SOME ISSUES THAT ARISE IN ESTIMATING DEMAND FUNCTIONS

- **WHAT TYPE OF DATA: CROSS-SECTION OR TIME SERIES?**

- **WHAT TIME FRAME?** Total annual demand, or average monthly demand, or average daily demand, or peak daily demand, or peak hourly demand etc

Which you should use depends on the purpose of the analysis – eg if you are planning capacity for a water purification plant or sewage treatment plant, you might want to know peak hourly/daily demand. If you are planning a reservoir, you might want to know total annual (or seasonal) demand.

- **WHAT LEVEL OF AGGREGATION FOR THE UNITS OF OBSERVATION?** Do you want data from individual factories, individual companies, individual industries, or individual counties, states or census regions? Babin shows why you need to break things down at least by industry. Othe types of disaggregation may make a difference, also.

- **WHAT TYPE OF DISAGGREGATION AMONG WATER USE?** One classification is intake, amount treated prior to use, amount recirculated (gross use) amount treated prior to discharge, and amount discharged. Do you just use total intake, say, or do you break it down by *end use*? - Examples of end use are: cooling, processing, steam generation, and sanitary use. Renzetti finds that intake and discharge are complements, while intake and recirculation are substitutes.

- **HOW MEASURE THE PRICE OF WATER?** For publicly supplied firms, there is the issue of sewer prices along with water supply prices. Moreover, there may not be *a single price* – ther may be several prices (connection fee, volumetric charge etc). Even if there is just a volumetric charge, it may not be constant - there can be *increasing or decreasing block prices* For self-supplied firms, there is not a price of water *per se*. In effect, water is produced commodity for the firm; the prices it faces are the price of electricity, labor, chemicals, etc.

WHAT DO YOU ESTIMATE?

In this context, one can estimate

- a single input demand function (this could be a conditional or unconditional input demand function; it could be a short or long run function)

- a *system* of input demand functions

These could be estimated alone or together with:

- the output supply function

- the profit function

- the cost function

- the production function

There are at least three aspects to this choice:

1. What you want to know

2. What data are available

3. What can reliably be estimated from the data in the light of potential econometric complications, especially the issue of exogeneity

HOW DO YOU ESTIMATE IT?

APPROACHES TO EMPIRICAL MODELING OF INPUT DEMAND

The literature contains two approaches for analyzing the demand for inputs in manufacturing and agriculture:

1) The mathematical programming (optimization) approach. Calibrate (or statistically estimate) a production function. Use mathematical optimization to derive the input demands/output supply associated with the production function.

2) Statistical estimation of input demand function. Use actual data on input use and estimate the best fitting demand function using regression or a related statistical technique.

Both approaches have been used with both industrial and agricultural input demand for water; the programming approach is used more frequently with agricultural demand.

Follows Grebenstein & Field, who used state-by-state Census data to estimate a demand for water for US manufacturing industry as a whole. They disaggregate into separate water using sectors.

They use the translog cost function?

> The four-input model used in this investigation includes capital services (K), water (W), production employees (P) and nonproduction employees (N) as factors of production. Thus the translog cost function, for a nonhomothetic production structure, can be expressed as
>
> $$(1)\quad \ln C = \ln \beta_0 + \beta_q \ln Q + \Sigma\beta_i \ln P_i + \tfrac{1}{2}\delta_{qq}(\ln Q)^2 + \Sigma\delta_{iq} \ln P_i \ln Q + \tfrac{1}{2}\Sigma\Sigma\delta_{ij} \ln P_i \ln P_j,$$
>
> where $i,j = k,p,n,w$, C is total cost, Q is output, P_i are input prices, and the β_i and δ_{ij} are the parameters to be estimated.

Why translog?

Because it is flexible.

Why not a production function? Or a profit function?

No useful quantity measure of output is available from the Cenusus

No useful price of output is available either.

Why? Because of the high level of aggregation associated with the use of 2 digit sectors This would be much less of a problem is used plant level data

They assume production is linearly homogenous

What does this imply?

Constant returns to scale.

Constant average = marginal cost

No well-defined supply curve

Therefore, not meaningful to use profit function or unconditional demand function – only the conditional demand function

They impose on the translog cost function the restrictions of symmetry and homogenity.

We therefore impose on our model the restrictions of symmetry ($\delta_{ij} = \delta_{ji}$, $i \neq j$, $i,j = k,p,n,w$) and linear homogeneity ($\Sigma\beta_i = 1$, $\beta_q = 1$, $\delta_{qq} = 0$; $\Sigma_j\delta_{ij} = 0$, $\delta_{iq} = 0$, $i = k,p,n,w$). The function can now be expressed as a translog unit cost function:

$$\text{(2)} \quad \ln C/Q = \ln \beta_0 + \Sigma\beta_i \ln P_i + \tfrac{1}{2}\Sigma_i\Sigma_j\delta_{ij} \ln P_i \ln P_j.$$

They estimate the conditional ddemand functions in share form.

