The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

May 2003

The Dynamics of Productivity Growth in U.S. Agriculture

Alejandro Onofri
Department of Economics
Universidad de La Plata, Argentina
aonofri@fullzero.com.ar

Lilyan Fulginiti
Department of Agricultural Economics
University of Nebraska-Lincoln
lfulginiti@unl.edu

Walter Enders
Department of Economics
University of Alabama
wenders@cba.ua.edu

Paper prepared for presentation at the American Agricultural Economics Association Annual Meetings, Montreal, Canada, July 27-30, 2003.

Copyright 2003 by Alejandro Onofri, Lilyan Fulginiti and Walter Enders. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

The Dynamics of Productivity Growth in U.S. Agriculture

Abstract

A dynamic model of productivity measurement that incorporates public goods is developed. Cointegration is used to estimate dynamic derived demands and economies of scale in US agriculture, 1948-1994. The impact of public inputs on the steady state stocks of private capital and their shadow prices are estimated.

Introduction

Neoclassical models of growth (Solow, Ramsey) have been widely criticized because they cannot explain productivity changes. According to these models, growth is exogenously given by an unexplained rate of technical change. As a response, endogenous growth theories prove that continuous growth is possible because of the existence of non-rival inputs of production (i.e., inputs that can be used by many firms at the same time or by the same firm repeatedly without additional cost). In these models, two necessary conditions for endogenous growth are: increasing returns to scale over all inputs, and positive impacts of non-rival inputs on the returns to investment. The main contribution of this study is to introduce a dynamic model of productivity measurement that incorporates public goods (non-rival by definition) as external factors to the firms. It also rationalizes the provision of public inputs by a benevolent social planner that internalizes the effects of them. Estimable functions that allow testing the necessary conditions for endogenous growth are obtained.

Many other papers have focused on the effects of public goods on private production, and most of them have found positive impacts[1]. For example, Aschauer's (1989) pioneer work estimates a single production function for the U.S. economy including public infrastructure as factor of production. Lynde and Richmond (1992) and Berndt and Hansson (1992) have also used duality theory to estimate the role of infrastructure in private production in the U.S. and Sweden, respectively. Nadiri and Mamuneas (1994) estimate the impacts of public capital and research and development (R&D) on the cost structure of twelve U.S. manufacturing industries, and Morrison and Schwartz (1996) study the regional effects of public infrastructure on the U.S. manufacturing sector. Both papers adopt a dual approach and find, in general, positive effects of public inputs on manufacturing productivity. The last paper also finds increasing returns to scale over all inputs (including infrastructure), but it does not include R&D.

For the agricultural sector, papers like Antle (1983) and Craig et al. (1997) find positive effects of public infrastructure and research on agricultural productivity but their approach is based on estimating a single production function. Binswanger et al. (1993) estimates the impacts of infrastructure and R&D in India. They consider, in a static framework, that public infrastructure investments are regionally allocated toward areas that are more productive. In contrast, the present study develops a dynamic model of productivity measurement. This approach, based on duality theory, maintains producer rationality and allows examination of the impacts of public inputs on producer's behavior.

---

[1] Exceptions are Garcia-Mila and McGuire (1992) and Holtz-Eakin (1994). They find insignificant effects

The model is tested with data for the U.S. agricultural sector. United States agricultural productivity has increased at an annual average rate of two percent over the 1948-1994 period. Some authors have found that productivity growth has been the main factor contributing to economic growth of the agricultural sector (Ball et al., 1997). Additionally, the provision of public goods in the form of public research and extension, and infrastructure has been sizable in this country. In an atomistic environment, these public expenditures are traditionally justified because of their low degree of appropriability and high costs. Theoretically consistent firms' dynamic demands for inputs are then estimated for U.S. agriculture including stocks of public capital and R&D as quasi-fixed factors. The existence of economies of scale and the likely positive impact of public inputs on the steady state stocks of private capital can be tested.

There are several reasons to undertake this study. First, the possibility of endogenous growth in the agricultural sector may imply spillovers to other sectors and, in particular, may have important effects on the growth of regional economies based on agricultural activities. Second, by determining the substitution or complementarity between public and private inputs, one may explain the recent evolution of private factors in the U.S. agricultural sector. Ball et al. (1997) show the increasing use of materials and the decreasing use of labor by the sector. Finally, the estimation of shadow prices for public capital and R&D stocks may provide an indicator to policy makers of the optimal provision of public investment.

This chapter develops as follows. Section II presents a summary of the endogenous growth theory involving publicly provided goods and the related testable hypotheses

---

of public infrastructure on private production.

using a dual approach. Section III introduces a dynamic model in which both producers' and government's behaviors are rationalized. The testable hypotheses are then revisited. Section IV introduces the empirical model and section V presents the preliminary results. Finally, conclusions and future lines of research are stated in section VI.

Growth Theory and Testable Hypothesis

In the neoclassical models of growth (Solow, Ramsey), the rate of growth of per capita output is a decreasing function of the per capita stock of private capital. Without technical change and with a well-behaved neoclassical production function, the level of per capita output converges to a steady state where the growth of per capita private capital eventually stops. This result, implied by the assumption of decreasing returns to capital, has been one of the major criticisms to these models.

As a response to these empirically unsustainable results, endogenous growth theory arose proposing different hypotheses. These theories incorporate into the models the reasons for technical change to occur based on the presence of externalities that originate nonconvexities.

Nonconvexities play an important role in new theories of growth. They are generally due to the presence of nonrival goods. Following Romer (1990), nonrivalry can be interpreted in two ways. First, nonrival factors of production are valuable "inputs that can be used simultaneously in more than one activity." Under this definition, public goods, like public infrastructure for instance, are nonrival inputs that can be used by many producers at the same time. Alternatively, one can define a nonrival input as that input that can be used repeatedly in the same activity. With this definition, a new chemical process, for example, is an input that can be used more than once in the production of a certain product. In this case, nonconvexities are intrinsically associated to this input: there is a high cost of producing the first unit, but the cost of producing subsequent units is zero. In any case, since the presence of nonrival inputs generates nonconvexities, the production function can be characterized by increasing returns to scale:

$$F(\lambda R, \lambda N) > F(\lambda R, N) = \lambda F(R, N), \text{ with } \lambda > 0$$

where R and N stand for rival input and nonrival inputs, respectively. Thus, if rival and nonrival inputs are doubled ($\lambda = 2$), output is more than doubled.

One of the pioneer studies in the endogenous growth literature has been that by Romer (1986). In this paper, Romer specifies a production function $F(k_i, K, \mathbf{x}_i)$, being k_i and $\mathbf{x}_i$ firm-specific inputs ($\mathbf{x}$ can be seen as a vector of inputs) and K an input external to the firm, like "the level of knowledge" defined as a function of the "firm-specific knowledge" ($K=g(\Sigma k_i)$). If F is increasing in K and linear homogeneous in k_i and $\mathbf{x}_i$, a perfect competitive equilibrium is still possible, but the factor k_i no longer exhibits diminishing returns. Consequently, permanent endogenous growth of output per capita is allowed.

Barro (1990) has developed a similar model where K can be interpreted as the stock of public capital (hereafter G). The intuition is that publicly provided capital (like roads, sewer capital, etc.) has positive impact on private production affecting the productivity of the firm-specific inputs. Public capital is assumed a public input that can be used by additional producers without cost. Consequently, total stocks of public goods enter in the production function of each individual firm. In this context, two necessary conditions for the hypothesized endogenous growth are: existence of increasing returns to scale over all inputs, and existence of constant returns to scale over factors that can be accumulated (private and public capital). This second condition implies that private capital is continuously accumulated and there is an optimal ratio between private to public capital. A weaker requirement, alternative to this condition, would be a positive impact of G on the demand for capital. Although not ensuring continuous growth, the presence of this nonrival input would imply a positive government's contribution to growth.

