WPS 0920

Policy Research

WORKING PAPERS

Public Economics

Country Economics Department
The World Bank
June 1992
WPS 920

Economic Incentives and Point Source Emissions

Choice of Modeling Platform

Raymond J. Kopp

A modeling platform for quantifying information about alternative regulations for controlling point source emission — that is, for choosing among incentive-based, market-oriented, and command-and-control regulatory policies that might adopt rapidly developing and transitional economies.

Policy Research

WORKING PAPERS

Public Economics

WPS 920

This paper — a product of the Public Economics Division, Country Economics Department — is part of a larger effort in the department to assess alternative instruments for controlling pollution in developing countries. This research was funded by the World Bank's Research Support Budget under research project "Pollution and the Choice of Policy Instruments (RPO 678-48)." Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington DC 20433. Please contact Carlina Jones, room N10-063, extension 37754 (June 1992, 46 pages).

Among economists, it is generally taken as given that environmental regulation based on the use of economic incentives and the competitive market will be more efficient (attain the same environmental goal at lower social cost) than traditional command-and-control approaches represented by environmental regulations in the United States and to a lesser extent the European Community. The sulfur trading provision of the recent amendments to the U.S. Clean Air Act suggests that some attention is being paid to economists' claims, but it would be unrealistic to assume that the command-and-control structure of U.S. environmental policy will be displaced by economic incentives any time soon.

While incentive- and market-based regulation may have penetrated very little in the United States, their potential use for environmental improvement in the rapidly developing countries of the Pacific Rim and the transitional economies of Eastern Europe and the Commonwealth may be great. These countries have no history of command-and-control regulation of the environment and the relative efficiency of regulation may be important to their capital-constrained economies.

The relative efficiency of incentive- and market-based regulation depends on the specific nature of the activity to be regulated. In some cases it may be greatly superior to command-and-control regulation; in others, only marginally so. Moreover, to be effective, incentive-based regulation may require greater effort for monitoring and enforcement — effort that offsets gains in efficiency. So the choice of regulatory approach is not always clear cut, and some analytical means of distinguishing between options is required.

Kopp reviews and recommends a modeiing platform for analyzing regulations designed to control point source emissions. The platform is intended to provide quantitative information on the efficiency of several alternative incentive-based, market-oriented, and command-and-control regulatory policies might adopt rapidly developing and transitional economies. In addition to discussing the model, Kopp pays considerable attention to such a model's informational requirements and to techniques for dealing with inadequate data.

Economic Incentives and Point Source Emissions:
Choice of Modeling Platform

by
Raymond J. Kopp

Table of Contents

I.	Introduction and Statement of the Problem		1
II.	Complexities in Modeling Point Sources		4
	II.1	Multiplicity of Inputs and Outputs	4
	II.2	Jointness of Production	5
	II.3	Input and Output Substitutability and Complementarity	6
	II.4	Brownfield and Greenfield Plant Designs	7
III.	A Model of Input Demand, Output Supply, and Residual Generation		9
	III.1	The Restricted Profit Function	9
	III.2	Modeling Regulatory Policies with the Restricted Profit Function	13
	III.3	Second Derivative Effects: Substitution in the Restricted Profit Function	15
	III.4	Evaluating the Economic Impact of Environmental Regulations on Point Sources Using the Restricted Profit Function	21
IV.	Data Availability Issues		22
	IV.1	Evaluating the Adequacy of Data	22
	IV.2	An Engineering Example	28
	IV.3	Aggregation Over Productive Units	32
	IV.4	Summary and Options for Dealing with Data Inadequacies	39
References			45

Economic Incentives and Point Source Emissions: Choice of Modeling Platform[1]

Raymond J. Kopp

I. INTRODUCTION AND STATEMENT OF THE PROBLEM

The purpose of this study is to identify the best modeling platform for the analysis of alternative environmental policy instruments designed to reduce the emission of pollutants from point sources, most notably, central power generating stations and manufacturing facilities.[2] The primary analysis of concern is a cost-effectiveness investigation of the policy; where for the most part, the cost of compliance is a multidimensional variable that includes the private costs incurred by the owners of the facility, measures of the change in the cost of providing the facility's product, and estimates of the change in facility capacity factors (i.e., measures of unemployed resources).[3] The range of pollutants under consideration include the usual menu of air- and waterborne emissions as well as solid and liquid wastes finding their way to landfills and other such disposal options (i.e., incineration). The range of policies considered

---

[1] The helpful comments of Shantayanan Devarajan, Gunnar Eskeland, Emmanuel Jimenez, Alan J. Krupnick, and Lili Liu are gratefully acknowledged; however, all errors and omissions are the sole responsibility of the author.

[2] The term modeling platform is not rigorously defined, but it is meant to be suggestive of the underlying theoretical and empirical structure of a model. For example, two platforms that are candidates for the types of analyses we want to consider are the engineering process model platform and the econometric platform. These types of models differ in terms of both concept and empirical content. The econometric model postulates optimizing behavior on the part of economic agents and analyzes policies by simulation, while the engineering process model says nothing about actual behavior and examines policies by explicit optimization of a specified objective. The econometric model relies on past behavioral observation and statistical estimation to parameterize the model, while the process model relies on generally accepted physical and engineering relations. Platforms can also be thought of as having a nested hierarchy. For example, under the econometric platform heading one might find production, cost, or profit function models. Each of these shares the same paradigm underlying all econometric models, but differs with regard to assumptions concerning the objectives of the economic agents.

[3] Cost-effectiveness implies that each of two competing policies produces the same environmental results but a differing cost. However, we will not constrain ourselves to this type of analysis and will certainly want to consider platforms that permit the evaluation of cost and emission reductions simultaneously.

include: (1) tariffs on the emission of pollutants, (2) tariffs and subsidies applied to the inputs or the outputs of the point source activities under consideration, (3) limits on the pollutant emissions themselves, and (4) directives regarding the installation of particular equipment and/or the alteration of process activities.[4]

It is important to note at the outset the difficulty of the task we are asking candidate model platforms to perform. We are concerned with policies that can alter every price the firm faces, alter the firm's ability to dispose of waste, and require the firm to undertake costly investments in abatement and treatment activities. Compare this analysis to the types of analyses that were undertaken during the heydays of the energy crisis when the modeling of manufacturing activity reached its zenith both in terms of effort levels and intellectual contributions. In these analyses the only variables of concern were energy prices, and the big issue being addressed was the relationships between energy and capital.[5] The analyses we wish to consider will build on this early intellectual base, but must broaden the base to encompass an enormous amount of additional complexity.

Everyone who attempts to construct an empirical model for the analysis of public policies faces a common dilemma. It is invariably the case that the degree of model sophistication, complication and detail called for by the nature of the policy to be analyzed, exceeds the data available for the empirical parameterization of the model. Among economists, the majority response to this problem has been to simplify and aggregate until the model has few enough parameters that it can be estimated with the scant available data.

---

[4] We note that the policies identified in (1) are broad enough to include tradable permits and deposit refund systems.

[5] We are obviously oversimplifying the energy demand research that was undertaken during the 1970s and early 1980s, but the fact remains that in terms of complexity, the task we have before us is at least an order of magnitude greater.

There are two explanations for this majority response. The first view assumes that data suitable for model parameterization must come from published government time series (e.g., the national income and product accounts) or cross-sectional surveys (e.g., household consumption surveys and plant level manufacturing surveys). It is hard to pinpoint the origin of this predilection among economists, but it indicates a strong preference for data that come from past historical experience and actual observation. Since such data are very expensive to collect, they are always in short supply and rarely in the form suitable for model estimation. The second reason is somewhat more sophisticated than the first and does not rely on simple preference. It implicitly assumes that the errors one makes in the evaluation of a policy are greater if one relies on low confidence data than the errors one would make by simplifying the policy model.[6] Unfortunately, few empirical studies exist to support this view.[7] However, in the early 1980s Kopp and Smith initiated a series of such investigations of models for manufacturing behavior using an experimental technique.[8] These experiments focused on input aggregation, constraints on the firm (e.g., pollution control regulations) and technological progress. Generally speaking, our results suggested that highly simplified models not only produced noninformative (statistically imprecise) estimates of important physical relationships, but often produced statistically significant estimates of production features that were simply false. This experience has led me to adopt the minority view when it comes to model and data tradeoffs. Simply stated, it is my opinion that it is better to use low

---

[6] Low confidence data can be defined by an example. Suppose I wanted to know the mean height of all students in a particular class. I could bring each student up to the front of the class and measure them and then calculate the mean. My confidence in the mean would be quite high and moderated only by my confidence in the mechanics of my height measurement method to which I might assign an error of plus or minus .25 inches. On the other hand, I could simply look at the class and guess at the mean. My confidence in this case might be low and I may believe that I could be off by as much as 6 inches. The only difference between these data is the confidence (error bounds) that I would assign to the estimates.

[7] If the information existed to estimate the full model, no one would care to estimate some lesser version.

[8] See Kopp and Smith (1980a, 1980b, 1982, 1983, 1984).

confidence data than a model that has been overly simplified through the use of restrictive assumptions (e.g., those necessary to support aggregation).

The remainder of this paper is organized as follows. In section II we discuss a nonexhaustive set of issues associated with the modeling of point source emissions and policies for their control. Given the state of our theoretical and empirical knowledge, some of these issues will be addressed by the modeling platforms we consider, but some of them will simply outstrip our current intellectual capital and are topics for research. Section III introduces the restricted profit function as a modeling platform, and discusses the usefulness of the restricted profit function in the analysis of environmental policy. In the closing section we examine informational requirements. The section proposes criteria for evaluating the adequacy of available data sets, discusses types of data commonly found, and considers options for dealing with data inadequacies.

