The Translog Cost Function

A general introduction

Suppose we have no testable information about the specific **properties** of a *production function* (2 inputs and 1 output)

$$f: R_+^2 \to R_+$$

describing the technology permitting to obtain a positive output using the two inputs (e.g. *L>0* e *K>0*). Yet, we are requested to estimate the the scale regime, the efficiency conditions, ecc. The problem is how to proceed.
Alternatively, we may face the problem of describing empirically the technology governing a *transformation function*

$$g(\boldsymbol{x}, \boldsymbol{y})$$

which says (by means of an implicit function of inputs ($\boldsymbol{x}$) and outputs ($\boldsymbol{y}$)) that the *n* inputs jointly produce the *m* positive outputs with the same production process.
How should we proceed in both cases?
The first problem could be solved, in principle, by using some approximation to a flexible version of function *f*: a C-D function, a CES version, a Leontief version and so on. After imposing the appropriate and theoretically plausible properties/restrictions to the relation between inputs and output resulting from the approximation, we can subject it to empirical testing.
On the contrary, the second problem is more difficult to deal with. We should in fact define an output index before proceeding to the empirical study of the technology. This would introduce elements of arbitrariness.

Yet, we know that, **given the prices of inputs**, the minimization of the expenditure necessary to produce a given value of output (or outputs) generates a continuous and differentiable function called (minimum) cost function having the property to describe technology as good as it is done by the above production and transformation functions. For the 2 input – 1 output case we can write

$$C: R_+^3 \to R_+$$

Unfortunately, absence of specific information on the functional form of the production function *f* implies absence of specific information on the functional form of *C* because **each production function generates (as the outcome of an expenditure minimization process) a specific cost or expenditure function**. Hence, a P_2 Taylor expansion centered in the origin can help exploit the duality existing between production maximization and cost minimization for empirical purposes[1]. Let the center of expansion be (0,0,0). Then, using Taylor formula

$$C(w,r,Q) \approx C(0,0,0) + \left[\frac{\partial C(0,0,0)}{\partial w}(w-0) + \frac{\partial C(0,0,0)}{\partial r}(r-0) + \frac{\partial C(0,0,0)}{\partial Q}(Q-0)\right]$$
$$+\frac{1}{2!}\left[\frac{\partial^2 C(0,0,0)}{\partial w^2}(w-0)^2 + \frac{\partial^2 C(0,0,0)}{\partial r^2}(r-0)^2 + \frac{\partial^2 C(0,0,0)}{\partial Q^2}(Q-0)^2\right]$$
$$+\frac{1}{2!}\left[2\frac{\partial^2 C(0,0,0)}{\partial w \partial r}(w-0)(r-0) + 2\frac{\partial^2 C(0,0,0)}{\partial w \partial Q}(w-0)(Q-0) + 2\frac{\partial^2 C(0,0,0)}{\partial r \partial Q}(r-0)(Q-0)\right]$$

[1] Clearly, this could be done to *f* too. However, when the firm produces more than one outputs (transformation function) this is almost impossible.

where *w* and *r* are input prices and *Q* is output. Here all the partial derivatives are evaluated at (0, 0, 0).

This expansion should be considered as the best second order polynomial approximation to *C*. Note that by the general assumptions, *C* is well defined defined and twice differentiable in the origin and so we may say that function *C* is (locally) determined by its derivatives at the origin. Recall that Taylor's theorem is of asymptotic nature: it only tells us that the error in the approximation tends to zero faster than, in our case, any other second-degree polynomial (*k*.th order polynomial, in the general case), when $w \to 0$; $r \to 0$; $Q \to 0$. In this sense, it is the best approximation. However, it does not tell us how large the error is in any *neighbourhood of the center of expansion*. In principle, one may think to select several Taylor polynomials with different centers of expansion to have reliable Taylor-approximations of the original function *C*. However, there is an economic reason justifying the choice of (0, 0, 0) as a center of expansion. It will be discussed later.
To make the above expansion a bit more understandable from a economic point of view, we may rewrite it as follows:

$$C(w,r,Q) \approx C(0,0,0) + \begin{bmatrix} \frac{\partial C(0,0,0)}{\partial w} & \frac{\partial C(0,0,0)}{\partial r} & \frac{\partial C(0,0,0)}{\partial Q} \end{bmatrix} \begin{bmatrix} w-0 \\ r-0 \\ Q-0 \end{bmatrix}$$