The conditions mentioned above (i.e., increasing returns to scale over all inputs and positive impact of public inputs on private capital accumulation) can be rationalized using the theory of the firm. The following section introduces a model in which firms respond to changes in public inputs provided by a benevolent social planner. Estimable functions that allow testing for the hypothesized endogenous growth conditions are then obtained in a model that maintains producer rationality.

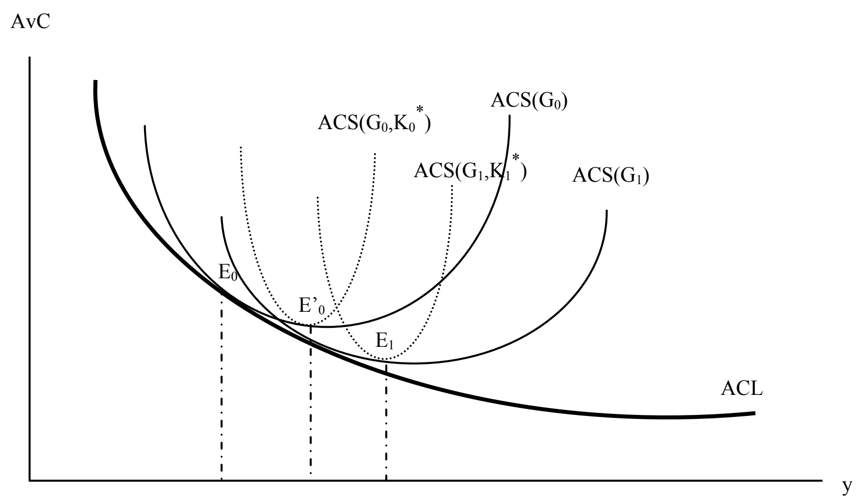
The Model

A dynamic dual model of the firm is used to explain growth based on the existence of public inputs. As was hypothesized, public goods might have positive effects on firms' production. If the dual problem of the firms is considered, public inputs reduce cost of production given the level of firms' output. In this manner, increases of public inputs increase firms' productivity.

The model assumes that economic agents are intertemporal optimizers: firms minimize intertemporal costs of production. Instantaneous adjustment of inputs is not possible because of the existence of cost of adjustment.

In their optimizing behavior, firms take public inputs as given. Public inputs are considered quasi-fixed inputs of production that they cannot adjust to obtain the minimum possible cost.

The following figure shows the dynamics of the firms' behavior.



G represents the stock of the public input. K is the stock of private capital. Three average cost curves (faced by the firms) are shown in the graph. ACS(G_t, K_t) represents a very short-run average cost curve when private inputs (capital in this case) and public inputs are fixed. ACS(G_t) is the short-run average cost curve when only public inputs are fixed. Finally, ACL is the long-run average cost curve when all inputs are adjusted.

At each period t, the firms observe the public input stock G and choose the optimal path of investment (I) that allows them to reach the optimal steady state (SS) stock K*. Starting at E_0 and with a stock of public inputs G_0, firms choose an optimal path of I that allows the firm to reach K_0^* at the minimum cost. The firm moves from E_0 to E'_0. The path is adjusted the next period when the stock G_1 implies a new SS stock K_1^*. The firm then moves to E_1. The two conditions for the hypothesized endogenous growth of the firms can then be seen in the graph:

I. Increasing returns to scale over the long-run average cost curve (ACL): negative slope of ACL.

II. Positive effects of G on the SS stocks of the private capital (i.e. the private input "that can be accumulated"): the SS stock of K increases from K_0^* to K_1^* when G grows from G_0 to G_1.

More formally, firms solve the following problem:

$$\begin{gathered}\underset{I(t)>0}{\text{Min}} \int_0^{\infty} e^{-\rho t}[C(y, Z, I; G) + p'Z]dt \\ \text{subject to } \dot{Z} = I - \delta Z \\ Z(0) = Z_0 \\ Z(t) > 0 \quad \forall t\end{gathered} \tag{1}$$

where C(y, Z, I; G) is the variable cost function; y is the only output; Z is the vector of stocks of quasi-fixed inputs; p is the rental price vector corresponding to Z; I is the vector of gross changes in quasi-fixed inputs; δ is the diagonal matrix containing the depreciation rates of Z; G is the vector of public inputs; and $\rho > 0$ is the firm's real rate of discount. It is assumed that there is one perfectly variable input whose price (w) is normalized to one.[2] Thus, the elements of p are relative rental prices.

[2] Given w = 1, the variable cost function is C(1, y, Z, I; G). For simplification, C(1, y, Z, I; G) = C(y, Z, I; G) is used.

Define now J(Z, y, p; G) as the value function that solves problem (1). Assuming that C(y, Z, I; G) satisfies the set of regularity conditions (A.1) – (A.6) and J(Z, y, p;G) satisfies properties (B1) – (B5) (see Appendix 1), duality between C(y, Z, I; G) and J(Z, y, p; G) can be established.

Duality between C(y, Z, I; G) and J(Z, y, p; G):[3] *any J Z, y, p; G.) satisfying conditions (B) is the value function corresponding to C(y, Z, I; G) that satisfies conditions (A) and is defined by*

$$C(y, Z, I; G) = \underset{p}{\mathrm{Max}}[\rho J(Z, y, p; G) - p'Z - J_z'(Z, y, p; G)(I - \delta Z)] \qquad (2)$$

or

$$\rho J(Z, y, p; G) = \underset{I}{\mathrm{Min}}[C(y, Z, I; G) + p'Z + J_z'(Z, y, p; G)(I - \delta Z)] \qquad (3)$$

These two equations provide the relationship between the cost function C(y, Z, I; G) and the value function J(Z, y, p; G). They allow expressing the parameters of C(y, Z, I; G) in terms of the parameters of J(Z, y, p; G) when firms minimize intertemporal costs. Thus, the derivative properties that characterize C(y, Z, I; G) can be recovered from the parameters of J(Z, y, p; G).[4] Therefore, the two previously mentioned endogenous growth conditions can be tested through estimation of parameters of J(Z, y, p; G).

---

[3] Epstein (1983).
[4] See Appendix for the derivative properties.

Conditions for Endogenous Growth

1) The impact of G on

   a) The cost function: this is provided by the fifth derivative property explained in the Appendix 1. The following expression represents this effect:

$$C_G(y, Z, I; G) = \rho J_G(Z, y, p; G) - J_{ZG}(Z, y, p; G)\dot{Z}^*(Z, y, p; G)$$

   which is the shadow price of G when the firms are out of the SS. At the SS, the shadow price is

$$C_G(y, Z, I; G) = \rho J_G(Z, y, p; G)$$

   If this expression is negative, the shadow price of G is positive, meaning that public inputs reduce cost of production.

   b) The dynamic demand for private capital: it can be shown that the dynamic demand for the quasi-fixed inputs Z can be expressed as

$$\dot{Z}^*(Z, y, p; G) = M(p, G)[Z - \bar{Z}(p, G)] \tag{4}$$

   where $\bar{Z}(p, G)$ is the SS stock of Z and M(p,G) is a stable adjustment matrix. This expression yields a flexible accelerator adjustment path for the stocks Z and is the reason for these dynamic models to be called "multivariate flexible accelerator models" (Epstein(1983)). The form of M(p,G) is determined by the functional form of C(y, Z, I); however, only under certain conditions, it can be successfully expressed as an explicit function of the parameters of C(y, Z, I).[5]

   The effect of G on the dynamic demand for Z can then be decomposed in the effect on the adjustment matrix and the effect on the SS stock of Z. The condition for endogenous growth would be for G to increase the SS stock of private capital K (one of the quasi-fixed factors of the firms). The effect on the adjustment matrix is only an effect on the speed of adjustment toward the SS. However, it is still required for this adjustment to be stable.