II. COMPLEXITIES IN MODELING POINT SOURCES

II.1 Multiplicity of Inputs and Outputs

We postulate that the point sources we wish to investigate are plants engaged in anthropogenic production of marketed goods or services. Each point source is assumed to produce at least one such good or service and most likely produces several. We further postulate that each point source employs two or more inputs and that the input list may include both anthropogenic and biogenic inputs. Finally, each point source produces a set of residual products.[9]

---

[9] This paper employs specific definitions for some commonly used and not so commonly used terms. These definitions are provided below.

<u>Production</u> - An activity that serves to transform energy and matter into alternative forms of energy and matter.

(Footnote 9 continued on following page)

II.2 Jointness of Production

At the most general level, the production of goods, services, and residuals is considered to be fully joint. The realization among economists that the first law of thermodynamics intimately links inputs, outputs and residuals in a rather complex and most often joint manner is due to Ayres and Kneese (1969). Technically speaking, jointness means that: (1) the marginal cost of producing any particular good or service is not independent of the level of production of all other goods, services, or residual products, and (2) the marginal rate of transformation (MRT) between any two goods, services, or residuals, or the MRT between a good or service and a residual, is not independent of the levels of all other goods, services, or residuals. In nontechnical terms this means that the cost of producing an additional unit of any particular good will vary with the level of production of all other goods and the amounts of residuals produced. Joint production is the most general case of the neoclassical

---

Anthropogenic Production - A production activity that is under the control of humans. e.g., electricity production, textile manufacture, etc.

Plant - The site at which anthropogenic production takes place, a term to characterize the totality of integrated production activities.

Biogenic Production - A production activity that occurs naturally, e.g., the production of salt water commercial and game fish by off-shore ocean fisheries, or the purification of surface water supplies performed by wetlands.

Goods and Services - Outputs of anthropogenic or biogenic production activities that generate positive utility for individuals when individuals are provided access to them.

Market Goods and Services - Goods and services that are exchanged in organized markets. The feature that enables them to be "marketed" is the ability to provide individuals with exclusive access to the good or service. Thus the marketed good or service can only generate utility for those individuals granted access. Such goods are often termed private goods and are generally provided by anthropogenic production.

Input - A form of matter or energy that is transformed during the production activity or that provides services to facilitate the transformation.

Anthropogenic Inputs - Inputs of human labor service, inputs of producible market goods and services, and inputs of producible capital service.

Biogenic Inputs - Inputs of nonmarket goods and service, inputs of primary energy forms (either market or nonmarket, e.g., coal and solar energy), inputs of animate and inanimate natural resources (e.g., plants and animals provided by biogenic production, and elemental and mineral resources).

Residual Products - Outputs from anthropogenic production activities that do not enhance individual utility are termed residual products. Residual products can also be defined in a materials balance sense as the residual difference between the matter and energy inputs that are input to the production activities and the matter and energy that are embodied in the produced marketed goods and services. Whether items are classified as residual products, inputs or marketed goods and services depends upon market conditions. For example, at a given set of market prices, spent sulfuric acid used in metal cleaning is a residual product, while at higher prices for sulfuric acid the spent acid is recycled and becomes an input or is sold to others and becomes a market good.

production model. Examples of nonjoint production are simply special cases of the general joint production characterization described above.

II.3 Input and Output Substitutability and Complementarity

Substitution among inputs and outputs is the normal assumption in simple production models and indeed the only possibility when inputs and outputs are limited to two each. However, in more complex models, where inputs and outputs exceed two, input and output complementarity is possible. Complementarity on the input side suggests, holding the output vector constant, that a decline in the price of one input will cause its employment to rise and the employment of its complement to rise as well. Similarly, for a multi-output plant, complementarity in production between outputs means that decreasing the price of one output decreases its quantity and the quantity of its complement holding inputs constant. Since one might reasonably expect particular goods and services to be complementary to and not substitutes with, particular residual products, the general framework used to characterize anthropocentric production must be sufficiently general to encompass both input and output substitution and complementarity.

It is important to note that the concepts of substitution and complementarity discussed above are defined holding something constant, either input levels or output levels. While it is useful to think of the relationships between input and output pairs in this manner, one should be cognizant of the fact that the analysis of regulatory policies directed toward point sources must be conducted in a context in which input and output levels will adjust. When firms optimally adjust these levels in response to regulatory policies, the resulting relationship between input and output pairs may be different than the constrained relationships we normally have in mind when we think of substitution and complementarity. We will have more to say about these relationships when we consider explicit modeling approaches.

II.4 Brownfield and Greenfield Plant Designs

It is useful to distinguish between the production technology characteristics of plants that are in existence (brownfield) and those that are merely blueprints for new plants (greenfield). The degree to which input mixes may be altered, while holding the output mix constant, is severely limited in existing plants. For example, in the manufacture of petro-chemicals, the specification of the products to be produced largely determines the inputs required. Moreover, if the specified outputs determine the inputs to be employed, one can generally conclude that the residual products generated are also determined although additional processes may reduce or transform residual streams.

The important question to ask at this point is how environmental policies based on economic incentives can affect the production of residuals at brownfield plants. Take, for sake of argument, a tax on a polluting input, say high sulfur coal used to generate process steam. If there exists a higher cost, but available lower sulfur coal that may be combusted in the boilers, a sufficiently high tax on the sulfur content of the coals will induce the plant to switch. However, if the only alternatives are fuel oil or natural gas, then the plant must scrap its brownfield coal fired boilers and install fuel oil or gas boilers. In this case one can expect that the tax on coal will need to be sufficiently high to cause the coal fired boilers to be retired and to cover the added cost of raising the finance to purchase the new equipment. In that case, therefore, prices of capital equipment will play a role, in addition to full prices. In a technical language, the cost function will not be separable between capital and energy in this case. The point to note is the dichotomy between a simple, reversible input substitution, driven by changing input prices as compared to a combined scrappage and investment decision that has intertemporal consequences due to its irreversible nature.

An alternative example of the effect of an input tax on a brownfield plant would be a tax on a material input, for example, sulfuric acid used to clean scale from sheet steel prior to further processing. There are substitutes for sulfuric acid, but few are as effective as the acid. The result of a switch to another cleansing agent will be a poorer cleaning job that will degrade the quality of the final product. In this case the plant will make decisions based on competitive market conditions, whether to continue using the acid and pay the tax, use a substitute for the acid and suffer lost sales, undertake a capital investment to recycle waste acid, or discontinue using the acid altogether and discontinue producing the product dependent upon acid cleaning. Each of these decisions will have a different impact on the amounts of residuals generated, the quality of products produced, and the competitive position of the plant. The costs of the chosen response, whether borne by the company or its clients, or others, should be included in the measure of the social cost of the policy change.

Continuing for the moment with the input tax example, there is a second pathway through which substitution and changes in residual production can come about in brownfield plants. Let us assume that there is more than one brownfield plant producing any particular set of products, and that the production technologies vary by plant due perhaps to capital vintage, access to particular input markets, etc. Under such assumptions, an input tax, forcing the plant decisions noted in the paragraph above, can affect the plant's competitive position and lead to a redistribution of output among the plants. This redistribution of output will lead to altered aggregate residual production as well as the spatial distribution of these residuals.

During the design of a new plant all things are variable, i.e., inputs, outputs and residuals. Since all decisions are investment decisions expectations about the future economic environment of the plant will be very important. During this design stage, the post-construction flexibility of the plant to modify its input, output and residual mix will be determined. Obviously, environmental policies based on economic incentives will be most effective in greenfield plant designs and their effectiveness will be more pronounced the more confidence the plant designers have in the stability of the policies over time.

III. A MODEL OF INPUT DEMAND, OUTPUT SUPPLY AND RESIDUAL GENERATION

Given the multiple input, output and residual production characteristics of most point sources, the assumed jointness of production that we wish to maintain and the range of regulatory policies we intend to consider, the neoclassical profit function appears to be the best conceptual platform for our analysis.

Since our goal is to analyze policies affecting existing plants, and since we might operationalize our conceptual model using pooled cross section - time series data and econometric estimation techniques, we shall restrict our attention to: (1) the class of profit functions termed "restricted" profit functions indicating that some subset of the inputs are quasi-fixed (e.g., capital), and (2) functional specifications lending themselves to econometric estimation.

III.1 The Restricted Profit Function

We shall be assuming that plant managers have under their control a production technology that employs multiple variable inputs to produce multiple marketed outputs and residual products using a fixed capital stock and perhaps other fixed inputs, subject to overall constant returns to scale.[10] The inputs are assumed to be available at fixed prices and with constant quality, and the outputs sold at fixed prices in competitive markets.[11] The feasible set of variable input, quasi-fixed input, variable output and residual vectors is represented by the set T, such that $(x, y, r, q) \varepsilon T$, where $x \equiv (x_1, ..., x_n)$ is an 1xn row vector of variable inputs, $y \equiv (y_1, ..., y_m)$ is an 1xm row vector of marketed outputs, $r \equiv (r_1, ..., r_s)$ is a 1xs row vector of residual products and $q \equiv (q_1, ..., q_k)$ is a 1xk row vector of quasi-fixed productive factors.

There exists a fairly standard notation simplification scheme that is often used to characterize profit functions in lieu of the scheme above. This simplification results in classifying both variable inputs and outputs as components of the x vector, where the first n components are variable inputs and are negative in sign (suggesting that increases in the price of inputs, all other things constant, reduces restricted profit) while the remaining m components are the variable outputs that are positive in sign (suggesting that increases in output prices increase restricted profits).

Given this notational scheme, we might consider incorporating residual products as well as variable inputs and marketed outputs into the x vector. In an unregulated economy, with free disposability of residuals and holding all else constant, increases or decreases in such residuals have no effect on restricted profit. However, even in an unregulated world, disposal

---

[10] In this characterization, constant returns to scale pertains to proportional increases in all factors of production. If capital and other inputs are quasi-fixed, this technology then exhibits diminishing returns to scale with respect to the variable factors of production.