$$+ \frac{1}{2!}\begin{bmatrix} w-0 & r-0 & Q-0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 C(0,0,0)}{\partial w^2} & \frac{\partial^2 C(0,0,0)}{\partial w \partial r} & \frac{\partial^2 C(0,0,0)}{\partial w \partial Q} \\ \frac{\partial^2 C(0,0,0)}{\partial r \partial w} & \frac{\partial^2 C(0,0,0)}{\partial r^2} & \frac{\partial^2 C(0,0,0)}{\partial r \partial Q} \\ \frac{\partial^2 C(0,0,0)}{\partial Q \partial w} & \frac{\partial^2 C(0,0,0)}{\partial Q \partial r} & \frac{\partial^2 C(0,0,0)}{\partial Q^2} \end{bmatrix} \begin{bmatrix} w-0 \\ r-0 \\ Q-0 \end{bmatrix}$$

The expansion includes three parts.

1. $C(0,0,0)$ is a constant and corresponds to the value that the function takes when the independent variables are zero. With no production and no input (which makes their price equal to zero) only fixed costs are present. Therefore, we may say that $C(0,0,0)$ performs the role of a constant/intercept in a regression estimation included to allow for fixed cost. This is the reason justifying the choice origin as a center of expansion.

2. In the second and third parts, each derivative (any order) is evaluated at the point of expansion and then they are **constant terms**. The linear part, if the third quadratic part were absent, corresponds to the hypothesis that the technology is linear.

3. The third term is quadratic with a 3 × 3 matrix. It includes the second derivatives (own and cross) of *C*. Terms on the main diagonal are clearly negative because cost function is concave in input prices; second derivatives correspond to first derivatives of compensated factor demand w.r.t. their own prices. Due to Young's theorem (reversibility of the order of differentiation) the matrix is **symmetric** (for instance, $\frac{\partial^2 C(.)}{\partial w \partial r} = \frac{\partial^2 C(.)}{\partial r \partial w}$). Since it corresponds to the Hessian of the cost function, it should be negative semi definite for the cost function to be quasi-concave. But this cannot be said a priori because we do not know the sign of the second own derivative w.r.t *Q*.

The function can be further illustrated by replacing each (fixed value) derivative with a coefficient. Call

$$C(0,0,0) = \alpha_0$$

$$\alpha_w = \frac{\partial C(0,0,0)}{\partial w}$$

$$\alpha_r = \frac{\partial C(0,0,0)}{\partial r}$$
$$\alpha_Q = \frac{\partial C(0,0,0)}{\partial Q}$$
$$\alpha_{ww} = \frac{1}{2}\frac{\partial^2 C(0,0,0)}{\partial w^2}$$
$$\alpha_{rr} = \frac{1}{2}\frac{\partial^2 C(0,0,0)}{\partial r^2}$$
$$\alpha_{QQ} = \frac{1}{2}\frac{\partial^2 C(0,0,0)}{\partial Q^2}$$
$$\gamma_{wr} = \frac{\partial^2 C(0,0,0)}{\partial w \partial r} = \frac{\partial^2 C(0,0,0)}{\partial r \partial w} = \gamma_{rw}$$

$$\gamma_{wQ} = \frac{\partial^2 C(0,0,0)}{\partial w \partial Q} = \frac{\partial^2 C(0,0,0)}{\partial Q \partial w} = \gamma_{Qw}$$
$$\gamma_{rQ} = \frac{\partial^2 C(0,0,0)}{\partial r \partial Q} = \frac{\partial^2 C(0,0,0)}{\partial Q \partial r} = \gamma_{Qr}$$

Then the above function rewrites

$$C(w,r,Q) \approx \alpha_0 + \alpha_w \times w + \alpha_r \times r + \alpha_Q \times Q$$
$$+\frac{1}{2}\left[\alpha_{ww} \times w^2 + \alpha_{rr} \times r^2 + \alpha_{QQ} \times Q^2\right]$$
$$+\frac{1}{2}\left[2\gamma_{wr}[(w) \times (r)] + 2\gamma_{wQ}[(w) \times (Q)] + 2\gamma_{rQ}[(r) \times (Q)]\right]$$

Or in a more compact way

$$C(w,r,Q) \approx \alpha_0 + \sum_{i=1}^{2} \alpha_i w_i + \alpha_Q Q$$
$$+\tfrac{1}{2}\left[\alpha_{ww} w^2 + \alpha_{rr} r^2 + \alpha_{QQ} Q^2\right] +$$
$$+\gamma_{wr}[(w \times r)] + \gamma_{wQ}[(w \times Q)] + \gamma_{rQ}[(r \times Q)] \qquad (1)$$