2) Scale Effects: there must be increasing returns to scale over all factors of production (public and private factors). Increasing returns to scale can be evaluated by considering the elasticity of cost with respect to output (ε_{CY}). It is well known in the production economics literature that the elasticity of cost with respect to output is the dual expression of the elasticity of scale (η_y): $\varepsilon_{cy}=1/\eta_y$.[6] When the elasticity of cost with respect to output is less than one, firms exhibit economies of scale. However, in the presence of factors external to the firm, some adjustments should be made in order to obtain ε_{cy}. Morrison and Schwartz (1996) show how to adjust the elasticity of cost with respect to output when there are quasi-fixed inputs in a static cost minimization framework.[7] This approach is extended here for the case of intertemporal optimization.

---

[5] See Epstein (1983) for details.
[6] See Chambers (1988) for details.
[7] The approach is based on Le Chatelier principle. Taking the derivative with respect to Y on both sides of the identity $C^A(\mathbf{P}, \mathbf{Pg}, Y) \equiv C(\mathbf{P},\mathbf{G}(\mathbf{P},\mathbf{Pg},Y),Y)$ gives

$$\frac{\partial C^A}{\partial Y} = \frac{\partial C}{\partial Y} + \sum_G \frac{\partial C}{\partial G}\frac{\partial G}{\partial Y}$$

Finally, completing elasticities gives

$$\varepsilon_{CY}^A = \varepsilon_{CY} + \sum_G \varepsilon_{CG}\,\varepsilon_{GY}$$

Define the shadow price of the public input $P_G = P_G(Z, y, p; G)$. This shadow price can be interpreted as an "inverse demand" for the public input. Solving for G, given P_G, gives the direct shadow demand for G that can be substituted into (4) to get

$$\begin{aligned} C(y, Z, I; G(\bar{P}_G, Z, y, p)) = \underset{p}{\mathrm{Max}}[\rho J(Z, y, p; G(\bar{P}_G, Z, y, p)) - p'Z - \\ - J_z^{'}(Z, y, p; G(\bar{P}_G, Z, y, p))(I - \delta Z)] \end{aligned} \tag{5}$$

Taking the derivative with respect to y, we obtain the adjusted effect of output on cost when the 'shadow demand' for G also changes with firms' output:

$$\frac{\partial C^A}{\partial y} = \rho J_y - J_{zy}^{'} \dot{Z}^* + (\rho J_G^{'} - \dot{Z}^{*'} J_{zG}) G_y(\bar{P}_G, Z, y, p) \tag{6a}$$

At the SS, this expression becomes

$$\begin{aligned} \frac{\partial C^A}{\partial y} &= \rho J_y + \rho J_G^{'} G_y(\bar{P}_G, Z, y, p) \\ &= C_y + C_G^{'} G_y(\bar{P}_G, Z, y, p) \end{aligned} \tag{6b}$$

Completing elasticities gives the following equation

$$\varepsilon_{CY}^A = \varepsilon_{CY} + \sum_G \varepsilon_{CG} \varepsilon_{GY} \tag{7}$$

which is the elasticity of cost with respect to output adjusted for the presence of public quasi-fixed inputs. Note that ε_{CG} is the elasticity of cost with respect to external factors, and ε_{GY} is the elasticity of "demand for external factors" with respect to output. This demand elasticity should be interpreted as a long-run one representing the change in external factors necessary to maintain the firm on the envelope long-run average cost curve after a change in output. Therefore, if ε_{CY}^A is less than one, then there are increasing returns to scale over all inputs.

Empirical Implementation

This section presents the empirical implementation of the model introduced above. The contribution of public capital and public R&D to U.S. agricultural growth and the conditions for the hypothesized endogenous growth can still be tested through estimation of the firms' cost and demands for private inputs. Adopting a flexible functional form for the value function of the firms, all parameters of interest can be recovered from the estimation of the dynamic demands for private quasi-fixed inputs and the demand for the variable input.

The study covers the period 1948 – 1994. Variables needed for estimation include quantity indexes of capital (K), labor (L), materials (M), and output (Y); implicit prices of the three inputs; and stocks of quasi-fixed public inputs (public capital (G) and R&D (R)).[8] K is an aggregate measure of capital and land. Capital and labor are assumed quasi-fixed inputs, while materials are the only variable input.[9] Output is an index of all crops and livestock products. Public capital stocks are values of federal, state, and local structures. Public R&D stocks are constructed from R&D spending using Chavas and Cox's method (1992).[10]

Consider the following normalized quadratic value function:

[8] See Ball et al. (1997) for details on all agricultural data. Public capital stocks are from Survey of Current Business and include buildings, highways, streets, sewer structures etc. Military structures are excluded. Public R&D spending is from Alston and Pardey (1996).

[9] The adoption of materials as a variable factor in agricultural production is consistent with the findings of previous studies, for example, Vasavada and Chambers (1986).

[10] With this method, the stock for a given year is constructed as a weighted sum of the last thirty years of expenditures, in which the weights follow an inverted 'V' pattern. Huffman and Evenson's (1989) methodology, which consists of a trapezoidal pattern of thirty-five years of expenditures, was also tried. Results show no significant differences.

$$J(\mathbf{P},\mathbf{Z},\mathbf{Q}) = a_0 + \left[\mathbf{A}_{\mathbf{P}}' \ \mathbf{A}_{\mathbf{Z}}' \ \mathbf{A}_{\mathbf{Q}}'\right]\begin{bmatrix}\mathbf{P}\\ \mathbf{Z}\\ \mathbf{Q}\end{bmatrix} + \frac{1}{2}\left[\mathbf{P}' \ \mathbf{Z}' \ \mathbf{Q}'\right]\begin{bmatrix}\mathbf{B}_{\mathbf{PP}} & \mathbf{B}_{\mathbf{PZ}}^{-1} & \mathbf{B}_{\mathbf{PQ}}\\ \mathbf{B}_{\mathbf{PZ}}^{-1} & \mathbf{B}_{\mathbf{ZZ}} & \mathbf{B}_{\mathbf{ZQ}}\\ \mathbf{B}_{\mathbf{QP}} & \mathbf{B}_{\mathbf{QZ}} & \mathbf{B}_{\mathbf{QQ}}\end{bmatrix}\begin{bmatrix}\mathbf{P}\\ \mathbf{Z}\\ \mathbf{Q}\end{bmatrix} \qquad (11)$$

This is a second order Taylor series expansion of J in (**P**, **Z**, **Q**), where **Z** is the vector of quasi-fixed factors, **P** is the corresponding vector of normalized rental prices, and **Q** is the vector of output and public inputs; $\mathbf{A}_i$ and $\mathbf{B}_{ij}$ are parameter matrices of appropriate order; a_0 is a scalar parameter. Then, the vectors **P'**, **Z'**, and **Q'** are equal to

$$\mathbf{P}' = \left[P_K \ \ P_L\right]; \quad \mathbf{Z}' = \left[K \ \ L\right]; \quad \mathbf{Q}' = \left[G \ \ R \ \ Y\right];$$

where P_K and P_L are the prices of capital and labor, respectively, normalized by the price of materials.

