a choice of adjustment paradigm.

The choice of aggregation level involves at least three different dimensions of the production structure: technologies, factors and firms.

There is obviously a limit to the degree of technological detail that could and should be included in an economic production model. The use of approximate (or "generalized") descriptions of technologies in the form of production functions and of aggregations of those over different technologies has the advantage of small data requirements and computational ease. The disadvantage is the introduction of approximation and aggregation errors in the numerical results. The risk of distorted and biased results caused by the functional approximation is particularly great if the functions used are such as to place a priori restrictions on the factor substitution to be measured. Fortunately, theoretical research in the 60s and 70s has provided us with a family of flexible functional forms — like the generalized Leontief and the translog — which i.a. do not assume any restrictions on substitution elasticities. (For a short survey of these developments cf Field-Berndt, 1981). Simulation experiences indicate, however, that the aggregation error may still be large enough to make it difficult to get stable and consistent estimates on substitution relationships from aggregate descriptions of technologies (cf. Kopp-Smith, 1981).

The conditions for subsuming different machines and constructions under an aggregate capital measure in the production function, or for aggregating different labor inputs, have been thoroughly discussed over the past years. In the present studies we have to deal with yet another kind of factor aggregation — the aggregate treatment of different sources of energy: petroleum products, solid fuels and electricity. This brings into focus two new and interesting questions. Is the supply of energy to industrial plants so flexibly designed that energy production from different sources, the primary energy allocation system, can be treated as completely independent — or "separable" — from other technological decisions?

The second question is concerned with causation in the opposite direction. How much will the optimal internal allocation of "composite capital" depend on the relative price of energy and the way it is produced? As stated above already, our intuitive reading of the postwar experience indicates a rather strong dependence both ways. The rapid substitution of oil for coal in the 50s was probably not only motivated by labor saving but in part dictated by the new technologies imported from the U.S. On the other hand it seems evident that the oil price hikes in the 70s had an important and different impact on the profitability of different types and vintages of existing production capital. Any results of studies dealing with aggregate capital and energy in the conventional way must therefore be interpreted with great caution.

Aggregating over firms adds yet another problem dimension, since i.a. the rate of utilization in the firms may vary with their different positions and strategies in the market and the resulting allocation of production between firms may not be stable over time.

Having decided on the proper level of aggregation, one is still left with the choice of adjustment paradigm, i.e., the decision about how to model the way production is adjusted to market changes.

One part of this choice is concerned with modeling the market on which the producers are supposed to operate. If one should take into account the particular kind of say oligopolistic market structure involved, should discern both sides of the mutual adjustment of supply and demand, and consider also the possibility of disequilibrium pricing, the modeling ambitions could easily outrun the available resources for estimation. The market is therefore usually treated in a very simplified manner, e.g., by assuming the producers to be price takers and/or to operate within fixed market shares.

Changes in market conditions should call forth two kinds of supply adjustment. The short-run adjustment is concerned with accommodating the market changes within existing capacity by changing the current production and with that the cost-minimizing input demands. The long-run adjustment is initiated by capacity changes, due both to technological changes and to investment/ scrapping activities. Since adjustment is costly and time-consuming the relevant decisions will stretch far into the future, and will depend on expectations which in turn may be built on historical experience going far back in time. Because of the obvious difficulties involved, very few attempts have been made so far to model this adjustment process explicitly as an intertemporal optimization under uncertainty. In most studies the dynamic element is simply represented by some rather ad hoc lag structure or accelerator relation. (For a thorough discussion of the various stages of dynamic adjustment representation, see Berndt, Morrison, Watkins, 1981). Indeed the majority of studies on factor substitution documented so far are based on static models which completely disregard the existence of adjustment constraints, assuming instead full and instant adjustment. As for technical change, whether embodied or disembodied, it is usually, for the sake of computational convenience, treated as neutral. This means disregarding the possibility — observed in many econometric studies — that the rate and direction of productivity change may depend on relative factor prices.