This means that we can empirically estimate the values of the above parameters and use them to evaluate competing hypotheses about technology.
In fact, the expanded version of the cost function retains the properties of the general minimum cost function resulting from the constrained expenditure minimization. Some properties are

(a) $C(\boldsymbol{w}, Q)$ is increasing in Q and non-decreasing in $\boldsymbol{w}$;
(b) $C(\boldsymbol{w}, Q)$ is linear homogeneous in $\boldsymbol{w}$: $C(k\boldsymbol{w}, Q) = kC(\boldsymbol{w}, Q)$;
(c) $C(\boldsymbol{w}, Q)$ is continuous and concave in each w_i;
(d) Shephard's lemma: $\partial C(\boldsymbol{w}, Q))/\partial w_i = x_i(\boldsymbol{w}, Q)$.

where **w** is the vector of the 2 input prices and **x** the vector of their quantities and k is a scalar.
Therefore, the above equation (1) must be estimated considering the implications of a) – d. This means that (i) estimations must include restrictions and (ii) estimated parameters should be consistent with the above properties. For instance, the above 3 × 3 symmetric matrix rewrites

$$\begin{bmatrix} \frac{\partial^2 C(0,0,0)}{\partial w^2} & \frac{\partial^2 C(0,0,0)}{\partial w \partial r} & \frac{\partial^2 C(0,0,0)}{\partial w \partial Q} \\ \frac{\partial^2 C(0,0,0)}{\partial r \partial w} & \frac{\partial^2 C(0,0,0)}{\partial r^2} & \frac{\partial^2 C(0,0,0)}{\partial r \partial Q} \\ \frac{\partial^2 C(0,0,0)}{\partial Q \partial w} & \frac{\partial^2 C(0,0,0)}{\partial Q \partial r} & \frac{\partial^2 C(0,0,0)}{\partial Q^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial x_1^h(0,0,0)}{\partial w} & \frac{\partial x_1^h(0,0,0)}{\partial r} & \frac{\partial x_1^h(0,0,0)}{\partial Q} \\ \frac{\partial x_2^h(0,0,0)}{\partial w} & \frac{\partial x_2^h(0,0,0)}{\partial r} & \frac{\partial x_2^h(0,0,0)}{\partial Q} \\ \frac{\partial^2 C(0,0,0)}{\partial Q \partial w} & \frac{\partial^2 C(0,0,0)}{\partial Q \partial r} & \frac{\partial^2 C(0,0,0)}{\partial Q^2} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{ww} & \gamma_{wr} & \gamma_{wQ} \\ \gamma_{rw} & \alpha_{rr} & \gamma_{rQ} \\ \gamma_{Qw} & \gamma_{Qr} & \alpha_{QQ} \end{bmatrix} = \begin{bmatrix} - & - & + \\ - & - & + \\ + & + & ? \end{bmatrix}$$

The main diagonal has the first 2 negative elements (they correspond to the fist derivatives of the compensated factor demands) and the element outside the diagonal positive (they correspond to the first derivative of the compensated factor demand w.r.t. *Q*). The last element in position (3,3) corresponds to the derivative of the marginal cost and can be either positive (increasing marginal cost) or negative (decreasing marginal cost) or zero (constant marginal cost, i.e. C-D technology). Estimation will tell the signs and help us infer the technology and the scale regime.

As it was stressed above, restrictions must be imposed but they will be considered when the most commonly employed cost expansion used in applied works, namely the *Translogaritmic Cost Function* (TCF), will be analysed.

Meanwhile let us refresh our memory. Scale economies are tested (in the one output case) by the ratio MC/AC. If the ratio is >1we have diseconomies of scale (Deceasing Returns). If it is < 1 we have Economies of Scale. (Revise the graphical illustration; see Gravelle-Rees, page 120). Applied cost analysis may be directed at estimating the presence or the absence of Increasing Return to Scale **for different levels of the output**. TCF allows researchers to investigate empirically this issue, and many others.