The dynamic demands for quasi-fixed inputs are then[11,12]

$$\dot{\mathbf{Z}}^*(\mathbf{P},\mathbf{Z},\mathbf{Q}) = J_{pz}^{-1}(\mathbf{P},\mathbf{Z},\mathbf{Q})[\rho J_p(\mathbf{P},\mathbf{Z},\mathbf{Q}) - \mathbf{Z}] \qquad (12a)$$

and the demand for the variable input (X*) is calculated from

$$X^*(P,Z,Q) = \rho[J(P,Z,Q) - J_p'(P,Z,Q)P] - [J_Z'(P,Z,Q) - p'J_{PZ}]\dot{Z}^* \qquad (13a)$$

In terms of the postulated value function, (12a) and (13a) become

$$\dot{\mathbf{Z}}^* = \rho\mathbf{H} + (\rho\mathbf{u} - \mathbf{B}_{\mathbf{PZ}})\mathbf{Z} + \rho\mathbf{B}_{\mathbf{PZ}}\mathbf{B}_{\mathbf{PP}}\mathbf{P} + \rho\mathbf{N}\mathbf{Q} \qquad (12b)$$

$$X^* = \rho[a_0 - \frac{\mathbf{1}}{\mathbf{2}}\mathbf{P}'\mathbf{B}_{\mathbf{PP}}\mathbf{P} + \mathbf{A}_{\mathbf{Z}}'\mathbf{Z} + \mathbf{A}_{\mathbf{Q}}'\mathbf{Q} + \frac{\mathbf{1}}{\mathbf{2}}\mathbf{Z}'\mathbf{B}_{\mathbf{ZZ}}\mathbf{Z} + \mathbf{Z}'\mathbf{B}_{\mathbf{ZQ}}\mathbf{Q} + \\ + \frac{\mathbf{1}}{\mathbf{2}}\mathbf{Q}'\mathbf{B}_{\mathbf{QQ}}\mathbf{Q}] - [\mathbf{A}_{\mathbf{Z}}' + \mathbf{Z}'\mathbf{B}_{\mathbf{ZZ}} + \mathbf{Q}'\mathbf{B}_{\mathbf{ZQ}}']\dot{\mathbf{Z}}^* \qquad (13b)$$

[11] To clarify notation, note that only subscripts in the value function J denote gradient vectors. B and A are matrices of parameters.
[12] See Appendix 1 for derivation.

where $\mathbf{H} = \mathbf{B}_{\mathbf{PZ}}\mathbf{A}_{\mathbf{P}}$, $\mathbf{N} = \mathbf{B}_{\mathbf{PZ}}\mathbf{B}_{\mathbf{PQ}}$, and $\mathbf{u}$ is a 2x2 identity matrix. Note that equations (12b) constitute the flexible accelerator with constant adjustment coefficients and can be rewritten as in Equation (4)

$$\dot{\mathbf{Z}}^*(\mathbf{Z}, \mathbf{P}, \mathbf{Q}) = \mathbf{M}[\mathbf{Z} - \bar{\mathbf{Z}}(\mathbf{P}, \mathbf{Q})] \tag{4b}$$

where

$$\mathbf{M} = (\rho\mathbf{u} - \mathbf{B}_{\mathbf{PZ}})$$

$$\bar{\mathbf{Z}} = -(\rho\mathbf{u} - \mathbf{B}_{\mathbf{PZ}})^{-1}\rho[\mathbf{H} + \mathbf{B}_{\mathbf{PZ}}\mathbf{B}_{\mathbf{PP}}\mathbf{P} + \mathbf{N}\mathbf{Q}]$$

being $\bar{\mathbf{Z}}$ the steady state values of private quasi-fixed inputs.

The model until now has been described in terms of continuous time. For estimation purposes, however, a discrete approximation to $\dot{\mathbf{Z}}$ must be used. Being $\mathbf{Z}_{-1}$ the lag of $\mathbf{Z}$, (12b) can be expressed as

$$\mathbf{Z} = \rho\mathbf{H} + (\mathbf{u} + \rho\mathbf{u} - \mathbf{B}_{\mathbf{PZ}})\mathbf{Z}_{-1} + \rho\mathbf{B}_{\mathbf{PZ}}\mathbf{B}_{\mathbf{PP}}\mathbf{P} + \rho\mathbf{N}\mathbf{Q} \tag{12c}$$

Joint estimation of (12c) and (13b) gives all the parameters needed for testing the effects of public inputs on firms' costs, steady state stocks of capital, and scale.[13,14]

---

[13] This estimation assumes that farmers expect the current input prices to prevail in the future. In this way, optimization plans are revised each period when new information is obtained (i.e., when farmers observe the new prices).

[14] Note that the theory presented here is a theory of the firm. Nevertheless, the data used for estimation is highly aggregated. Consistent linear aggregation would require

$$J(P, Z, Y, G, R) = \sum_i J(P, Z_i, Y_i, G, R),$$

$$Z = \sum_i Z_i, \text{ and } Y = \sum_i Y_i$$

where the sum is across firms. The linear aggregation is over private quasi-fixed stocks and output because they are different across firms. For public inputs, however, this is not required because they are non-rival by definition: the same input (as long as they are not local public goods) can be used by many producers at the same time. Hence, for the quadratic value function presented above, consistent aggregation across firms requires linearity in Z and Y, i.e., $J_{ZZ} = B_{ZZ} = 0$, $J_{ZY} = B_{ZY} = 0$, and $J_{YY} = b_{yy} = 0$, where B_{ZY} is a partition matrix of $B_{ZQ} = [B_{ZG}\ B_{ZR}\ B_{ZY}]$, and b_{yy} is one element of B_{QQ}. For the estimation presented below

Results

With three private inputs, estimation of the system (12c)-(13b) implies joint estimation of three equations: two dynamic demands (for labor and capital) and the demand for the variable input. Additionally, the theoretical model implies that public inputs are simultaneously determined by **P**, **Z** and Y. Therefore, instrumental variables for the public inputs must be used. Accordingly, predicted values of G and R were then adopted for estimation of (12c)-(13b) by iterative nonlinear seemingly unrelated regressions (nonlinear ITSUR).[15]

Table 1a presents the parameter estimates. The necessary conditions presented in Appendix 1 and other regularity conditions of the economic theory can be checked using the parameter estimates. The list of conditions include: conditions (B), long-run demand for inputs that are decreasing in their own prices and increasing in output, and positive shadow prices of public inputs (monotonicity condition in public inputs). Some of the conditions, like concavity of the quadratic value function and stability adjustments, can be directly tested and checked if they are satisfied globally. Others, however, have to be checked locally at each data point.

Nonlinear ITSUR estimates imply that concavity of the value function holds, i.e., the matrix $\mathbf{B_{pp}}$ is negative semidefinite. The stability requirement is also satisfied, i.e., the eigenvalues of $(\mathbf{u} + \rho\mathbf{u} - \mathbf{B}_{\mathbf{PZ}})$ were inside the unit circle. In terms of the rest of the conditions, conditions (B2)(i) and (B2)(ii) for capital exhibit four and thirty-six

aggregation conditions were not imposed. When those conditions are imposed, there is no qualitative change in the results.

violations, respectively. While violation of (B2)(i) means that the Euler equation does not hold for capital, violation of (B2)(ii) implies violation of the adjustment cost condition for that input. Additionally, (B2)(iii) is not satisfied for one observation (year 1983), implying negative estimated marginal cost for that year. All these condition violations mean that the parameter estimates in Table 1A are not consistent with the dynamic theory of the firm.

Estimated shadow prices of public infrastructure and public R&D by decade are presented in Table 2A. A positive shadow price implies that the corresponding public input reduces agricultural costs of production. While positive shadow prices of public research were obtained for the whole sample period, shadow prices of infrastructure were all negative. Hence, the monotonicity condition on public infrastructure is not satisfied, which contradicts the assumption of rational government behavior in the provision of this public good.