Most of these problems of simplification and model choice are exemplified in the four studies of industrial substitution possibilities contained in Part II. Table 1 gives some indication of the variety of models and methods used by the different authors. The first five columns are concerned with characteristics related to aggregation, the next two columns reflect the choices of "adjustment paradigm" while the last indicates the method of estimation.

Dargay uses time-series data to estimate factor cost shares for twelve manufacturing branches. Her translog cost functions include capital, labor, intermediate goods and energy as arguments, with the aggregate energy input being alternatively measured directly in terms of physical energy units or estimated indirectly as a cost-minimizing mix of primary energy inputs. Her model is essentially static with Hicks neutral technical change. Both a homothetic and a non-homothetic functional form were tried, with the non-homothetic formulation giving more significant and consistent results. Dargay did not, however, succeed in producing separate estimates of rates of return to scale and technical change. An FIML estimation program was used in two stages — firstly to estimate the cost-minimizing energy mix and secondly to estimate the cost-shares of the aggregate factors.

The Jansson study of the Swedish iron and steel industry is based on the same time-series data as the Dargay study, deals with the same aggregate factors — although assuming proportionality in

Table 1 **Four approaches to measuring factor substitution**

Author	Data	Coverage	Level of aggregation	Production factors in estimation	Form of cost function	Adjustment constraints	Technical change	Method of estimation
J. Dargay	Time-series 1952-76	Total manufacturing	Twelve branches	C,L,M,E, E(e,o,s)[a]	Translog	-	Hicks neutral	Two-stage FIML
L. Jansson	Time-series 1952-75	Iron and steel	Capital vintages of branch	C,L,E[a]	-"-	Vintage capital, demand and profit development	-"-	FIML
L. Hultkrantz	Cross-section statistics and engineering data, 1979	Wood, pulp and paper in northern Sweden	Production activities within plants	I,L,M,e,o[a]	Linear	Supply constraints, current capacity and available investment options	-	LP
S. Lundgren	Engineering data	Iron and steel	Production activities	C,L,M, e,o,s[a]	-"-	Short run: capacity constraints Long run:-	-	LP

[a] C = Capital
L = Labor
E = Energy
M = Intermediate goods
I = Capital investment
e = electricity
o = oil
s = solid fuels

the use of intermediate goods — and used the same FIML estimation program. However, there are several important differences between his study and that of Dargay. In the Jansson study the different annual vintages of production capital are distinguished in the model. Even more important, Jansson's model, which is of the "putty-clay" type, attempts to explain adjustment in terms of gross investment and scrapping of production capacity, with the technology of the new capacity reflecting current factor prices. In such a dynamic model, "substitution effects" in terms of technological adjustment in new capacity to short-term changes in factor prices, may be overlaid and dominated by "vintage effects", resulting from adding new capacity to an existing stock which reflects techniques and prices over the past thirty years. Two inputs — like capital and energy — may then be substitutes in the technological sense and yet be complementary over time, even without non neutral technical change. In this way, the Jansson study reconciles the diverging and controversial results obtained in earlier studies of the elasticity of substitution between capital and energy in various manufacturing branches.

The two following studies both use a radically different approach. Instead of time-series data they use cross-section and/or engineering data and are then able to model individual production activities within the branch in question. For the same reason there is no need for them to try to aggregate the diverse kinds of primary energy resources. To be able to handle this mass of technological information, they are forced to linearize all relations, so that optimal production plans can be computed with linear programming techniques. While in the preceding studies elasticities of substitution could be computed from parameters of the estimated functions, they can now only be very roughly approximated by comparing the outcome of different runs of the LP-models.

Hultkrantz' study of the wood, pulp and paper industry in northern Sweden encompasses two periods and includes different packages of investment options for the two time horizons. The options are those currently considered by the firm at the time of the enquiry (1979). In terms of this multiperiod model Hultkrantz can define a concrete and specific meaning and measure for the distinction between short-run and long-run adjustment. A special feature of the Hultkrantz model is the fact that the paper and pulp industry is here embedded within a larger model, which takes explicit account of alternative uses of wood — for the sawing industry and more particularly for heat generation. One of his main conclusions, of great importance and relevance for current Swedish energy policy, is that only very drastic further increases in the relative oil price could make wood-based heating stations a serious competitive threat for the forest-products industries. This and related results are derived by maximizing the quasi-rents to industrial capacity and the price of stumpage subject to the constraints set by industrial capacity, investment opportunities and available volumes of wood of different kinds.