The estimating equations are share equations in which the dependent variable (S_i) is the ratio of cost of factor i to total cost for each industry and state. Because of the identity $\Sigma_i S_i = 1$, only three of the share equations are independent, and hence we estimate equation (2) with the above restrictions through the three share equations:

$$S_i = \beta_i + \Sigma_j \delta_{ij} \ln(P_j/P_w) + u_i, \quad (3)$$

where $i,j = k,p,n$.

.... The dependent variables (S_k, S_p, and S_n) were defined as the ratios of the total cost of the particular input in a given industry and state to the total cost for the industry in that state. In turn, the denominator was calculated for each industry and state as the sum of total payroll plus the capital service price multiplied by quantity of capital plus water price multiplied by water utilized.

Why?

Because no meaningful quantity units are available for characterizing input quantity at this level of aggregation

Using expenditure on inputs finesses this.

They don't actually have data on the price of water from the Census of Manufacturing.

So they extrapolate from the American Waterworks Assoc survey data

This could be problematical

The price of water in each state was calculated from data published by the American Water Works Association in 1972. Data on quantities and revenues for various water utilities permitted the computation of an average price within each state. These price data are an improvement over the Grebenstein and Field data, which were regional average prices. Like Grebenstein and Field, we were unable to develop a marginal price. Data limitations also forced us to assume that the price levied by public and private water supply agencies was a useful shadow price for water produced directly by firms. If this is untrue, our estimations may be biased.

Some insight into changes in factor utilization likely to result from relative price changes is provided by the estimated price elasticities of demand, table 2. Estimated own-price elasticities of demand generally carry the expected negative sign. They are positive for SIC groups 35 and 36, implying that the estimated cost functions are not concave in input prices over the data range. However, the standard errors are quite large in these cases, so that the elasticities are not significantly different from zero.

The differences among sectors also are striking. For example, the estimated relationship between capital and water is complementary in three sectors and in the pooled model. However, these two inputs appear as substitutes in sectors 20 (food and kindred products), 34 (fabricated metal products), and 36 (electric and electronic equipment). While the estimated relationship between capital and production workers shows substitutability for all sectors, their magnitudes and standard errors vary considerably. Accordingly, the appropriateness of pooling these sectors was tested. T[illegible] The resulting test [illegible] ...tically, as F with the usual degrees of freedom. The calculated F statistic of 19.10 rejects the hypothesis of equal parameters across sectors at the 1% level (the critical value of F, with 54 numerator and 541 denominator degrees of freedom, is approximately 1.5).

Summary and Implications

From the individual sector estimates, it appears that changes in the price of water will have little effect on the usage of other inputs. However, the converse is not true, as indicated by the ϵ_{wi} values. For example, a 1% rise in the price of water would result in a 0.16% drop in the use of capital for the paper and pulp industry (SIC 26). Such a small response makes sense because of the relatively small cost share of the water input. On the other hand, a 1% rise in the cost of capital services for the same industry would lead to a 0.71% reduction in water use.

An important result is that one cannot assume that the industries studied here can be pooled on the basis of identical parameters. Our estimates found capital and production labor to be substitutes in all industry groups analyzed. However, all other pairs of inputs were found to be substitutes for some industries and complements for others. Thus, water planners are likely to find that elasticities based on pooled data are unreliable indicators of price responsiveness for their specific industrial customers.