In order to obtain reliable estimates consistent with the economic theory of the firm, new estimations imposing the set of required conditions were done. Those restrictions imply the local imposition of inequality constraints, that is, the restrictions must be imposed at each data point.[16] One way of doing this is by using Bayesian estimations to introduce the desired conditions as prior beliefs.

Bayesian estimation entails calculation of the joint posterior distribution of the parameters. Analytical calculation of that distribution is, however, not possible, and sampling algorithms are generally used to simulate that joint posterior distribution.

---

[15] Instruments include total U.S. population, number of non-farm workers, interest rate of federal bonds, and total non-agricultural exports.

Recently, different algorithms have been developed. This study follows the Metro-Hastings (MH) algorithm, a Markov Chain Monte Carlo (MCMC) simulation method that has already been used by Griffiths et al. (1999) and O'Donnell et al. (1999) in previous empirical economic studies.

As other sampling algorithms, the MH simulation method consists of generating draws of the parameters of interest from their conditional distribution. Because some restrictions want to be imposed in this case, the algorithm contains an accept–reject step in which new draws are included in the sample if those conditions are satisfied. In this way, the estimation is constrained to the parameter space that is consistent with the economic theory. Additionally, iterations characterized the process in which each random draw is conditioned on the last draw. After a certain number of iterations, that process converges to a random sample from the joint posterior distribution. The MH parameter estimates are then the mean of that random sample.[17]

The MH estimation was first done imposing the required conditions on all data points. In this case, no draw satisfying all the conditions could be obtained, i.e. the parameter space was empty. The conditions were then relaxed and, due to potential measurement errors, they were required to be satisfied only at 80% percent of the observations (Atkinson and Dorfman (2001)). Since this relaxation was not enough to get a nonempty set, the conditions implied by the Euler equations and adjustment costs were

[16] Diewert and Wales (1987) show that, to impose those conditions globally, non-flexible functional forms must be adopted.

[17] Appendix 3 presents a brief description of the MH algorithm for Bayesian estimation. A detailed explanation is presented in Griffiths et al. (1999) and O'Donnell et al. (1999).

not imposed.[18] Table 1B shows then the MH parameter estimates without imposing these two conditions. Additionally, convergence of the MH algorithm to the joint posterior distribution has been rejected based on the convergence diagnostic developed by Geweke (1992).[19]

Using these MH estimates, the value function was negative in the last seven years of the sample (i.e., condition (B1) was satisfied in more than 80% of the cases). The Euler equation for capital was not satisfied in the first five years, while, for labor, it was not satisfied in the first seventeen years. The adjustment-cost condition was not satisfied in thirty-four years for capital and forty-three years for labor. In contrast, positive shadow prices of public inputs were obtained for all data points. It seems, therefore, that forcing the parameter estimates to satisfy monotonicity in public inputs makes them difficult to satisfy the Euler equations and the adjustment-cost conditions.

Table 2B presents the shadow prices of public infrastructure and research by decade and their respective standard deviations. The shadows are positive for all decades and most of them are significantly different from zero. This could be interpreted as a positive contribution of public inputs to productivity growth of the US agricultural sector.

Tables 3B to 5B show the short- and long-run elasticities of demand and the elasticities of cost with respect to output. Long-run elasticities of demand would indicate that, while infrastructure has had positive impacts on private capital accumulation, public research has substituted private capital. Finally, elasticities of cost with respect to output, even after adjusting for the presence of public inputs, are larger than one, meaning that

---

[18] This was determined by trial and error examination of the conditions. It was found that the Euler equation and adjustment cost conditions were the conditions more difficult to be satisfied.

[19] See Appendix 3 for details.

the US agricultural sector has exhibited decreasing returns to scale, contrary to the postulated endogenous growth condition.

All these results, however, cannot be taken as conclusions of this study given that the estimated parameters are not consistent with the economic theory and they cannot be interpreted as being drawn from the posterior distribution due to no convergence of the iterations. Reliability of the data, measurement errors, and aggregation biases could be named as possible reasons for those conditions not to be satisfied. Therefore, either the non-sample (prior) information (i.e., the restrictions concerning the microeconomic theory of the firm) is not correct or the sample information (the available data set of the U.S. agricultural sector) is not enough information to successfully obtain the posterior distribution.

V. Conclusions

This chapter has presented a dynamic model to measure the contribution of public inputs to productivity growth. It has also shown testable hypotheses related to the main postulates of a version of endogenous growth theory (‘AK’ models with public goods) using duality theory. In particular, two conditions have been postulated and tested. One is the existence of increasing returns to scale over all inputs (private and public). The other is the positive effect of public inputs on the long-run demand for private factors that can be accumulated (steady state stocks of capital).

The estimates presented in this study do not meet the conditions implied by the economic theory of the firm and the assumed rational behavior of government. Numerous reasons could be mentioned for this, like data reliability, measurement errors, aggregation biases, and model identification. Related to this last reason is the possibility that the U.S. agricultural sector has experienced technological changes that have been too fast for this empirical model to correctly capture them. Independently of the reason, Bayesian estimation allows concluding that either the prior information (i.e., the restrictions concerning the microeconomic theory of the firm) is not correct or the sample information (the available data set of the U.S. agricultural sector) is not enough information to successfully obtain the posterior distribution.

Finally, more work is needed to overcome the limitations of this study. The use of time series may cause problems due to the presence of nonstationary data. One way of overcoming this problem is to consider a cointegration approach. However, the large number of parameters to estimate, relative to the sample size, makes this task difficult. Another alternative approach is the use of panel data, estimating the model at the state level. This approach can improve this study in both theoretical and econometrical aspects. In terms of theoretical aspects, panel estimation at the state level could also allow for the presence of spillover effects as well as different patterns of growth in each state. In terms of econometrical aspects, a larger number of degrees of freedom is introduced, which, as is well known, improves statistical estimations. Clearly, a model introducing panel data to this dynamic duality model is the direction to follow.

References

Alston, J. and P. Pardey, *Making Science Pay*, AEI Press, 1996.

Antle, J. "Infrastructure and Aggregate Agricultural Productivity: International Evidence." *Economic Development and Cultural Change* 31 v3 (1983): 609-19.

Aschauer, D. A. "Is Public Expenditure Productive?." *Journal of Monetary Economics* 23 (March 1989): 177-200.

Ball, E., J. Bureau, R. Nehring, and A. Sonwaru. "Agricultural Productivity Revisited," *Amer. J. Agr. Econ.* 79 (November 1997): 1045-63.

Barro, R., "Government Spending in a Simple model of Endogenous Growth." *Journal of Political Economy*, 98 (October1990), Part II: S103-S125.

Barro, R., and X. Sala-I-Martin. *Economic Growth*, McGraw-Hill, 1995.

Barten, A. "Maximum Likelihood Estimation of a Complete System of Demand Equations" European Economic Review, 1 (1969): 7-73.

Berndt, E. and B. Hansson. "Measuring the Contribution of Public Infrastructure Capital in Sweden." *Scandinavian Journal of Economics* 94 (Supplement 1992): s151-72.

Binswanger, H., S. Khandker and M. Rosenzweig, "How infrastructure and financial institutions affect agricultural output and investment in India." *Journal of Development Economics* 41 (1993): 337-66.

Chalfant, J. and K. White. "Estimation and Testing in Demand Systems with Concavity Constraints." Working Paper. Department of Agricultural and Resource Economics, University of California at Berkeley, 1987.