Lundgren's study of the iron and steel industry is entirely based on engineering data and blueprints for future technologies. His model is essentially a static one-period model with explicit capacity constraints. Long-run adjustment can be defined and measured by eliminating all capacity constraints. While Hultkrantz' experiments are based on maximizing profits or quasi-rents, Lundgren's simulations all deal with cost-minimization, holding the output mix constant.

Some numerical results

Four different ways to model reality lead to four different modes of designing questions about factor substitution — and imply four different types of answers. We will make no attempt here to survey or summarize the numerical results recorded in the four studies in Part II. The examples presented in Table 2 below merely serve the purpose of illustrating the variety of numerical experiments performed and of substitution mechanisms investigated.

Table 2 Elasticities of substitution - some numerical results*

Author	Type of elasticity	Branch	Energy(oil) - - Capital	Energy(oil) - - Labor	Oil - Electricity	Energy (oil)
J. Dargay	Allen partial elasticity of substitution = σ	Total manufacturing	$\sigma_{EC} = -1.43$[1]	$\sigma_{EL} = 0.12$[1]	$\sigma_{oe} = 0.21$[2]	$\varepsilon_{oo} = -0.29$[3]
	Price elasticity = ε	Wood, pulp and paper	$\sigma_{EC} = -0.59$[1]	$\sigma_{EL} = 0.02$[1]	$\sigma_{oe} = 0.22$[2]	$\varepsilon_{oo} = -0.28$[3]
		Iron and steel (Primary metals)	$\sigma_{EC} = -0.66$[1]	$\sigma_{EL} = -0.61$[1]	$\sigma_{oe} = 0.24$[2]	$\varepsilon_{oo} = -0.26$[3]
L. Jansson	-"-	Iron and steel	$\sigma_{EC} = 0.82$[4]	$\sigma_{EL} = 2.63$[4]		$\varepsilon_{EE} = -0.98$[4]
L. Hultkrantz	Arc cross-price elasticity = e (profit maximization under supply and capacity constraints)	Wood, pulp and paper in northern Sweden	$e_{Io} = -0.57$[5]	$e_{Lo} = -0.29$[5]	$e_{eo} = -0.72$[5]	$e_{oo} = -0.49$[5]
S. Lundgren	Arc cross-price elasticity = e (cost minimization for given output mix)	Iron and steel			$e_{eo} = -0.26$[6]	$e_{oo} = -4.3$[6]

[1] Homothetic cost function, direct estimates.

[2] Partial substitution effects, total energy consumption constant.

[3] Total own-price elasticity, non-homothetic total cost function.

[4] Elasticities of the ex ante production function

[5] 50% oil price increase, long-run adjustment including output change.

[6] 50% oil price increase, long-run adjustment without investment constraints, output constant.

The different approaches are reflected in different notions and measures of the elasticity of substitution.

The concept of elasticity of substitution, as originally introduced by Joan Robinson (Robinson, 1933), is intended to measure the ease of substituting between two inputs, when output is held constant. It is usually defined as the derivative of the ratio of two input levels with respect to the ratio of the two corresponding input prices. For a production function with two inputs, $Q = f(x_1, x_2)$, and the corresponding input prices, p_1, p_2, the elasticity of substitution can be written as:

$$\sigma_{12} = \sigma_{21} = \frac{d\ \ln(x_1/x_2)}{d\ \ln(p_2/p_1)}$$

σ_{12} here grows larger as substitution becomes easier. Also, when $\sigma_{12} > 11$, the cost share of input 1 becomes larger relative to the cost share of input 2 when input 2 becomes relatively more expensive.