[11] At this point we are not permitting a subset of the inputs, namely the biogenic inputs, to be degraded or improved in quality due to environmental regulation or the lack thereof.

is rarely free (sweeping up scraps at the end of the day costs firms something) and thus one might imagine that increases in residuals would lower restricted profit. Certainly in a regulated world, where firms using the environment for the disposal of residuals would face constraints or would be charged a fee, this would be the case. Thus, in the case of residuals fees, variable residuals look to the restricted profit function like variable inputs, even though we tend to think of them as outputs. In this case, taxes or fees on the release of residuals drive up the price of disposal into the environment. Increasing the cost of disposal induces firms to reduce their emissions by reducing the generation of residuals or incorporating treatment into their production activities in order to transform the residuals into products whose release is not taxed (e.g., treated waste water) or is taxed at a lower rate.[12]

The prices the firm faces for its residuals are interpreted as the prices for disposal to the common environment. Anything that causes these prices to rise, causes the firm to reduce its environmental disposal. One can also imagine a set of marketed products that the firm can produce by routing its residual emissions from environmental disposal through a recycling activity. *Ceteris paribus* (i.e., holding all prices and quasi-fixed inputs constant), as the price of this recycled product increases, residual emissions and environmental disposal decrease.

The plant managers are assumed to face fixed positive prices for variable inputs, outputs (including recycled products) and the environmental disposal of residuals. These positive prices are denoted by the 1x(n+m+s) row vector $p \equiv (p^x_1, \ldots, p^x_n, p^y_1, \ldots, p^y_m, p^r_1, \ldots, p^r_s)$. The managers then choose components of the expanded x vector $x \equiv (x_1, \ldots, x_n, y_1, \ldots, y_m, r_1, \ldots, r_s)$ to maximize profits subject to the quasi-fixed technology set T.

---

[12] A final technical assumption concerns the disposability of the quasi-fixed inputs q. We shall maintain that if $q' \geq q''$ and $(x,q') \varepsilon T$, then $(x,q'') \varepsilon T$.

$$\max_x \{px' : (x,q)\epsilon T\} \equiv \Pi(p,q,T). \tag{1}$$

The objective function in (1) is simply the maximization of variable revenue minus variable costs as given in (2).

$$\Sigma p_i^y y - \Sigma p_j^x x - \Sigma p_h^r r \tag{2}$$

where: i=1,n; j=1,m; h=1,s

Subject to the technology T. The restricted profit function has several mathematical properties that we note without discussion.[13] $\Pi(p,q,T)$ is linearly homogeneous and convex in p for fixed q and is nondecreasing and concave in q for fixed p.

The most useful economic properties of $\Pi(p,q,T)$ are its derivative properties. In particular, the partial derivatives of $\Pi(p,q,T)$ with respect to the vector p generates the system of variable input demand equations, the system of output supply equations, and the system of optimal (i.e., restricted profit maximization) residual generation equations.[14] These systems are displayed below.

---

[13] Readers interested in pursuing the technical details further are directed to Diewert (1985).

[14] When one allows for the adjustment of both inputs and outputs the definitions of input and output substitution and complementarity provided in section II.3 are no longer strictly maintained. However, one can still examine the signs of the respective cross Allen price elasticities, but one should be aware that permitting output to adjust in response to either input or output price changes will cause the signs of these elasticities to be more positive suggesting more complementarity.

$$x^{x*} = \partial\Pi(p^x, p^y, p^r, q, T)/\partial p^x = x^x(p^x, p^y, p^r, q, T) \quad (3)$$

$$x^{y*} = \partial\Pi(p^x, p^y, p^r, q, T)/\partial p^y = x^y(p^x, p^y, p^r, q, T)$$

$$x^{r*} = \partial\Pi(p^x, p^y, p^r, q, T)/\partial p^r = x^r(p^x, p^y, p^r, q, T)$$

III.2 Modeling Regulatory Policies with the Restricted Profit Function

Using the system (3), one can analyze three of the four regulatory policies we wish to consider by simply altering the components of the price vector p. In particular, one can analyze taxes or subsidies on: (a) residuals, (b) inputs, and (c) outputs. For example, a vector of residuals taxes equal to t^r will lead to altered levels of input use, output supply and residuals generation by increasing the price of residuals from p^r to (p^r+t^r). This tax scheme would appear in the system of input demand and output supply equations as,

$$\tilde{x}^x = x^x(p^x, p^y, p^r+t^r, q, T) \quad (4)$$

$$\tilde{x}^y = x^y(p^x, p^y, p^r+t^r, q, T)$$

$$\tilde{x}^r = x^r(p^x, p^y, p^r+t^r, q, T).$$

Similar taxes applied to the input price vector p^x or the output price vector p^y, will lead to adjustments in profit maximizing input, output and residuals.

Virtually any policy that relies on economic incentives to affect the firm's residual emission activity can be modeled with the restricted profit function. However, policies of a command and control nature, that specify emission limitations or the installation of particular control equipment, are not modeled as neatly by the restricted profit function as they might be by an engineering process model approach.

To examine command and control policies, let us first consider a policy that would force a firm to install and operate a particular set of emission abatement or residual treatment equipment. Analysis of the policy will require engineering information regarding the additional quantities of all factors of production (capital, labor, energy and purchased intermediate inputs) that will be required by the regulation. We will also require information on the amounts by which each residual's emissions will be reduced when the above quantities of productive factors are employed.

Define Δx^x as the vector of additional factors of production called for by the regulation and Δx^r as the vector of reduced emissions. One can then consider the modified system of profit maximizing input, output and residuals choices as,

$$\tilde{x}^x = x^x(p^x,\ p^y,\ p^r,\ q,\ T) + \Delta x^x, \tag{5}$$

$$\tilde{x}^y = x^y(p^x,\ p^y,\ p^r,\ q,\ T),$$

$$\tilde{x}^r = x^r(p^x,\ p^y,\ p^r,\ q,\ T) + \Delta x^r.$$

The system (5) can then be integrated back to obtain the restricted profit function under the command and control regulatory policy.[15]

$$\Pi(p,q,T) \quad + p^x \Delta x^{x\prime} + p^r \Delta x^{r\prime} \tag{6}$$

An emissions limitation policy can be modeled using techniques of simulation. Suppose the policy sets upper limits on the vector of residuals. These policy limits are denoted as Lx^r. Using the system (3), one searches over the residuals disposal price space p^r for a vector of disposal prices that would lead to a vector of residuals equal to the policy limits. It is important to note that setting the disposal prices at these shadow values will most likely lead to altered input demand and output supplies as well as emission reductions. The ability of the restricted profit function to track the effect of policies on the full range of firm decision making is one of its major advantages.

III.3 Second Derivative Effects: Substitution in the Restricted Profit Function

We have stated above that the first partial derivatives of the restricted profit function with respect to prices generates the system of input, output and residuals emission equations by which one can evaluate a range of environmental policies. While this system of first derivatives is most useful for the policy analysis itself, the matrix of second derivatives (the Hessian) provides the insight behind the adjustments identified by the first derivatives.

Let the Hessian matrix denoted by $\nabla^2_{pp}\Pi(p,q,T)$ with typical element $\partial^2\Pi(\cdot)/\partial p^2$, indicating the change in the optimal quantity of the ith variable input, variable marketed

[15] This approach to command and control regulatory policy modeling was first introduced in Hazilla and Kopp (1990) in the context of a single output, long-run cost function. In that particular case, concern was focused on the increase in long-run production cost brought about by the regulatory policy and not concerned about the generation of residuals. In contrast, our current effort is more broad and lies with residuals generation,
(Footnote 15 continued on following page)

output or variable residual emission, due to a change in the jth variable input, variable marketed output or variable residual emission price. Multiplying each element of the Hessian by the ratio of the ith quantity (i.e., variable input, variable marketed output or variable residual emission) to the jth price, generates the system of price elasticities denoted η_{ij} and presented in **table 1.**

The elasticity matrix η_{ij} contains crucial information that indicates the range of response possibilities the firm has open to it in the face of economic incentive based environmental quality regulations. The first block of the table (rows x_1^x ... x_n^x) indicates how the optimal demands for the variable factors x_1, ..., x_n will change in a response to any price in the model, including charges placed on the residual emissions. From economic theory we know that the diagonal elements (η_{ii}) will be negative, indicating that, *ceteris paribus*, a rise in the price of an input will lower its demand. Input price increases have both substitution and output effects. A rise in an input's price makes it more expensive vis-a-vis other inputs, and thus the firm will attempt to substitute away from the input whose price has risen. In addition, an input's price rise will increase the marginal cost of production, leading the firm to lower rates of profit maximizing output, implying less input demand. Thus, both the substitution and output effects serve to reduce the demand for the input whose price rises.[16]

---

output production, input employment, as well as production cost. The proper modeling of command control policies in this context depends to great extent on the exact form the policy takes. The case exemplified by (5) is quite simplified and stylistic. It assumes that: (1) the policy requires the installation of a specific piece of equipment that is captured in fixed cost, (2) variable cost of operating the equipment is zero, (3) residual emissions are reduced by a specific amount from the levels that would obtain in an uncontrolled, profit maximizing, short-run equilibrium. Under these assumptions, the firms marginal cost remains unchanged and therefore so too does its employment of factors of production (other than the required control capital) and its output.

[16] The output effect in this system is analogous to the income effect in consumer theory. For firms' input demands, however, the output effect will always go in the same direction as the substitution effect.