Chalfant, J., R. Gray and K. White. "Evaluating Prior Beliefs in a Demand System: The Case of Meats Demand in Canada." Working Paper. Department of Agricultural and Resource Economics, University of California at Berkeley, 1989.

Chambers, R. *Applied Production Analysis - The Dual Approach*. Cambridge University Press 1988.

Craig, B., P. Pardey, and J. Roseboom. "International Productivity Patterns: Accounting for Input Quality, Infrastructure, and Research." *Amer. J. Agr. Econ.* 79 (November 1997): 1064-76.

Epstein, L. "The Multivariate Flexible Accelerator Model: Its Empirical Restrictions and an Application to U.S. Manufacturing." *Econometrica*, Vol. 51, No 3 (May, 1983): 647-674.

Epstein, L. "Duality Theory and Functional Forms for Dynamic Factor Demands." *Review of Economic Studies*, 48 (1981): 81-95.

Garcia-Mila, T. and McGuire, T. "The Contribution of Publicly Provided Inputs to States' Economies." *Regional Science and Urban Economics* 22 (June 1992): 229-42.

Geweke, J. "Exact Inference in the Inequality Constrained Normal Linear Regression Model." Journal of Applied Econometrics, 1 (1986): 127-141.

Holtz-Eakin, D. "Public Sector Capital and the Productivity Puzzle." *Review of Economics and Statistics* 76 (February 1994): 12-21.

Huffman, W. and R. Evenson, "Supply and Demand Functions for Multiproduct US Cash Grain Farms: Biases Caused by Research and Other Policies." *Amer. J. Agr. Econ.* 71 (August 1989): 761-73.

Lynde, C. and J. Richmond. "The Role of Public Capital in Production." *Review of Economics and Statistics* 74 (February 1992): 37-44.

Morrison, C. and E. Schwartz. "State Infrastructure and Productive Performance." *The*

American Economic Review 86 (December 96): 1095-1111.

Munnell, A. "How Does Public Infrastructure Affect Regional Economic Performance?." *New England Economic Review* (Sept./October 1990): 11-32.

Nadiri, I. and T. Mamuneas. "The Effects of public Infrastructure and R&D Capital on the Cost Structure and Performance of U.S. Manufacturing Industries." *Review of Economics and Statistics* 76 (February 1994): 22-37.

Romer, P. "Increasing Returns and Long-Run Growth." *Journal of political Economy* 94 (October 1986): 1002-37.

Romer, P. "Endogenous Technological Change." *Journal of political Economy* 98 Part II (October 1990): S71-S102.

Vasavada, U. and R. Chambers. "Investment in U.S. Agriculture." *Amer. J. Agr. Econ.* 68 (November 1986): 950-60.

APPENDIX 1

This appendix presents conditions (A) and (B) that guarantee duality between cost and value functions of the firms.

Conditions (A)

It is assumed that C(y, Z, I; G) satisfies the following set of regularity conditions:

(A.1) C(y, Z, I; G) ≥ 0.

(A.2) C(y, Z, I; G) is increasing in y and decreasing in Z. Additionally, $C_I > 0$ when $I > 0$ and vice versa, which follows from the assumption of adjustment costs.

(A.3) C(y, Z, I; G) is convex in I.

(A.4) For each (Z_0, y, p; G) a unique solution exists for (1). This means that there are well-defined factor demand functions associated with (1).

(A.5) For each (Z_0, y, p; G), problem (1) has a unique steady state (SS) stock $\bar{Z}(y, p; G)$ that is globally stable. This condition establishes the uniqueness and stability of the steady state.

(A.6) For any (Z_0, y, p; G), there exists $\hat{p}$ such that $\hat{I}$ is the optimal gross investment vector at t = 0 in (1) given (Z_0, y, p; G).

Conditions (B)

It is assumed that the value function J(Z, y, p; G) satisfies the following properties:

(B.1) J(Z, y, p; G) ≥ 0.

(B.2) (i) $(ru + \delta)J_z(Z, y, p; G) - p - J_{zz}(Z, y, p; G)\dot{Z}^*(Z, y, p; G) < 0$, where u is an identity matrix. This expression is dual to $C_z < 0$.

(ii) $J_z(Z, y, p; G) < 0$ when $I^*(Z,y,p;G) \equiv \dot{Z}^*(Z,y,p;G) + \delta Z > 0$ and vice versa.

This condition is dual to $C_I > 0$ when $I > 0$ and vice versa.

(iii) $\rho J_y(Z, y, p; G) - J'_{yz}(Z, y, p; G)\dot{Z}^*(Z, y, p; G) > 0$, where

$\dot{Z}^*(Z,y,p;G) = J_{pz}^{-1}(Z, y, p; G)[\rho J_p(Z, y, p; G) - Z]$. This condition is dual to $C_y > 0$.

(B.3) The following expression is concave in p:

$$\rho J(Z,y,p;G) - p'Z - J'_z(Z,y,p;G)\dot{Z}^*(Z,y,p;G)$$

Under some specific functional forms (like the normalized quadratic presented above), $J_z(Z, y, p; G)$ is linear in p and the curvature requirement reduces to concavity of J(Z, y, p; G) in p. This condition is dual to (A.3).

(B.4) The demand for the variable input, $X^*(Z, y, p; G)$, is positive.

(B.5) The stock Z that solves $\dot{Z}^*(Z, y, p; G) = J_{pz}^{-1}(Z, y, p; G)[\rho J_p(Z, y, p; G) - Z]$, with $Z(0) > 0$, has a unique globally stable steady state $\bar{Z}(y,p;G)$.

Then, under conditions (A) and (B), duality between C(y, Z, I; G.) and J(Z, y, p; G) can be established as in equations (2) and (3). The following derivative properties then hold:

Derivative Properties

1. With respect to I:

   $C_I(y, Z, I; G) = - J_z(Z, y, p; G)$. From (A.2) or (B.2.ii), this expression must be positive when I > 0 and vice versa.

Testing for $J_z(Z, y, p; G) = 0$ is equivalent to testing for adjustment costs in inputs Z.

2. With respect to Z:

$$C_z(y, Z, I; G) = (\rho u + \delta)J_z(Z, y, p; G) - p - J_{zz}(Z, y, p; G)\dot{Z}^*(Z, y, p; G) < 0 \text{ from (A.2).}$$

This expression gives the shadow price of quasi-fixed inputs.

3. With respect to y:

$$C_y(y, Z, I; G) = \rho J_y(Z, y, p; G) - J'_{zy}(Z, y, p; G)\dot{Z}^*(Z, y, p; G) > 0 \text{ from (A.2).}$$

This expression represents the output supply of the firms.

4. With respect to p:

$$0 = \rho J_p(Z, y, p; G) - Z - J_{zp}(Z, y, p; G)\dot{Z}^*(Z, y, p; G)$$

Then,

$\dot{Z}^*(Z, y, p; G) = J_{pz}^{-1}(Z, y, p; G)[\rho J_p(Z, y, p; G) - Z]$, which is the dynamic demand for Z.

5. With respect to G:

$$C_G(y, Z, I; G) = \rho J_G(Z, y, p; G) - J_{ZG}(Z, y, p; G)\dot{Z}^*(Z, y, p; G)$$

This expression represents the shadow price of G when the firms are out of the SS. At the SS, the shadow price is

$$C_G(y, Z, I; G) = \rho J_G(Z, y, p; G)$$

If this expression is negative, the shadow price of G is positive, meaning that public inputs reduce cost of production.

APPENDIX 2

This appendix presents conditions (C) to (D) that guarantee duality between the value function of the firms and the value function of the government.