The *Translogaritmic Cost Function*

(Based on: L. Christensen and W. Green, Economies of Scale in the U.S. Electric Power Generation (1976), Journal of Political Economy, pages 655-676

The purpose of the translog cost function is to identify for empirical application a functional form for the expansion of the cost function that could be so general as to embody **all the assumptions and results of our cost minimization model**. In particular, we want a cost function that allows for **U-shaped average cost** in order to evaluate scale regimes for different levels of output. The other conditions include the following:

1. Input demand is downward sloping.
2. Cross price effects are symmetric.
3. The shift in marginal cost w.r.t. an input price is equal to the shift in the input's demand w.r.t. output.
4. The sum of own and cross price elasticities is equal to zero.
5. A proportional increase in all input price must shift cost by the same amount holding output constant

The *N* input and 1 output in natural log form is

$$lnC(\mathbf{w},Q) \approx \alpha_0 + \sum_{i=1}^{N} \alpha_i \times lnw_i + \alpha_Q ln(Q) + \frac{1}{2}\, \alpha_{QQ} [ln(Q)]^2 + \sum_{i=1}^{N} \gamma_{Qw_i} [lnQ \times lnw_i] + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \gamma_{ij} [\, (lnw_i) \times (lnw_j)\,] \qquad (2)$$

Where the last term encompasses the squared of each log of input price.
If we log differentiate (2) w.r.t. *ln(w_i)* we have

$$\frac{\partial \ln C(\mathbf{w},Q)}{\partial \ln(w_i)} = \frac{\partial C(\mathbf{w},Q)}{\partial w_i}\frac{w_i}{C} = \underbrace{x_i^h(\mathbf{w},Q)}_{\text{Shephard-s Lemma}} \frac{w_i}{C} = \frac{x_i^h w_i}{C} = S_i \equiv \text{Share of factor i's expenditure} \qquad (3)$$

As a result

$$\frac{\partial \ln C(\mathbf{w},Q)}{\partial \ln(w_i)} = S_i = \alpha_i + \gamma_{Qw_i} \ln Q + \sum_j \gamma_{ij} \ln w_j \quad \text{with } \gamma_{ij} = \gamma_{ji} \qquad (3')$$

The translog cost function (2) is estimated together with N – 1 shares (one is dropped to avoid the problem of a singular variance-covariance matrix for the disturbances). The **data** generated by the firm's behavior that we observe and use for empirical purposes **are**

- **(i)** **total cost**
- (ii) the allocation of total cost across the various inputs (i.e., input expenditure **shares**)
- (iii) the firm's **output** level
- (iv) the **input prices** that the firm is **supposed to passively face**. This (general) assumption is fundamental to cost minimization. But can be questionable. Consistently with the theoretical model of cost minimization subject to a quantity constraint, this very neoclassical assumption maintains that input prices are beyond firm's influece i.e., that firms have no market power in the input markets.

One necessary restriction on the parameter estimates across equations is that imposed by **linear homogeneity** of the cost function w.r.t. input prices. Then, we must impose

$$\sum_{i=1}^{n} \alpha_i = 1 \qquad \sum_{i=1}^{n} \gamma_{Qw_i} = 0 \qquad \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} = 0$$

Of particular interest is the scale economy effect. The translog function allows for both positive and negative scale effects, that is, average cost can both decrease and increase across the range of the cost function. In this sense, the translog function can represent a production function that can even be not homogeneous.
If we log differentiate (2) w.r.t. *ln(Q)* we obtain the elasticity of Cost w.r.t. output

$$\frac{\partial \ln C(\mathbf{w},Q)}{\partial \ln Q} = \underbrace{\frac{\partial C(\mathbf{w},Q)}{\partial Q}}_{MC} \underbrace{\frac{Q}{C}}_{1/AC} = \frac{MC}{AC} = \alpha_Q + \alpha_{QQ} \ln Q + \sum_{i=1}^{n} \gamma_{Qw_i} \ln w_i$$

Recall that if AC is decreasing AC > MC and we have IRS (average cost is falling). Given the estimated parameters, if the above equation is fitted with different values of *Q* and, for example, the mean of prices, then one can evaluate the possible change of the scale regime emerging from calculation run with different quantity values.

If the estimated cost elasticity is constant because the estimated parameters $\alpha_{QQ} = \gamma_{Qw_i} = 0 \quad \forall i$, then $\frac{\partial \ln C(\mathbf{w}, Q)}{\partial \ln Q} = \frac{MC}{AC} = \alpha_Q$ which is typical of a C-D technology.

The conclusion is that the translog function allows for estimation of parameters that embody all of the relations that are derivable from the general model of cost minimization subject to an output constraint. Results extend to the case of multiunit production. Yet it requires **strict exogeneity** of input prices.
Pros and Cons of this methodology are discussed during classes.
Actual application (with all the variants) of the methodology are beyond the purposes of these notes whose goal is only to underline the **duality existing between productions and cost, under the assumption that input prices are perfectly competitive.**

An important reference is
R. Sickles and V. Zelenyuk (2019), Measurement of Productivity and Efficiency. Theory and Practice, Cambridge University Press; In particular: Ch. 6.