TABLE 1. Elasticity with respect to price of

	Input Prices		Output Prices		Residual Prices	
	x_1^x, ...,	x_n^x, ...,	x_1^y, ...,	x_m^y, ...,	x_1^r, ...,	x_s^r
x_1^x	$\eta_{1,1}^x$	$\eta_{1,n}^x$	$\eta_{1,n+1}^x$	$\eta_{1,n+m}^x$	$\eta_{1,n+m+s}^x$	$\eta_{1,n+m+s}^x$
⋮						
x_n^x	$\eta_{n,1}^x$	$\eta_{n,n}^x$	$\eta_{n,n+1}^x$	$\eta_{n,n+m}^x$	$\eta_{n,n+m+s}^x$	$\eta_{n,n+m+s}^x$
x_1^y	$\eta_{1,1}^y$	$\eta_{1,n}^y$	$\eta_{1,n+1}^y$	$\eta_{1,n+m}^y$	$\eta_{1,n+m+s}^y$	$\eta_{1,n+m+s}^y$
⋮						
x_m^y	$\eta_{m,1}^y$	$\eta_{m,n}^y$	$\eta_{m,n+1}^y$	$\eta_{m,n+m}^y$	$\eta_{m,n+m+s}^y$	$\eta_{m,n+m+s}^y$
x_1^r	$\eta_{1,1}^r$	$\eta_{1,n}^r$	$\eta_{1,n+1}^r$	$\eta_{1,n+m}^r$	$\eta_{1,n+m+s}^r$	$\eta_{1,n+m+s}^r$
⋮						
x_s^r	$\eta_{s,1}^r$	$\eta_{s,n}^r$	$\eta_{s,n+1}^r$	$\eta_{s,n+m}^r$	$\eta_{s,n+m+s}^r$	$\eta_{s,n+m+s}^r$

Next consider the cross price effects among inputs, e.g., the effect that an increase in the price of energy can have on the demand for capital. Once again, there are substitution and output effects. If capital and energy were substitutes, an increase in the price of energy would lower its demand and cause the demand for capital to rise; however, when there are more than two inputs, the relationship between the inputs can be complementary, indicating that an increase in the price of energy can actually bring forth less capital demand rather than more. Whether substitution or complementarity reigns after the firm has fully adjusted its output as well as inputs, depends on the strength of the output effect which, due to increased marginal cost, decreases input demand when input prices rise. Given an increase in energy prices, if capital and energy were complements (holding output constant), the output effect on capital would reinforce the complementary relationship and the sign of the capital-energy elasticity would be negative (indicating that an increase in the price of energy would drive down capital demand). On the other hand, if the relationship between capital and energy (output constant) were one of substitution, then the relative strength of the opposing substitution and output effects would determine whether capital demand would rise or fall with an increase in the price of energy. It is important to note that unlike analyses of factor demand conducted with a cost function, where one examines the relationship among inputs (substitution and complementarity) while output is held fixed, the profit function permits the firms to adjust output as well as input levels, and thus the relationships among inputs are more complicated than one would find in fixed output models, but also more realistic.

As one continues to read across the first block of rows of the elasticity table, one finds the elasticities that indicate how input demand varies in response to changes in the price of marketed outputs. Corner solutions aside (i.e., assuming maximum profit occurs at an interior point in input-output space), the signs of these elasticities are positive, indicating that an increase in the price of an output will induce (due to the output effect) increased demand for inputs.

The last set of elasticities in the first block of the table are the elasticities that measure the responsiveness of input demand to changes in the disposal cost of generated residuals.

Suppose one of the residuals is sulfur dioxide and the input of interest is sulfur-bearing coal used to produce process steam. If a tax is placed on the emission of sulfur from the plant and the plant has the ability to alter fuel types on the basis of their carbon content, the elasticity relating a change in the price of sulfur emissions to the quantity of high sulfur coal demanded will be negative (an increase in the price of sulfur emissions will reduce high sulfur coal use). The larger in absolute value this elasticity, the more sensitive will the demand for high sulfur coal be to taxes on sulfur emissions. Remember, that the demand for high sulfur coal will decline with a rise in sulfur emission taxes due to both substitution of low sulfur fuels for high sulfur coal and the tax induced increased marginal cost of production that will decrease profit maximizing output levels and thus factor demand. Extending the example of the sulfur tax to the fuel that substitutes for high sulfur coal, the cross price elasticity between sulfur emission charges and the demand for low sulfur-bearing fuels may have either a positive or negative sign, depending on the relative strengths of the substitution and output effects.

The second block of the elasticity table indicates how the supply of marketed output varies in response to price changes. With respect to input price changes, the sign of these elasticities will be negative (higher input prices will induce lower levels of output due to the output effects). With respect to an output's own price, the sign will be positive (increased prices bring forth increased supply), while with respect to other output price changes the sign can be either positive or negative, depending on the relative strength of the output substitution effects (outputs can be substitutes or complements, holding inputs constant) and the output expansion effect (jointness in production can cause the increase in inputs devoted to the production of one output to reduce the marginal cost of producing a joint output).

The last section of the second block describes how increased residual disposal cost can affect the supply of marketed outputs. In this case, increased disposal cost leads to an output effect that dampens profit maximizing output levels and one can expect to find elasticities with negative signs (but not necessarily for all, when there are several outputs). The absolute magnitude of the elasticities will depend upon the engineering relationship between the output of interest and the generation of residuals subject to the tax.

The third block of the table provides the information necessary to understand how an input, output, or residual tax affects the emission of residuals. Taxes on inputs, outputs and residuals all cause the marginal cost of production to rise, dampening output and the generation of residuals. Thus, the output effect of any tax on residuals generation is generally negative.[17] While one might be inclined to generalize from this statement that the signs of the residual elasticities with respect to all environmental taxes is negative, it once again depends on the relative strength of the output effect vis-a-vis the very complex engineering relationships among residuals, input use, and output production. For example, output effects aside, a tax on sulfur emissions might also decrease nitrogen oxide emissions if the production technology is such that these residuals are complements. In contrast, if the firm responds to the increased cost of sulfur dioxide disposal by increasing capital use (installing sulfur scrubbers), sulfur emissions may be reduced but solid waste (scrubber sludge) might increase. If this substitution effect overpowers the output effect, the sign of the elasticity between the tax on sulfur and the emission of solid waste may be positive.

---

[17] An exception to this general rule would be a fully nonjoint multiple output production technology, where one output is produced using a "dirty" input giving rise to a residual, while the other outputs are produced using the same capital equipment, but without use of the dirty input. If a tax were placed on the residual caused by the dirty input, the output of the corresponding product would be reduced thus freeing fixed capital. This additional capital could push the short- run marginal cost curve for the other outputs out to the right, causing their production to expand.

III.4 Evaluating the Economic Impact of Environmental Regulations on Point Sources Using the Restricted Profit Function

The restricted profit function is a very useful tool for the comparative analysis of alternative regulations because it provides sectoral estimates of changes in output, capacity utilization (both capital and labor) energy demand, costs of production, and welfare theoretic measures of the loss in producer surplus. The factor demand and output supply system (3) provides all the information necessary to evaluate sectoral employment, output and cost effects given any imposition of environmental taxes, fees, or subsidies. Searching for a price vector p^r that would generate a residuals emission vector Lx^r, by simulation of the (3), again provides the information necessary to evaluate sectoral employment, output and cost effects for emission limits; while system (4) is used to examine command and control technology regulations.

For each of the regulatory approaches discussed above, one may integrate the systems of factor demand, output supply, and residual emission equations back to form the restricted profit function. Changes in the levels of restricted profit are true welfare theoretic measures of changes in producer surplus and may be combined with changes in household well-being to measure the relative costs of any array of policy alternatives. For example, the change in producer surplus, given a change in residual emissions taxes from p^{r0} to p^{r1} is expressed as,

$$\Delta PS = \Pi(p^x, p^y, p^{r1}, q, T) - \Pi(p^x, p^y, p^{r0}, q, T). \qquad (7)$$

The producer surplus expression (7) is defined for an individual firm. Summing over all firms subject to the regulation provides a measure of the direct cost of the regulation, but neglects the secondary producer surplus changes (general equilibrium effects) that would be brought about by the inter-industry transfer of goods and services from regulated to unregulated industries if the output prices of the regulated firms rise.

IV. DATA AVAILABILITY ISSUES

IV.1 Evaluating the Adequacy of Data

The purpose of this closing section is to investigate items needed to operationalize the restricted profit function model, in particular, the informational requirements of the modeling approach. As a stepping off point for this discussion we should recognize that an econometric model is constructed for the purpose of "mapping or projecting" from one set of information to another that is more useful in the investigation and understanding of a particular economic issue.

For example, if we wish to consider the responsiveness of energy demand to a tax on energy we may wish to map from observations on plant-level input choices and their associated prices to matrices of input substitution elasticities. The reliability of the estimated elasticities (i.e., how well they will approximate actual plant-level response) will depend on how well the theoretical model used to characterize the elasticities matches the actual characteristics of the plant. If for example, we adopt a model that assumes all inputs can be costlessly adjusted when in fact energy using capital is costly to adjust, the estimated elasticities may be unreliable. Similarly, if we adopt a model that assumes structures and equipment may be aggregated into a homogenous index of capital (i.e. structures and equipment form a weakly separable subset of the factor inputs) and find out this assumption is untrue, the elasticities estimates will again be unreliable. When one is forced to assume costless input adjustment because data on actual costs of adjustment are not available, or one assumes weak separability because disaggregate structures and equipment data are unavailable, we refer to the available data as being "inadequate."

It is difficult to investigate quantitatively data inadequacy, and therefore, difficult to state with any precision the effect of data inadequacy on model results. Some investigative experimental work was conducted in the 1980s, but even under experimental conditions, much of the analysis was qualitative rather than strictly quantitative.[18] Regardless of the quantitative or qualitative character, investigation of data inadequacy begins with a description of the economic issues to be addressed and the modeling framework to be employed. From this description one determines the ideal information set necessary to drive the model (i.e., the information set providing a one-to-one correspondence), and compares it to the information available.