Conditions (C)

It is assumed that $J(y, Z, p; G) + AC(I_g)$ satisfies the following conditions:

(C.1) $J(y, Z, p; G) + AC(I_g) \geq 0$

(C.2) (i) $J(y, Z, p; G) + AC(I_g)$ is increasing in I_g. Given that $J(y, Z, p; G)$ is independent of I_g, $AC(I_g)$ must be increasing in I_g.

(ii) $J(y, Z, p; G) + AC(I_g)$ is decreasing in G. Given that $AC(I_g)$ is independent of G, $J(y, Z, p; G)$ must be decreasing in G.

(C.3) $J(y, Z, p; G) + AC(I_g)$ is convex in I_g. Then, $AC(I_g)$ must be convex in I_g.

(C.4) For each (Z, p, y, r, G_0), there exists a unique solution for (8). This means that there are well-defined supplies of public inputs.

(C.5) For each (Z, p, y, r, G_0), (8) has a unique steady state stock $\bar{G}(Z,p,y,r)$ that is globally stable.

(C.6) For any (Z, p, y, r, G_0), there exists $\hat{r}$ such that $\hat{I}_g$ is the optimal public gross investment vector at t = 0 in (8), given (Z, p, y, r, G_0).

Conditions (D)

It is assumed that $J^g(y, Z, p; r, G)$ satisfies the following conditions:

(D.1) $J^g(y, Z, p; r, G) \geq 0$

(D.2) (i) $J^g_G(y, Z, p; r, G) < 0$. This condition is dual to (C.2)(i) and means that there are adjustment costs in the provision of public inputs.

(ii) $(\theta u + \delta_g)J^g_G(y, Z, p; r, G) - J^g_{GG}(y, Z, p; r, G)\dot{G}^* < 0$. This expression is dual to (C.2)(ii): $J_G(y, Z, p; r, G) < 0$ (positive shadow prices of public inputs). Given $J^g_G(y, Z, p; r, G) < 0$, it is sufficient for (D.2)(ii) to hold that $-J^g_{GG}(y, Z, p; r, G) < 0$ (that is, increases of the public good decrease the shadow price of it).

(D.3) $\theta J^g(Z, y, p, r, G) - r'G - J^g_G{}'(Z, y, p, r, G)\dot{G}^*$ must be concave in r. This is dual to condition (C.3).

(D.4) $I^*_g(y, Z, p; r, G) \equiv \dot{G}^*(y, Z, p; r, G) + \delta_g G$ is positive.

(D.5) The stocks G that solve

$$\dot{G}^*(y, Z, p; r, G) = J^{g-1}_{Gr}(y, Z, p; r, G)[\theta J^g_r(y, Z, p; r, G) - G],$$

with G(0) > 0, has a unique globally stable steady state $\bar{G}(Z, p, y; r)$.

Then, under conditions (C) and (D), duality between J^g(y, Z, p; r, G) and J(y, Z, p; r, G) +AC(I_g) can be established as in equations (9) and (10). The derivative properties presented below then hold.

Derivative Properties

1. With respect to I_g:

$$0 = AC_{I_g} + J^g_G(p, Z, y; r, G)$$

or

$$-J^g_G(p, Z, y; r, G) = AC_{I_g} > 0\,,$$

This is positive given $AC_{Ig} > 0$.

2. With respect to G:

$$\theta J^g_G(p, Z, y; r, G) = J_G(Z, y, p; G) + r + J^g_{GG}(p, Z, y; r, G)\dot{G}^* - \delta_g J^g_G(p, Z, y; r, G)$$

or

$$J_G(Z, y, p; G) = (\theta u + \delta_g)J^g_G(p, Z, y; r, G) - r - J^g_{GG}(p, Z, y; r, G)\dot{G}^*$$

This expression is the firms' willingness to pay for G (shadow price) when the firms are at the steady state. If the expression is negative (condition (D.2)(ii)), then public inputs reduce cost of production. When the government is also at the SS, that expression can be rewritten as

$$-J^g_G(p, Z, y; r, G) = (\theta u + \delta_g)^{-1}(-J_G(Z, y, p; G) - r)$$

which could be interpreted as a 'social' shadow price: the net social benefit (the firms' shadow price of G minus the government's cost of providing G) adjusted by the 'social' discount rate plus the depreciation rate of public inputs.

3. With respect to r:

$$\theta J^g_r(p, Z, y; r, G) = G + J^g_{Gr}(p, Z, y; r, G)\dot{G}^*$$

or

$$\dot{G}^{*} = J_{Gr}^{g-1}(p, Z, y; r, G)[\theta J_{r}^{g}(p, Z, y; r, G) - G]$$

which gives the optimal path of G.

4. With respect to Z:

$$\theta J_{z}^{g}(p, Z, y; r, G) = J_{z}(Z, y, p; G) + J_{Gz}^{g}(p, Z, y; r, G)\dot{G}^{*}$$

or

$$J_{z}(Z, y, p; G) = \theta J_{z}^{g}(p, Z, y; r, G) - J_{Gz}^{g}(p, Z, y; r, G)\dot{G}^{*} < 0$$

where the sign is given by condition B.2(ii): the value function of the firm is decreasing in Z.

5. With respect to y:

$$\theta J_{y}^{g}(p, Z, y; r, G) = J_{y}(Z, y, p; G) + J_{Gy}^{g}{}'(p, Z, y; r, G)\dot{G}^{*}$$

or

$$J_{y}(Z, y, p; G) = \theta J_{y}^{g}(p, Z, y; r, G) - J_{Gy}^{g}{}'(p, Z, y; r, G)\dot{G}^{*} > 0$$

where the sign is given by condition B.2(iii): the value function of the firm is increasing in y. Finally, at the SS level of G (or with no adjustment cost of G),

$$J_{y}(Z, y, p; G) = \theta J_{y}^{g}(p, Z, y; r, G) > 0$$

Appendix 3

Bayesian Estimation

In Bayesian statistics, the parameters to be estimated are treated as random variables associated with a subjective probability distribution that describes the state of knowledge about the parameters. The knowledge either may exist before observing any sample information or might be derived from both prior and sample information. In the former case, the associated probability distribution is a prior distribution. In the latter, that distribution is a posterior distribution. Thus, different from the classical statistics that concentrates on point estimates of a (set of) parameter(s), the objective of Bayesian statistics is usually the achievement of the posterior distribution of a (set of) parameter(s).

Using the Bayesian Theorem, it can be shown that the joint posterior distribution of a set of parameters can be obtained from the combination of sample information and the joint prior distribution of the parameters (see Judge et al., 1988). That is,

$$f(\boldsymbol{\beta}, \boldsymbol{\Sigma} / \mathbf{y}) \propto L(\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) p(\boldsymbol{\beta}, \boldsymbol{\Sigma}) \tag{A3.1}$$

where $\boldsymbol{\beta}$ is the vector of parameters of interest, $\boldsymbol{\Sigma}$ is their variance-covariance matrix, and $\mathbf{y}$ is the matrix of sample observations. Expression (A.3.1) states that the posterior joint density function of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ (i.e., $f(\boldsymbol{\beta}, \boldsymbol{\Sigma} / \mathbf{y})$) is proportional to (‘$\propto$’) the likelihood function $L(\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\Sigma})$ (which contains all the sample information) times the prior density function $p(\boldsymbol{\beta}, \boldsymbol{\Sigma})$. Intuitively, the prior information about the parameters is modified by the available sample information (through the likelihood function) to obtain the posterior information about the parameters.