Table 2. Price Elasticities of Demand for Factors of Production

Price Elasticity	Pooled	Food and Kindred	Paper and Allied	Stone, Clay, Glass	Fabricated Metal	Machinery	Electric and Electronic
	Industry Group						
ϵ_{kk}	−0.46 (.0157)	−0.63 (.142)	−0.52 (.281)	−0.13 (.041)	0.10 (.238)	0.05 (.296)	0.27 (.194)
ϵ_{pp}	−0.32 (.052)	−0.49 (.101)	−2.58 (.089)	−0.18 (.039)	−0.02 (.101)	0.17 (.195)	0.11 (.546)
ϵ_{nn}	−0.08 (.102)	−0.91 (.193)	−0.50 (.273)	0.13 (.112)	−0.15 (.291)	0.40 (.340)	0.66 (1.328)
ϵ_{ww}	−0.56 (.104)	0.14 (.125)	−0.66 (.200)	−0.38 (.080)	−0.41 (.206)	0.14 (.282)	0.54 (.604)
ϵ_{kp}	0.64 (.112)	0.21 (.131)	4.14 (.201)	0.37 (.124)	0.65 (.268)	0.41 (.315)	0.64 (.310)
ϵ_{kn}	−0.05 (.129)	0.35 (.164)	0.26 (.196)	−0.21 (.166)	−0.09 (.337)	−0.46 (.430)	−0.94 (1.278)
ϵ_{kw}	−0.14 (.045)	0.06 (.049)	−0.16 (.085)	−0.03 (.023)	0.03 (.099)	−0.01 (.057)	0.03 (.029)
ϵ_{pk}	0.22 (.039)	0.08 (.049)	1.76 (.085)	0.15 (.050)	0.11 (.043)	0.08 (.061)	0.34 (.014)
ϵ_{pn}	0.45 (.037)	0.45 (.097)	0.01 (.074)	0.02 (.031)	0.09 (.106)	−0.35 (.190)	−0.22 (.560)
ϵ_{pw}	0.05 (.016)	−0.04 (.006)	0.07 (.038)	0.01 (.004)	0.01 (.006)	0.10 (.013)	0.00 (.000)
ϵ_{nk}	−0.03 (.081)	0.20 (.095)	0.30 (.223)	−0.20 (.155)	−0.03 (.124)	−0.14 (.129)	−0.29 (.391)
ϵ_{np}	0.81 (.067)	0.69 (.148)	0.03 (.197)	0.05 (.073)	0.21 (.240)	−0.54 (.294)	−0.37 (.925)
ϵ_{nw}	0.03 (.028)	0.02 (.009)	0.17 (.071)	0.01 (.017)	−0.02 (.000)	0.07 (.029)	−0.01 (.012)
ϵ_{wk}	−1.11 (.356)	0.82 (.635)	−0.71 (.373)	−0.92 (.785)	1.90 (6.012)	−0.26 (3.015)	1.48 (1.379)
ϵ_{wp}	1.18 (.356)	−1.35 (.212)	0.73 (.393)	0.76 (.314)	2.43 (2.405)	18.02 (3.518)	0.56 (.000)
ϵ_{wn}	0.33 (.356)	0.39 (.212)	0.64 (.275)	0.55 (.628)	−3.93 (.000)	8.90 (5.025)	−1.50 (1.839)

Note: Asymptotic standard error estimates are in parentheses. These were calculated as $var(\epsilon^*_{ij}) = var(\delta^*_{ij})/S^{*2}_i$.

Table I. Main estimations of industrial water demand.

Authors	Area	Method	Price elasticity
Grebenstein and Field (1979)	USA	Translog	−0.80/−0.33
Babin, Willis, and Allen (1982)	USA	Translog	−0.66/+0.14
Ziegler and Bell (1984)	Arkansas, USA	Cobb-Douglas	−0.08
Williams and Suh (1986)	USA	Log/Log	−0.97/−0.44
Renzetti (1988)	British Columbia, Canada	Cobb-Douglas	−0.54/−0.12
Schneider and Whitlatch (1991)	Columbus, USA	Log/Log	−1.16
Renzetti (1992a)	Canada	Translog	−0.59*/−0.15*
Dupont and Renzetti (2001)	Canada	Translog	−0.77*

*Intake water.

ticity of industrial demand and the substitutability/complementarity relationships between water and the conventional inputs. Grebenstein and Field (1979) and Babin, Willis, and Allen (1982) study water demand of the manufacturing industry in the United States. Both works estimate a translog cost function using aggregate data. Grebenstein and Field (1979) compute price elasticity values ranging from –0.33 to –0.80, depending on the water price specification adopted. The authors show that water and labor are input substitutes whereas capital and water are complements. Babin, Willis, and Allen (1982) find that price elasticity varies considerably across sectors, ranging from 0.14 for the food industry to –0.66 for the paper and wood industry. Substitution possibilities between water and other production inputs also depend on the industrial sector. Renzetti (1988) provides a deeper investigation of the role of water in industrial plants by breaking down water use into four components: intake, pre-treatment, recirculation, and discharge. According to the sector considered, price elasticity varies from –0.54 to –0.12. The author finds that water intake and recirculation are substitutes, providing some evidence that intake water charges may induce water use efficiency.[2] Dupont and Renzetti (2001) extend the previous analysis by incorporating information on other production inputs than water. They show again that water intake is a substitute to water recirculation, as well as to energy, labor and capital. Last, Reynaud (2003) investigates the structure of industrial water demand in France. Elasticity values are generally in line with the ones found for US and Canadian firms, varying from –0.10 to –0.79 across activities.