As noted in the first section of this paper, the economic issues to be investigated are those having to do with the policies directed at the control of point source pollutants. In particular, these policies include:

A. tariffs on the emission of pollutants,

B. tariffs and subsidies applied to the inputs or the outputs of the point source activities under consideration,

C. limits on the pollutant emissions themselves, and

D. directives regarding the installation of particular equipment and/or the alteration in process activities.

To evaluate these policies one needs to understand:

1. the physical relationships that permit agents to transform energy and materials into desired products and residual products,

[18] See Kopp and Smith (1980a, 1980b, 1982, 1983, 1985), and Hazilla, Kopp, and Smith (1982).

2. the choice variables and constraints facing the agents,

3. the factors motivating the agents' choices, and

4. the nature of the markets for the energy and material inputs, and the desired residual products (including the distinctions that serve to define a residual product from a desired product).

Setting item 1 aside for the moment, items 2-4 involve assumptions. For example, in item 2 we assume the agent makes choices over variable input usage, and desired and residual product production. Agents do not choose new capital items. We also assume that the agents are subject to no constraints other than those imposed by their production technologies (other constraints may take the form of existing environmental regulation or any other form of governmental regulation). In item 3 we assume the agents are motivated solely by the desire to maximize restricted profit, while in item 4 we assume the input and output markets are characterized by perfect competition and that desired and residual products are determined on the basis of their contribution to restricted profit.

Item 1 is distinct from items 2-4 in the sense that economic theory places few prior assumptions on the technology, although we often assume long-run constant returns to scale.[19] Instead, item 1 is driven by the need to understand the behavior of agents under various regulatory schemes. Since this behavior concerns the profit maximizing choice of variable inputs, desired outputs and residuals, subject to fixed capital and the physical relationships defining the technology of production, we require information on the agents' choices of variable inputs, desired outputs, residuals, and fixed capital. Thus, when gathering our "ideal"

[19] A few additional mathematical properties are assumed, but they are of trivial importance to our task.

information set, items 2-4 are supplied by assumption and item 1 comes from direct observation.

It is important to note that through estimation econometricians are attempting to gain an understanding of the physical relationships that permit agents to transform energy and materials into desired products and residual products by examining the agents' choices. For this inference to be valid, the physical relationships must remain constant across all of the econometricians' observations on choices, or the econometrician must know the manner in which the relationships have changed.

Observable information on variable inputs, desired outputs, residuals, and fixed capital may be evaluated on the basis of several attributes. For our purposes the following attributes and questions concerning those attributes are relevant.[20]

1. **Attribute:** enumeration of the variable inputs, desired outputs, residuals, and fixed capital.

   Questions: are all the items enumerated or is some class missing (e.g., residuals), or some elements of a class missing?

2. **Attribute:** disaggregation of the variable inputs, desired outputs, residuals, and fixed capital.

   Questions: is there only aggregate energy or are there differentiated types (electricity, coal, fuel oil, etc.)? How fine is the differentiation (e.g., high sulfur versus low sulfur coal)?

3. **Attribute:** number of observations and the homogeneity of the structure generating the observations.

   Questions: are there a large number of observations on the choices of the same agents, and across these observations has the production technology remained constant? If there are observations over different agents, do they share a common technology? If there are observations over time, has the technology changed? How much variation exists in the data, across observations?

---

[20] I am abstracting from the issue of data information quality, i.e., how accurate is the information provided?

Given the three attributes above, one can evaluate any information set, at least qualitatively, and determine how well it corresponds to the dictates of a particular model and the economic questions to be asked of that model. I consider each of these attributes below in a general way and discuss three particular common forms of data: (1) aggregate time series manufacturing KLEM (K, capital; L, labor; E, energy; and M, intermediate materials) data, (2) sectoral time series KLEM data with energy disaggregation, and (3) pooled time series - cross section plant level data.

Enumeration. Clearly, we would like to have all the variable inputs, desired outputs, residuals, and fixed capital accounted for in the data, but this is improbable. The class of variables most likely to be missing will be the residuals. If all other classes exist, and one is willing to assume that residuals are weakly separable from the other classes, then one may still estimate a substitution matrix for the existing variable inputs and outputs that is meaningful for policy analyses (more on separability below). Of course, one cannot say anything meaningful about the residuals, and thus, if this information is totally missing, one would have no desire to estimate a substitution matrix for the available inputs and outputs.

However, suppose some information on the substitution among residuals were available from say a "nontraditional" source (i.e., not direct observation, but say opinion). If one could conjure up the substitution between aggregate residuals and: (1) each of the variable inputs and outputs, or (2) the aggregate residuals and aggregate variable inputs and aggregate variable outputs, one could use this information together with the substitution matrix estimated from observed data. The drawback to this approach is that one cannot examine the substitution between one of the variable inputs and one of the residuals. At best one can consider only the relationship between a variable input and the aggregate of all residuals.

A similar type of situation exists if one is missing information on energy or purchased intermediate materials. If any class is missing, the appropriateness of employing data from other sources is dependent on the validity of the weak separability assumption noted above.

Aggregation and Observations. Aggregation of different inputs, different outputs and different residuals is justified if: (1) Leontief aggregation conditions hold (all items to be aggregated are in the same quantity proportions to each other across all observations), (2) Hicksian aggregation conditions hold (the prices of all items to be aggregated are in the same proportion to each other across all observations), or (3) the items to be aggregated form a weakly separable subset of the inputs, outputs and residuals. Conditions 1 and 2 rarely hold, and if they did we would generally be forced to aggregate the items due to collinearity problems. The condition of weak separability is often invoked to justify aggregation when aggregation is desirable. However, providing evidence to support the assumption of weak separability usually requires first estimating a model without aggregation and then performing a series of statistical hypotheses tests. If you could estimate the disaggregate model to begin with, you probably would not be interested in aggregating items.

The upshot of the aggregation quandary is to keep the model as disaggregate as possible. However, detailed models mean many parameters and thus require substantial quantities of data. While the system estimation of (3) increases the degrees of freedom by the addition of estimating equations, the data requirements are still large. We are anticipating that the functional forms chosen would be "Diewert flexible," that is, capable of providing a second order approximation to an arbitrary continuously differentiable profit function. If we chose such a functional form and the sum of the variable inputs, outputs, residuals, and quasi-fixed factors were N, we would face the estimation of $1 + N + N(N+1)/2$ parameters.[21]

---

[21] See Diewert and Wales (1986).

Maintaining maximum disaggregation minimizes the need to make separability assumptions, but does so at the expense of large data requirements. While aggregation does not eliminate the need to collect data on all costs incurred and revenues generated by the plant, it does permit the categories of cost and revenue to be decreased. Since the cost and feasibility of any data collection scheme is dependent upon the number of these categories, the smaller the number, the more likely it is that data will exist or might be collected. For example, in the interest of model integrity, one may wish to maintain disaggregate capital inputs, differentiated by process; but, in reality, capital data is usually collected in two categories (equipment and structures, and if one is lucky, perhaps electric motors as well). The issue that must be addressed by the modeler is how much integrity is lost at this higher level of aggregation.

As one proceeds to greater levels of aggregation, one benefits from more widely available data, but one runs the risk of compromising the model's ability to adequately characterize the relevant features of the underlying technology.[22] Ultimately, the decision on input, output, and residual aggregation will depend upon the policy questions we expect the model to address and the range of options the technology presents to the firm as it tries to respond to the policy.

IV.2 An Engineering Example

An actual example will be helpful. Suppose we wish to examine policies designed to reduce the emissions from coking plants employed to provide coke to integrated iron and steel-making facilities. Coking is a process by which metallurgical coals of varying

---

[22] In Kopp and Smith (1980) the issue of input aggregation and the adequacy of technology characterization was examined in a controlled experimental setting. Not surprisingly, Kopp and Smith found that models that aggregate inputs to very great extent provide very poor characterizations of technologies. However, models that maintained a level of disaggregation consistent with the input categories frequently found in U.S. plant level data, generally performed quite well.

characteristics are heated in ovens to drive off some of the coal constituents (volatile matter, water, and sulfur) and produce elemental carbon. In addition to the production of coke, these ovens generate several residuals which are discharged to the environment or recycled, depending upon prices for the recycled materials.

The coking process utilizes the inputs of coal and gas to fire the oven to produce the output coke and two residual streams, airborne particulates and sulfur, and flushing liquor blowdown.[23] In addition to coal and gas, the coking process requires capital, labor, steam, electricity, process water, and sulfuric acid. The airborne particulates and sulfur are fugitive emissions that are released at the time the oven is opened and the coke "pushed." The flushing liquor blowdown comes from water sprays that are used to cool and condense the coke oven gases in a primary cooling process and from the cooling tower blowdown where the gases are further cooled in a heat exchanger system. The flushing liquor and the gas condensate from the primary cooler are run to a decanting tank where tar is settled out before the liquor and gas is returned to the cooling sprays. Flushing liquor blowdown results in about 30 gallons of liquor that are withdrawn from the process for every ton of coal. This liquid residual product can be sent to environmental discharge, treatment and/or by-product recovery depending on market prices and environmental regulation. **Figure 1** describes a by-product coking process where coke oven gas flushing liquors are input to by-product recovery systems producing ammonium sulfate, sulfur, gas to fire the coke ovens, light oil, and other chemical compounds.[24]

---

[23] Blowdown is a term used to describe the liquid that is withdrawn from these process systems and replaced with make-up liquids to keep the concentrations of various chemicals within design tolerances.

[24] Figure 1 is a stylized version of figure 3.1 in Russell and Vaughan (1976). Readers interested in a the engineering process approach to the modeling of production activity and residuals generation are directed to this work.