With more than two inputs involved, however, different definitions of elasticity result from different choices of the ceteris paribus conditions under which the partial derivatives are obtained.

The most commonly used definition — the Allen-Uzawa partial elasticity of substitution — is simply a price cross elasticity weighted by the inverted value of the corresponding cost share:

$$\sigma_{ij} = \left.\frac{d\ \ln(x_i/x_j)}{d\ \ln(p_j/p_i)}\right|_{Q=\bar{Q}} = \frac{1}{p_j x_j / \sum_i p_i x_i} \cdot \left.\frac{d\ \ln\ x_i}{d\ \ln\ p_j}\right|_{Q=\bar{Q}} = \frac{1}{k_j}\ \varepsilon_{ij},$$

where k_j denotes the cost share of the jth input. In this definition all other inputs adjust optimally to the price change.

As shown in Table 2, Dargay's elasticity measures for total manufacturing show complementarity between capital and energy, while energy and labor appear to be relatively independent of each other. Oil comes out as a rather poor substitute for electricity, and also registers a low own-price elasticity.

Her results for the wood, pulp and paper industry, also shown in the table, are very similar to those for total manufacturing although the complementarity between capital and energy does not register as strongly for this branch.

For iron and steel and other primary metal industries the one main divergence in results, compared to the wood, pulp and paper industry, is the complementarity here registered also between energy and labor.

As for intermediate goods, Dargay's results seem to support the conclusion from Parks' earlier study of Swedish manufacturing 1870-1950, that capital and labor in most branches are not separable from intermediate goods (Parks, 1971).

In Jansson's production model for the iron and steel industry a distinction can be made between the "potential" substitution possibilities of the ex ante production function and the actual realized substitutions, which may to a large degree be determined by "vintage effects", i.e. by the inertia due to older capital vintages. This may explain why energy shows up in his study as a strong substitute in the more narrow technological sense for both capital and labor, while the opposite result is derived from Dargay's static model.

The elasticity measures recorded in the LP studies of Hultkrantz and Lundgren are quite different from those used in the preceding papers. Firstly, they are arc elasticities, which means that instead of being computed from parameters of estimated production functions, they are rough measures of average effects of intramarginal — and in fact quite drastic — price changes in the model simulations.

Secondly, there are in the simulations important constraints — concerning production capacity and raw material supply — on the adjustment of inputs. In this regard their elasticity measures are not so much related to the Allen-Uzawa elasticity as to the concept of "direct" elasticity of substitution, which holds constant other inputs than those directly concerned (McFadden, 1963).

Thirdly, what they compute are straightforward cross price elasticities and not elasticities of substitution, although these two concepts are closely related (cf. above).

Finally, in the case of Hultkrantz, the elasticities are not computed with output held constant, which means that the measured effects of input price increases are also influenced by shrinking total production.

This last point probably to a large extent explains why Hultkrantz finds energy to be a complement not only to capital but also to labor. For the case of regular neoclassic production functions it has been shown (Field-Allen, 1981) that a cross price elasticity with freely variable output can be defined as;

$$\eta_{ij} = \left.\frac{d\ \ln x_i}{d\ \ln p_j}\right|_{\substack{p_k = \bar{p}_k \\ k \neq j}} = \varepsilon_{ij} + k_j \eta \psi ,$$

where k_j is the cost share of the jth input, η denotes the (cost) price elasticity of output, while ψ represents a function of the rate of return to scale such that $\psi = 1$ when this rate is constant.

In Hultkrantz's model, output will decrease with rising costs while the rate of return to scale is non-increasing. Even if an oil price hike would mean that capital and labor tended to replace energy the consequent downscaling of production could therefore lead to complementarity being registered with this kind of elasticity measure. The same evidently is true in regard to the substitution relation between electricity and oil.

The weak complementarity between oil and electricity in Lundgren's model of the iron and steel industry seems instead to be caused mainly by switches between different technologies. In the short-run version, with effective constraints on investment, the sign of the elasticity is reversed due to the fact that the electric arc furnace is then still a viable option. The own-price elasticity for oil recorded by Lundberg for the long-run version seems surprisingly large, which may at least partly be due to his probably unrealistic assumption of flexible furnace equipment, making possible a costless switch from oil to internally generated fuels like coke-oven and blast-furnace gas.