Due to data availability problems, empirical evidence on industrial water demand in developing countries is particularly scarce. This lack is problematic, especially in the context of ongoing water policy reforms and increasing quality-related water problems that most developing countries experience. Moreover, given the significant differences

[2] Renzetti (1992) estimates the same model using an enlarged sample and finds similar results.

in firms' production technologies, one could expect water price elasticities to vary between developing and industrialized countries. Wang and Lall (1999) is the first econometric analysis applied to a developing country. They use plant-level information on approximately 1,700 Chinese industrial plants. In contrast to previous works, based on a dual cost function estimation, Wang and Lall (1999) adopt a marginal productivity approach and find an average price elasticity around –1.0, a higher value than those reported for developed countries. Onjala (2001) analyzes industrial water demand in Kenya. The author estimates a single water demand equation based on a dynamic adjustment model. The estimated price elasticities range from –0.60 to 0.37 with high variations across sectors. More recently, Kumar (2004) investigates the water demand of Indian manufacturing plants by adopting an input distance function approach. The author reports an average price elasticity equal to –1.11.

The main results of these studies are the following. First, price elasticities are small but in general higher than domestic ones. Second, estimates strongly depend upon the industry considered. Third, water and labor are mostly substitutes whereas capital and water are complementary inputs. Moreover, excepting Reynaud (2003), none of these papers integrates effluent emissions when estimating the industrial cost function. The implicit assumption is that production and water pollution control decisions are separable. This seems to be a strong assumption, as Reynaud (2003) tests and rejects this separability hypothesis. In what follows, by considering effluent discharge as a joint negative output of the production process, we can assess the impact of environmental regulation on firms' production decisions.

THE LITERATURE ON INDUSTRIAL WATER DEMAND

Babbin, US Census of manufacturing

Renzetti (1992) Canadian census of Manufacturing 1986.

Assumes demand for water is separable from demand for other inputs. But identifies four types of water demand, each with its own price:

- Intake
- Treatment of Intake
- Recirculation
- Treated/discharge of wastewater

Estimates a system of 4 demand functions for the quantities of each type of water, as a function of the costs of each type of water and the level of plant output. Uses translog with constant returns to scale.

Considers 8 sectors.

Key finding: intake and recirculation of water are substitutes.

Dupont and Renzetti (2001)

Canadian census of Manufacturing, 1981, 1986 and 1991

Considers demand for water intake; water recirculated; capital; labor; energy, and materials.

Examines whether water is a variable or a fixed input for firms. Determines that water is a variable input.

Finds that intake water is a substitute for most of the other inputs (recirculated water, labor, capital and energy).

The substitution relationship between intake and recirculated water is stronger when water intake is process-related than when it is related to cooling or to steam production.

Finds that technological change in Canadian manufacturing over this period has been biased towards greater reliance on water intake and less reliance on water recirculation.

Reynaud (2003)

Examines 56 manufacturing plants in southwest France; annual data fir 1994-1996 (153 observations)

Firms either use public water supply systems or are self-supplied.
Firms may or may not treat water prior to use.

Conceptualizes this as three input demands: public supplied water; self-supplied water; treated water.

Has data only on labor, not on other non-water inputs

Eight industry groups

Finds: price elasticities are generally small, but higher than residential demand elasticities.

The price elasticities vary greatly by industrial sector.

Water and labor are mostly substitutes

Feres and Reynaud (2005)

Brazil

Finds water demand definitely price elastic.

Concludes that pricing is likely to be an effective tool to promote both water conservation and reduction in pollution from discharge of industrial wastewater.

Table VI. Price-elasticities (ϵ_{jk}) by industry.

Industry	Obs.	ϵ_{NN}	ϵ_{AA}	ϵ_{TT}	ϵ_{NA}	ϵ_{NT}	ϵ_{AT}
Extractive industry	6	−0.734	0.060	–	−0.397	–	–
		(0.086)	(0.082)		(0.704)		
Metal fabrication	27	−0.241	0.656	−1.651	−0.695	0.944	−0.106
		(0.300)	(1.157)	(0.691)	(1.385)	(1.583)	(0.395)
Chemicals	12	−0.375	−0.063	−2.173	−0.130	0.505	0.210
		(0.134)	(0.078)	(1.624)	(0.194)	(0.327)	(0.275)
Alcohol	84	−0.095	0.155	−0.899	−0.149	0.244	0.050
		(0.243)	(0.222)	(0.656)	(0.177)	(0.324)	(0.280)
Food and beverages	6	−0.304	0.068	−1.474	−0.158	0.462	−0.049
		(0.027)	(0.197)	(0.138)	(0.006)	(0.033)	(0.953)
Paper and wood	9	−0.217	−0.017	–	−0.060	–	–
		(0.343)	(0.074)		(0.162)		
Commercial and services	3	−0.272	−0.046	–	0.077	–	–
		(0.022)	(0.029)		(0.016)		
Others	6	−0.787	0.135	–	−0.200	–	–
		(0.003)	(0.060)		(0.039)		

Notes:Values in parentheses correspond to standard-deviance of elasticities. The second column gives the number of observations for each industry.