BY-PRODUCT COKING

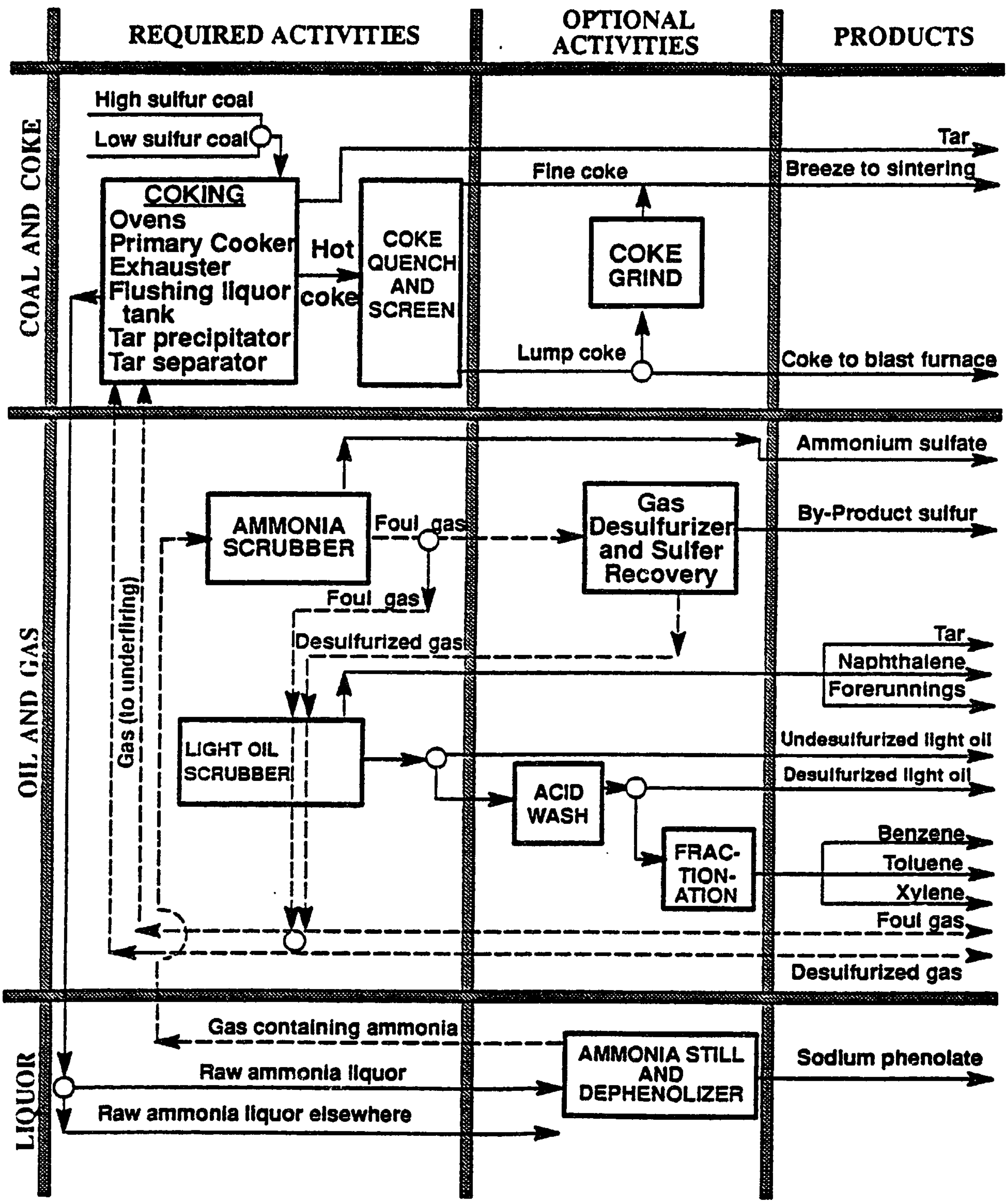

FIGURE 1

Given this very stylized description of coke production, suppose we wanted to look at policies to reduce waterborne emissions. These emissions are dependent upon the quantity of coke produced and whether and to what extent by-product recovery from the flushing liquor is practiced. We can reasonably aggregate over the types of coal used (C), labor employed (L), intermediate inputs: electricity, steam, water, sulfuric acid (INTER), and in-place capital (K). We can aggregate tar, breeze, and coke into a single output (TBC); aggregate ammonium sulfate, sulfur, naphthalene, forerunnings, light oil, benzene, toluene, xylene, foul gas, desulfurized gas, and sodium phenolate into another output (BY-PROD); and define ammonia liquor (LIQ) as the residual of interest.

The profit function would then contain input prices for L, C, INTER, and the disposal of LIQ, output prices for TBC and BY-PROD, and the quantity of K in-place. The model would then suggest that the quantity of LIQ released to the environment will decrease with taxes on environmental discharge, taxes on the final products, C and subsidies to the sale of BY-PROD. If instead of water emissions, we were concerned with releases of sulfur, we would have to disaggregate coal by sulfur content and disaggregate BY-PROD to break out sulfur as a residual. The more residuals of concern, the more disaggregate the model must become.

Unfortunately, it is not possible to state in any generic way, that is for any arbitrary process or set of residuals, the proper level of aggregation. The level of disaggregation embodied in any final model will be based on engineering knowledge of the technology, the types of questions the model is designed to address, and the constraints of available data.

<u>Observations, Homogeneity and Variation</u>. It goes without saying that more observations are preferred to fewer observations and that greater variation in the data is

preferred to less. While these are very important attributes of any information set, they do not require a great deal of further discussion.

The most serious issue in this section concerns heterogeneity across the production technologies serving to generate the data. If these observations are generated by very different production technologies, one must attempt to capture this heterogeneity either parametrically by using plant specific parameters within a single profit function model, or by sub-sampling observations into homogenous groups and estimating separate profit function models for each group. Failing to perform either a parameterization or sub-sampling, at best leaves one with estimates of an "average" production technology. Using this average technology to investigate economic issues has some shortcomings which are addressed later.

IV.3 Aggregation Over Productive Units

Throughout this paper we have assumed that the appropriate economic agent and unit for empirical analysis is the individual production organization (typically a plant). Consequently, the important data and aggregation issues investigated using the example of section IV.2 focused on the implications of input aggregation while maintaining the plant as the unit of observation. For taxonomy reasons we term this form of data aggregation "within-plant" aggregation. In this current section we wish to consider the case where the available data may be inconsistent with our maintained view of the economic agent (i.e., the individual plant as the unit of observation) and examine data sets that aggregate over plants. We term this type of data problem "across-plant" aggregation.

We consider three forms of across-plant aggregation that are consistent with generally available data sets. For each aggregation we assume that the same input and output disaggregation is available, that is, there exists three within-plant input aggregates; capital, labor, intermediate inputs; a disaggregate set of energy inputs and a single within-plant aggregate measure of output. The three categories of plant-level aggregation considered range from no aggregation to complete aggregation at the 2-digit level. These categories are provided below.

1. Plant level data observed in a pooled time series-cross section panel format (no across-plant aggregation),
2. Plants within a single 2-digit classification operating within a specific region, aggregated each year over some period,
3. All plants within a single 2-digit classification aggregated each year over some period.

To understand the implications of the above across-plant aggregation we must first identify the nature of the plant heterogeneity that one might expect.[25] First, within any 2-digit category, plants will likely produce very different products. For example, within the 2-digit category "food and kindred products" the U.S. Bureau of the Census identifies over 154 different outputs, ranging from cat food to orange juice. These output differences will likely be associated with different inputs and different residual products. Second, even if the plants are producing identical products, they will be located in different areas of the country where prices for factor inputs and taxes levied by local and regional political jurisdictions can be expected to be be different. In addition, local regulations, including but not limited to environmental regulations, may also vary. Differing prices and regulatory regimes will give rise to differences in input use and residual products. Finally, and again assuming that the plants are producing the same outputs, the plants may be using very different vintage

[25] We will assume that each plant is a profit maximizer unto itself (e.g., a single plant firm) and that it competes in competitive input and output markets.

technologies, operating very different size plants and operating these plants at very different rates of utilization.

As soon as one begins to aggregate over plants, one can expect the above problems may arise. There is no conceptual difference between the aggregation from category 1 to 2, or 2 to 3, the same problems can occur; however, the more you aggregate the more severe one can expect these problems to be.

Before we consider the implications of aggregating over plants, let us briefly review the estimation of a production model based on plant-level observations. Once one chooses a sample of plants for estimation one assumes that all the features of the plants not captured by the model specification are: (1) the same across plants, (2) unchanging with respect to the policy, or (3) uncorrelated with the features included in the model. If assumption (1) holds, the estimated model is truly representative of the production technology embodied in each plant, since each plant has the same technology, aside from differences captured by the model. When one employs this estimated model in a policy simulation, one has a high level of confidence that actual plants will respond to the policy in a manner consistent with the simulation results. If assumption (1) does not hold, but assumption (2) does, one's confidence in the simulation results depends on the degree to which the policy in question may effect any of the features held constant across plants during the estimation. If the policy does not affect these features, one can again have high confidence in the simulation results. However, if the policy influences these features, the the actual behavior of the plants will diverge from the simulation results. If only assumption (3) holds, the estimated production technology is a "composite" of the heterogeneous technologies that comprise the plant sample. If (3) holds, the composite is unaffected by plant differences not incorporated into the model and one can have confidence that the policy simulation results will generally characterize actual plant behavior.