Structural change and energy use in the future

One of the reasons for measuring substitution possibilities is the need to gauge the future energy requirements in Swedish manufacturing. To discern future trends in industrial energy demand, one must study the dynamics of industrial investment and growth, analyzing the effects on specific energy use and tracing the changing branch composition.

That is the aim of the study by Ysander-Nordström making up Part III of this volume. The authors try to accomplish it by simulations on a dynamic macro model of the Swedish economy, incorporating a vintage approach to industrial capital, and a relatively detailed description of the different mechanisms for energy substitution. Many of these mechanisms have been modeled using the estimates of price elasticities derived by Dargay and Jansson.

Some of the most interesting results of this study are summarized in Table 3. For each form of energy the change of total use in manufacuring during the period 1980-2000 is recounted as the change in production volume multiplied first by the change in energy coefficients (structure being held constant) and then by the change in energy use structure (energy coefficients being kept constant).

We see that for total energy the "structural" effect is of the same magnitude as the change in specific energy usage. The same is true for total fuels and for electricity. The change in specific usage varies, however, between the fuels as to both sign and magnitude. While specific usage is halved in the case of oil it increases almost by half for coal and by some thirty percent for domestic fuels.

Some rather dramatic changes in the energy system are moreover expected to occur during the period. The closing down of nuclear reactors, beginning in the 90s, will mean an end to the "electricity glut" and will imply higher electricity prices, which can be expected to cause a certain slow down both of mechanization and electrification and of oil saving in manufacturing.

Table 3 **Factors determining change of energy use in manufacturing, 1980-2000**

	Relative change 2000/1980 in:			
	total production volume •	specific energy usage[a] •	use structure[b] =	energy use
Oil	1.65	0.52	0.93	0.79
Coal	1.65	1.46	0.86	2.07
Domestic fuel	1.65	1.29	0.83	1.77
Total fuel	1.65	0.90	0.87	1.29
Electricity	1.65	0.92	0.91	1.38
Total energy	1.65	0.90	0.88	1.31

[a] Weighted average of specific energy usage with 1980 production shares as weights.

[b] Weighted average of production shares with specific energy usage in 2000 as weights.

One way of summarizing the findings reported in Table 3 would be to note that half the total energy savings up till the turn of the century would be realized even if the average energy-efficiency remained unchanged within each manufacturing branch. Having worked our way through the maze of econometric estimates of substitution possibilities within the manufacturing branches, we thus come back to the conclusion already derived intuitively from postwar experience. Energy saving and energy economy are not just matters of public and private energy policy. They depend as much on economic development in general and on the rate of industrial restructuring in particular.

REFERENCES

Berndt, E.R. and B.C. Field (eds.), 1981, Modeling and Measuring Natural Resource Substitution. The Mit Press, Cambridge, Mass.

Berndt, E.R., C.J. Morrison and G.C. Watkins, 1981, "Dynamic Models of Energy Demand: An Assessment and Comparison"; in Berndt-Field (eds.). (Cf above).

Field, B.C. and P.G. Allen, 1981, "A General Measure for Output - Variable Input Demand Elasticities", American Journal of Agricultural Economics, 63.

Field, B.C. and E.R. Berndt, 1981, "An Introductory Review of Research on the Economics of Natural Resource Substitution", in Berndt-Field (eds.) (Cf above).

Kopp, R.J. and V.K. Smith, 1981, "Measuring the Prospects for Resource Substitution under Input and Technology Aggregations", in Berndt-Field (eds.). (Cf above).

McFadden, D., 1963, "Constant Elasticity of Substitution Production Functions", Review of Economic Studies, 30.

Parks, R.W., 1971, "Responsiveness of Factor Utilization in Swedish Manufacturing 1870-1950", Review of Economics and Statistics, 53.

Robinson, J., 1933, Economics of Imperfect Competition, London.