REGRESSION ANALYSIS – ASSUMPTIONS

The Linearity Assumption

The first assumption is that the relationship between the dependent variable and the regressors is linear.

Assumption 1.1 (linearity):

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + \varepsilon_i \quad (i = 1, 2, \ldots, n), \tag{1.1.1}$$

where β's are unknown parameters to be estimated, and ε_i is the unobserved error term with certain properties to be specified below.

The part of the right-hand side involving the regressors, $\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}$, is called the **regression** or the **regression function**, and the coefficients (β's) are called the **regression coefficients**. They represent the marginal and separate effects of the regressors. For example, β_2 represents the change in the dependent variable when the second regressor increases by one unit while other regressors are held constant. In the language of calculus, this can be expressed as $\partial y_i / \partial x_{i2} = \beta_2$. The linearity implies that the marginal effect does not depend on the level of regressors. The error term represents the part of the dependent variable left unexplained by the regressors.

The Strict Exogeneity Assumption

The next assumption of the classical regression model is

Assumption 1.2 (strict exogeneity):

$$\mathrm{E}(\varepsilon_i \mid \mathbf{X}) = 0 \quad (i = 1, 2, \ldots, n). \tag{1.1.7}$$

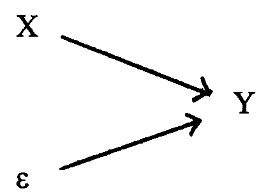
Here, the expectation (mean) is conditional on the regressors for *all* observations. This point may be made more apparent by writing the assumption without using the data matrix as

$$\mathrm{E}(\varepsilon_i \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = 0 \quad (i = 1, 2, \ldots, n).$$

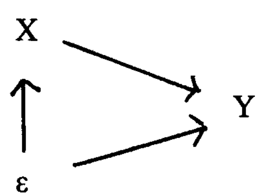
Suppose for simplicity that there is a single regressor, so that the regression equation is:

$$Y = \alpha + \beta X + \varepsilon$$

Suppose that the assumption of strict exogeneity is satisfied; then we can represent the situation diagrammatically as follows:



Instead, now suppose that the assumption of exogeneity is violated. In that case, we can represent the situation diagrammatically as follows:



In this case, some of the effect of ε on Y is transmitted through X. The results is that this *contaminates* the estimate of β, the coefficient showing the effect of X on Y. Some of the effect of ε on Y is wrongly attributed to X.

What does this have to do with the choice of what to estimate? The answer is that, ideally, you would choose to estimate an equation that is consistent with the assumption of strict exogeneity. In other words, you want to avoid a situation where the regressors might be *endogenous*.

In many circumstances, this would typically be an argument against estimating a production function, because the regressors – the input levels – are likely to be endogenous. By contrast, for a price-taking firm, its input and output prices are likely to be exogenous, which would typically argue for estimating unconditional input demand functions, the output supply function and/or the profit function.

As a very simple illustration of the endogeneity of the regressors in estimating a production function, consider the simple case of a Cobb-Douglas production function with a single input, labor:

$$Y = A L^{\beta} e^{\varepsilon}$$

or

(1) $$\ln Y = \ln A + \beta \ln L + \varepsilon$$

with unconditional input demand for labor given by

(2) $$L = \left(\frac{w}{p}\right)^{\left(\frac{1}{\beta-1}\right)} \left(\beta A e^{\varepsilon}\right)^{\left(\frac{1}{1-\beta}\right)}$$

Suppose that the random factors in the random term ε include *soil quality* or the farmer's *managerial ability*, both of which variables are unobserved by the econometrician. A higher level of ε (corresponding to better soil or better ability) influences output, Y, directly, it also influences output indirectly through the effect on the farmer's choice of labor input.

Since soil quality or managerial ability cannot be observed by the econometrician, they cannot be controlled when estimating the production function (1) and they will contaminate the estimate of β, the coefficient associated with ln L in estimating (1).

However, they do *not* contaminate the estimate of $1/(\beta - 1)$, the coefficient of (w/p) in estimating the unconditional demand function (2), because (w/p) is exogenous to the farmer.