If none of the assumptions above hold, the estimated model will still represent a composite of the heterogeneous technologies, but the policy simulation may not adequately characterize the behavior of the actual plants. In this case the policy will affect heterogeneous plant attributes not captured by the model, but which are correlated with behavior of concern (e.g., input substitution).[26] This set of circumstances will yield simulation results quite possibly in conflict with actual plant behavior. The above discussion is meant to demonstrate that even with disaggregate plant-level data there exist potential modeling problems. As we discuss below, when these problems are combined with aggregation over plants the problems are exacerbated.

Identical Plants - Different Outputs. We now turn our attention to across-plant aggregation and begin the discussion with the hypothetical case of identical plants (i.e., plants with identical production technologies) producing different outputs. It may stretch the imagination somewhat to assume that plants can be identical and yet produce different outputs, but there are examples of multiple output plants that have very high degrees of output substitution and can produce any of a number of different products.[27] In this example we assume such a flexible technology exists so that we do not confound our analysis of output differences with differences in technology.

If one were estimating an econometric production model using these data, assumption (1) would be satisfied and we would have high confidence in the simulation

[26] This situation is characterized as model misspecification and is probably more the norm than the exception in most modeling exercises.

[27] The synthetic cellulosic fiber industry (SIC 2823) is an example where the same plants can produce nylon, acrylic and polyester fibers.

results. However, in this example of across-plant aggregation we now aggregate over some or all of these plants, where the quantity of output and the factor inputs within each class (i.e., capital, labor, intermediate materials and energy forms) might simply be summed and indexes of input prices formed. In this case the across-plant aggregation can lead to unwanted problems.

Consider the outputs. Each plant produces a different output with a different market price. One accepted and common across-plant aggregation scheme is to compute the value share of each plant in the total value of the output produced by all plants and use the shares to weight the individual plant output prices and compute an aggregate output price.[28] One may then divide the output value aggregate by the aggregate price index and arrive at an aggregate output index. Unfortunately, as the price of one of the output's in the aggregate changes (e.g., rises) the aggregate output index diverges from the true value (in the case of rising output price the aggregate will fall). While in the limit (i.e., for very small price changes) their is no effect of price change on the quantity aggregate, the greater the price change, the greater the output aggregate divergence.

Under the assumption that these are identical plants, our plant level observations are realizations of differing output mixes (i.e., different expansions of the transformation frontier in output space). I would assert without proof that a consistent across-plant aggregate of these outputs exist if each plant's output expansions are linearly homogeneous, and I would guess that this might also be true under particular circumstances if the output expansions were homothetic.[29]

---

28 The econometrician rarely knows the exact process by which aggregate data was formed. For the purposes of this paper we will assume that a share weighted aggregation is generally employed.

29 Proving these results is beyond the scope of this paper and the interested reader is directed to Diewert (1976, 1988, 1989) and Blackorby, Primont and Russell (1978).

Differing outputs can also imply that plants will use different inputs to produce these differing outputs. This will be most apparent when examining the individual components of each plant's material aggregate. Differing individual components will lead to differing aggregate material price indexes and we are thus left with a situation in which we are aggregating across-plants with identical technologies, but which face differing input prices. This case is discussed in the next section.

Plants with differing outputs can likely be characterized by differing input mixes as well as differing inputs. For example, the production of one output may be more labor intensive than another. If these plants face the same input prices, across-plant input quantity aggregation results in the simple summation of like inputs, while the aggregate input price is the common price. Estimation of a production technology based on these data results in an estimated "composite" single output technology that to some extent summarizes the actual multiple output technology embodied in each of the disaggregate plants. As long as the policy in question affects all plants equally we can have confidence in the results; however, if the policy effects the plants producing one of the outputs differentially, the simulation results will diverge from actual behavior.

Identical Plants - Different Input Prices. We now turn our attention to aggregation over plants with the hypothetical case of identical plants facing different input prices. I assert without proof that if the plants share a common homothetic technology, across-plant aggregate price indexes formed as the weighted sum of the plant-level components, will lead to an estimated aggregate technology with the same features as the technology embodied in the individual plants. However, if the technology is not homothetic there is no single price index that will ensure the estimated aggregate technology mirrors the plant level technologies.[30]

[30] This assertion is based on within-plant aggregation results contained in Blackorby, Primont and Russell (1978), particularly section 3.3.

Different Plants - Same Outputs. Plants producing the same output, but that embody different technologies, pose very serious problems. Such plants will undoubtedly utilize different inputs due to their technologies and could also face differing input prices (that might have given rise to the choice of the different technologies). Estimation using this data results in a composite technology that will be sensitive to policies that affect plants differentially.

Aggregation Over Productive Units - Summary. Aggregation over productive units (plants) has two generic consequences. First, productive unit aggregation involves the aggregation of "like" inputs and outputs across-plants rather than the commonly discussed input and output aggregation that occurs when one aggregates heterogeneous entities within a plant to form within-plant aggregates. If the plants share common technologies, then this across-plant aggregation looks very much like within plant aggregation and perhaps some of the same theorems pertaining to the existence and properties of within plant aggregation apply to across plant aggregation. Second, when plants employ different input mixes or embody different technologies, across-plant aggregation results in data that give rise to estimated "composite" technologies. However, as long as the policies to be considered affect all plants equally (i.e., the plants do not respond to the policy differentially due to differences in their characteristics), the composite technology will adequately reflect aggregate plant behavior. By aggregate behavior we mean the general behavior of the industry comprised by the plants, but not the behavior of any individual plant.

One may attempt to quantify the impact of the above two aggregation problems on the performance of an econometric simulation models by appealing to modern aggregation theorems, but generally one will find these theorems provide only qualitative results that suggest the conditions under which it is inappropriate to aggregate. If one wishes to know how much of an error one is likely to make by aggregating when a particular condition obtains, one must turn to experimental methods. Unfortunately, very little experimental evidence exists that deals with across-plant aggregation.

IV.4 Summary and Options for Dealing with Data Inadequacies

This paper has presented a specific modeling framework (termed a platform) for the analysis of incentive based environmental policies designed to reduce the emissions from point sources. The chosen platform is based on an econometric restricted profit function and is shown to be capable of modeling all the incentive programs normally considered and a vast majority of command and control programs as well.

Like any econometric model, the restricted profit function platform contains parameters that must be specified and if one desires to maintain significant degrees of within- and across-plant disaggregation, the restricted profit function possess large number of parameters. If each parameter is to be econometrically estimated, the restricted profit function platform will be a voracious user of data.

In previous sections we have investigated some of the possible consequences of using aggregate data to estimate the parameters of a down sized aggregate profit function model. This investigation focused on the confidence one would have in the results of a policy simulation using a model that has been aggregated to match available data. Based on this investigation, it is our general belief that model aggregation, necessitated by data availability, can lead to serious problems and that more attention should be paid to the prospects for augmenting data rather than aggregating the model. In this closing section we wish to consider this alternative approach.

There are at least two options for dealing with inadequate data that maintain the neoclassical econometric approach outlined in this paper. The first is rather obvious and simply implies adopting an econometric model that requires less data. We are not suggesting that one abandon the "Diewert flexible" functional forms noted above in an effort to reduce the number of parameters in the model and thereby the required number of degrees of freedom. Rather, we are suggesting one might consider a model of firm decision making that economizes on the kinds of data required (i.e., various types of prices and quantities). The model we have in mind is the restricted cost function.

The restricted cost function model assumes that plant managers face fixed prices for their inputs and choose their variable inputs and residuals so as to minimize variable cost, subject to the existing technology, quasi-fixed capital, and fixed production rates for final products.

In contrast to the profit function system (3), the restricted cost function would contain only the factor demand and residual emission equations, while output supply equations would be absent. In such a framework one could examine command and control technology based regulations (see Hazilla and Kopp, 1990); emission fees, permits and taxes; and input taxes and subsidies. However, the fact that outputs are held constant reduces the flexibility of the technology characterized by the model to respond to such regulatory schemes through output adjustments.

While the restricted cost function is the next best modeling approach to the profit function, it economizes only on the need for output prices and does so by sacrificing the model's ability to adjust output. The importance of output adjustment will depend upon the technology and regulatory scheme under consideration, but *a priori* we would rather not exclude such adjustment. Thus, we propose a second approach that returns to the profit function framework, but considers the use of information that extends beyond the confines of available econometric data.

In the section of this paper describing the second derivative properties of profit functions, we emphasized the richness and usefulness of the information contained in the elasticity matrix. Indeed, given this matrix and observed levels of actual input use, output production, residual emissions, and levels of quasi-fixed capital stocks, we can derive all of the parameters of the restricted profit function (analytically or numerically depending on the functional form chosen) without any econometric estimation. One may reasonably ask, even if we could obtain current data on levels of input use, output production, residual emissions, and quasi-fixed capital stocks, where do we obtain information on the elasticity matrix?

The elasticity matrix describes how optimal levels of inputs, outputs and residuals are linked to the prices each firm faces for these items. In a sense, the matrix is a composite of engineering information that describes the technical relations between the items (how one may be transformed into another) and the economic adjustments that the firm makes to maximize profit. As microeconomists we should note that this information set (the engineering relations and the technical knowledge necessary to maximize profit) is assumed known to the firm. Perhaps this information is not known in the form of the elasticity matrix, but if our underlying information assumptions regarding neoclassical producer behavior are at all reasonable, the core of the knowledge necessary to construct the elasticity matrix is known to firms and individuals familiar with the industry.

This rather nebulous information we seek is, in fact, the basis of many process analysis models constructed for various manufacturing processes. In fact, since the level of technical specificity we require to construct the elasticity matrix is far less than that necessary to construct a process model, one can reasonably expect that the effort required to assemble the elasticity information is far less than that required to produce the process model.

In a series of papers in the early 1980s Kopp and Smith (1980a, 1980b, 1982, 1983, 1985) and Hazilla, Kopp, and Smith (1982), employed engineering information, originally developed to construct process models, to characterize elasticity matrices and other economic features of production, such as technical change. Their purpose was to compare this engineering-based information, describing fundamental economic concepts of neoclassical production, with similar information obtained from actual econometric estimation. While space does not permit a complete review of this research, suffice it to say that the results suggested what was expected *a priori*, but never before proved. Econometric estimates of engineering features of production was found to reflect reasonably well the physical relationships built into the more detailed process models if one does not stretch the underlying neoclassical assumptions too far.

Given this past experience, we are suggesting that it is not unreasonable to consider assembling engineering and plant management information in an effort to construct an elasticity matrix. This information would come from existing industry studies and perhaps models as well as the direct probing of experts (engineers, plant managers, and those knowledgeable of the industry). In fact, while the work of Hazilla, Kopp, and Smith drew from published studies, recent developments in the construction of expert systems and developments by economists and other social scientists in the field of contingent valuation, suggest that "expert encoding" may be extremely valuable.

Expanding on our suggestion to employ noneconometric data to fill out an elasticity matrix goes well beyond the scope of this paper; however, we would like to close this section by noting a few of the benefits of such an approach. One can argue that an elasticity matrix, derived from actual econometric estimation, using the finest data, is the ideal modeling tool. However, one must also argue that rarely do we have access to such ideal tools and most often we must perform analyses with the best available tools. Recognize however, that we do have at hand the platform for the ideal tool, that is, the theoretical concept of the restricted profit function, which given the analyses we wish to consider, seems quite ideal indeed.

What we lack is the quantification of the elasticity matrix. Suppose we sat down and "estimated" the elements of this matrix by simply guessing, and then performed the analyses we would like to undertake. Under such a scenario, we would assign quite large confidence intervals to the results since we believe that the probability distributions underlying our elasticity estimates (guesses) have large variances. In fact, we might convince ourselves that the only difference between this extremely low cost analysis and the obviously high cost analysis performed with the fully econometric ideal model, is the variance of the elasticity estimates employed. For the purposes of policy making, we would have more confidence in the low variance econometric estimates; but other than the variance in the estimates, we would consider the two analyses absolutely equal. For a relatively small expenditure (relative to the full econometric approach), we have an up-and-running policy evaluation model that is conceptually as solid as current economic thinking permits.

One can now think of ways that we might begin to improve our confidence in results of the "guessed" model. The first thing to recognize is that, generally, any additional information will narrow the variance of the results. Thus, moving from a guessed matrix to one informed by information found in the literature or obtained though expert encoding, will improve our confidence.

Second, not all elasticities are equally important for all policy analyses. One can imagine a Monte Carlo simulation exercise where a particular policy, say an emission charge, is repeatedly modeled by randomly choosing elasticity estimates from assumed probability distributions and designing the experiment around the variances of these distributions. The variance in the simulated policy results is calculated for each elasticity variance design point. Comparing the variance of the policy result with the variance of each elasticity will determine the sensitivity of the results to the variance of the elasticities. This information can then be employed to rank the elasticities for the purposes of research and direct research effort to those elasticities that are most crucial to the policy problem under consideration. The Monte Carlo study discussed above would provide the road map to actual data collection (either econometric or noneconometric).

We do not mean to paint the above approach as a panacea, there are surely difficulties to be encountered in implementing the approach and making it operational; but it nevertheless has desirable attributes, for example: (1) data inadequacies do not force, as is usually the case, bastardization of the theoretical framework, although we often see evidence of such in applied literature; (2) a "best concept" model is available immediately for policy analysis; (3) given any set of elasticities, counter-factual studies of complex policies can be performed, studies too complex for comparative static analysis; (4) as new data and information become available, it can be incorporated in the model; (5) should econometric data be assembled years in the future, the modeling framework requires no modification since it is an econometric framework; and (6) perhaps most important, while the data are being improved the framework is being utilized, providing insights and education to its users regarding the complexities of regulation and economic incentives.

REFERENCES

Ayres, R. U., and A. V. Kneese. 1969. "Production, Consumption and Externalities," American Economic Review, vol. 59, pp. 282-297.

Berndt, E. R., and D. O. Wood. 1975. "Technology, Prices and the Derived Demand for Energy," Review of Economics and Statistics, vol. 52, no. 2, pp. 259-268.

___. 1977. "Engineering and Econometric Interpretations of Capital-Energy Complementarity," American Economic Review, pp. 342-354.

Blackorby, C., D. Primont, and R. R. Russell. 1978. Duality, Separability and Functional Structure: Theory and Economic Applications, (New York, North - Holland).

Diewert, W. E. 1976. "Exact and Superlative Index Numbers," Journal of Econometrics, vol. 4, pp. 115-145.

___. 1985. "The Measurement of the Economic Benefits of Infrastructure Services," Discussion Paper No. 85-11, University of British Columbia, March.

___. 1988. "The Early History of Price Index Research," Discussion Paper 88-26, Department of Economics, University of British Columbia.

___. 1989. "Fisher Ideal Output, Input and Productivity Indexes Revisited," Discussion Paper 89-07, Department of Economics, University of British Columbia.

Diewert, W. E., and T. J. Wales. 1986. "Semiflexible Functional Forms," Discussion Paper No. 86-25, Department of Economics, University of British Columbia.

Fuss, M. A. 1977. "The Demand for Energy in Canadian Manufacturing," Journal of Econometrics, vol. 5, no. 1, pp. 89-116.

Hazilla, M., and R. J. Kopp. 1987. "Establishment-Level Data for Econometric, Engineering and Policy Analysis," Final Report to the Measurement Methods and Data Improvement Program, National Science Foundation, grant #SES-8500150.

___. 1990. "The Social Cost of Environmental Quality Regulations: A General Equilibrium Analysis," Journal of Political Economy, vol. 4, pp. 853-873.

Hazilla, M., R. J. Kopp, and V. Kerry Smith. 1982. "The Performance of Neoclassical-Econometric Models of Natural Resource Substitution in the Presence of Environmental

Constraints," in W. Eichhorn, R. Henn, K. Neumann, and R. W. Shephard, eds., Economic Theory of Natural Resources (Physica-Verlag).

Jorgenson, D. W., and B. M. Fraumeni. 1981. "Relative Prices and Technological Change," in E. R. Berndt, and B. Field, eds., Modeling and Measuring Natural Resource Substitution (Cambridge, Mass., MIT Press), pp. 17-47.

Kopp, R. J., and V. Kerry Smith. 1980a. "Measuring Factor Substitution with Neoclassical Models: An Experimental Evaluation," Bell Journal of Economics, vol. 11, pp. 632-655.

___. 1980b. "Input Substitution, Aggregation and Engineering Descriptions of Production Activities," Economic Letters, vol. 5, pp. 289-296.

___. 1982. "Neoclassical Measurement of Ex Ante Resource Substitution: An Experimental Evaluation," in J. R. Moroney, ed., Advances in the Economics of Energy and Resources, vol. IV (JAI Press).

___. 1983. "Neoclassical Modeling of Nonneutral Technical Change: An Experimental Appraisal," Scandinavian Journal of Economics, vol. 85, pp. 127-146.

___. 1985. "The Measurement of Nonneutral Technical Change," International Economic Review, vol. 26, pp. 135-159.

Russell, C. S., and W. J. Vaughan. 1976. Steel Production: Processes, Products and Residuals (Washington, D.C., Resources for the Future).

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS906	Bulgaria's Evolving Legal Framework for Private Sector Development	Cheryl W. Gray Peter Ianachkov	May 1992	CECSE 37188
WPS907	Institutional Reform in Emerging Securities Markets	Robert Pardy	May 1992	Z. Seguis 37664
WPS908	Tax Incentives, Market Power, and Corporate Investment: A Rational Expectations Model Applied to Pakistani and Turkish Industries	Dagmar Rajagopal Anwar Shah	May 1992	C. Jones 37669
WPS909	Parallel Markets, the Foreign Exchange Auction, and Exchange Rate Unification in Zambia	Janine Aron Ibrahim A. Elbadawi	May 1992	V. Barthelmes 39175
WPS910	Policy Issues in Financial Regulation	Dimitri Vittas	May 1992	W. Pitayatonakarn 37666
WPS911	Does Exchange Rate Volatility Hinder Export Growth? Additional Evidence	Ying Qian Panos Varangis	May 1992	D. Gustafson 33714
WPS912	Understanding the Investment Cycle in Adjustment Programs: Evidence from Reforming Economies	Andrés Solimano	May 1992	E. Khine 39361
WPS913	The Women's Development Program in Rajasthan: A Case Study in Group Formation for Women's Development	Maitreyi Das	May 1992	L. Bennett 82772
WPS914	Health Personnel Development in Sub-Saharan Africa	J. Patrick Vaughan	May 1992	O. Nadora 31091
WPS915	Trade Policy and Exchange Rate Issues in the Former Soviet Union	W. Max Corden	May 1992	CECTP 37947
WPS916	Measuring the Risk of Default in Six Highly Indebted Countries	Marc Chesney Jacques Morisset	June 1992	S. King-Watson 31047
WPS917	Creditor Country Regulations and Commercial Bank Lending to Developing Countries	Asli Demirgüç-Kunt	June 1992	K. Waelti 37664
WPS918	Tax Evasion and Tax Reform in a Low-Income Economy: General Equilibrium Estimates for Madagascar	Jaime de Melo David Roland-Holst Mona Haddad	June 1992	D. Ballantyne 37947
WPS919	Fiscal and Quasi-Fiscal Deficits, Nominal and Real: Measurement and Policy Issues	Roberto de Rezende Rocha Fernando Saldanha	June 1992	L. Ly 37352

Policy Research Work'ng Paper Series

	Title	Author	Date	Contact for paper
WPS920	Economic Incentives and Point Source Emissions: Choice of Modeling Platform	Raymond J. Kopp	June 1992	C. Jones 37754