Attainment of the posterior distribution requires specification of the prior distribution and the likelihood function. In the present study, the prior distribution must incorporate the restrictions implied by the economic theory. That is, the assumptions required for the adopted economic theory to be true are included as prior beliefs. The prior distribution must then assign probability zero to regions of the parameter space that do not meet the restrictions and positive probability otherwise. Following O'Donnell et al. (1999), the following *non-informative*[20] joint prior distribution was adopted

$$\mathrm{p}(\boldsymbol{\beta},\boldsymbol{\Sigma}) = \mathrm{p}(\boldsymbol{\beta})\mathrm{p}(\boldsymbol{\Sigma})\,\mathrm{I}(\boldsymbol{\beta}\in \mathrm{B}) \propto \left|\boldsymbol{\Sigma}\right|^{-\frac{N+1}{2}} I(\boldsymbol{\beta}\in \mathrm{B}) \tag{A3.2}$$

where B is the subspace of parameter vectors that satisfy the restrictions, N is the number of equations, and $I(\beta \in B)$ is an indicator function that takes the value 1 when a given vector of parameters β belongs to B and takes the value 0 otherwise.

The SUR model to be estimated is $\mathbf{y} = \mathrm{g}(\mathbf{X},\boldsymbol{\beta}) + \varepsilon$, where the N equations have been stacked. To specify the likelihood function, then, a distribution for ε must be assumed. Following Judge et al. (1985), it is assumed that the error vector has a multivariate normal distribution, i.e. $\varepsilon \sim \mathrm{MVN}(0, \Sigma \otimes \mathrm{I}_T)$. The likelihood function is then

$$\begin{aligned} \mathrm{L}(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Sigma}) &\propto \left|\boldsymbol{\Sigma}\right|^{-T/2} \exp[-.5(\mathbf{y} - \mathrm{g}(\mathbf{X},\boldsymbol{\beta})'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_{\mathrm{I}})(\mathbf{y} - \mathrm{g}(\mathbf{X},\boldsymbol{\beta})] \\ &\propto \left|\boldsymbol{\Sigma}\right|^{-T/2} \exp[-.5\,\mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}^{-1})] \end{aligned} \tag{A3.3}$$

where $\mathbf{A}$ is the N x N symmetric matrix with the $(i,j)^{th}$ element equal to $a_{ij} = (\mathbf{y}_i - \mathrm{g}(\mathbf{X}_i,\beta))'(\mathbf{y}_j - \mathrm{g}(\mathbf{X}_j,\beta))$.

[20] A non-informative prior distribution is a distribution that does not contain specific information about the parameter. Equation (A3.2), form example, does not specify any exact distribution $p(\beta,\Sigma)$. If $p(\beta,\Sigma)$ is said to be a normal distribution, in contrast, then that would be an informative prior distribution. As a consequence of adopting a non-informative prior, when this prior is combined with the likelihood function in equation (A3.1), the posterior distribution is dominated by the sample information.

Finally, having specified both the prior distribution and the likelihood function, the joint posterior distribution is

$$f(\boldsymbol{\beta}, \boldsymbol{\Sigma} / \mathbf{y}) \propto \left|\boldsymbol{\Sigma}\right|^{-(T+N+1)/2} \exp[-.5(\mathbf{y} - g(\mathbf{X}, \boldsymbol{\beta})'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_T)(\mathbf{y} - g(\mathbf{X}, \boldsymbol{\beta})] \, I(\boldsymbol{\beta} \in B) \quad \text{(A3.4)}$$

As in O’Donnell et al. (1999), the interest is on the characteristics of the marginal distribution of $\boldsymbol{\beta}$ and, then, $\boldsymbol{\Sigma}$ is considered a nuisance parameter that can be integrated out of equation (A3.4). This procedure yields

$$f(\boldsymbol{\beta} / \mathbf{y}) \propto \left|\mathbf{A}\right|^{-T/2} \; I(\boldsymbol{\beta} \in B) \qquad \text{(A3.5)}$$

Additional integration of this joint posterior distribution would eventually give the marginal distribution of β. However, this is not analytically possible. The way this is overcome is by adopting numerical methods. In particular, computer-intensive algorithms can be implemented in the estimation of the marginal distribution.

Markov Chain Monte Carlo (MCMC) algorithms, like the Gibbs sampler and the Metropolis-Hastings (MH) algorithm, have recently become very popular in applied economics. They constitute a technique for generating random variables from a marginal distribution without need of analytically calculating the density (Cassella and George, 1992). That is, instead of computing or approximating the posterior distribution $f(\boldsymbol{\beta} / \mathbf{y})$ directly, those algorithms allow generating a sample $\beta^1, \ldots, \beta^m \sim f(\boldsymbol{\beta} / \mathbf{y})$ without knowing $f(\boldsymbol{\beta} / \mathbf{y})$. The characteristics of the marginal density can then be calculated with a large enough sample. For example, the mean of $f(\boldsymbol{\beta} / \mathbf{y})$ is calculated using the sample mean (Cassella and George, 1992)

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} \boldsymbol{\beta}^i = \int_{-\infty}^{\infty} \boldsymbol{\beta} f(\boldsymbol{\beta} / \mathbf{y}) d\boldsymbol{\beta} = E[\boldsymbol{\beta}] \qquad \text{(A3.6)}$$

Hence, for a large enough sample size m, any population characteristic can be obtained from the generated observations.

The Bayesian results shown in this chapter are based on the Metropolis-Hastings algorithm presented in O'Donnell et al. (1999).[21] This algorithm consists of the following steps

1) An arbitrary starting value of the parameter vector $\boldsymbol{\beta}^i$, i = 0, is specified such that the constraints are satisfied; that is, $\boldsymbol{\beta}^0 \in$ B.
2) Given $\boldsymbol{\beta}^i$, a candidate value for $\boldsymbol{\beta}^{i+1}$, $\boldsymbol{\beta}^C$, is generated from a symmetric transition density q($\boldsymbol{\beta}^i$,$\boldsymbol{\beta}^C$).
3) $\boldsymbol{\beta}^C$ is used to evaluate the constraints; if any constraint is violated, then the function $\alpha(\boldsymbol{\beta}^i,\boldsymbol{\beta}^C)$ is set equal to zero and the algorithm jumps to step 5).
4) If the constraints hold, then $\alpha(\boldsymbol{\beta}^i,\boldsymbol{\beta}^C)$ is set equal to min[g($\boldsymbol{\beta}^C$)/ g($\boldsymbol{\beta}^i$), 1], where $\mathrm{g}(\boldsymbol{\beta}) = |\mathbf{A}|^{-\mathrm{T}/2}\, I(\boldsymbol{\beta} \in \mathrm{B})$ is the kernel of $\mathrm{f}(\boldsymbol{\beta}/\mathbf{y})$; that is, $\mathrm{f}(\boldsymbol{\beta}/\mathbf{y}) \propto |\mathbf{A}|^{-\mathrm{T}/2}\, I(\boldsymbol{\beta} \in \mathrm{B})$.
5) An independent uniform random variable (U) is generated from the interval [0,1].
6) The next value in the sequence, $\boldsymbol{\beta}^{i+1}$, is generated from the rule

$$\beta^{i+1} = \begin{cases} \beta^C \text{ if } U < \alpha(\beta^i, \beta^C) \\ \beta^i \text{ if } U \geq \alpha(\beta^i, \beta^C) \end{cases}$$

7) The value of i is set equal to i+1 (i = i+1), and the procedure continues in step 2.

---

[21] Detailed theoretical explanation of the Gibbs sampler and the MH algorithm is provided in Cassella and George (1992), Chib and Greenberg (1996), Gelfand et al. (1990), Gelfand et al. (1992), and Gelfand and Smith (1990). For empirical implementation, see Atkinson and Dorfman (2001), Griffiths et al. (1999), O'Donnell et al. (1999), and Terrel (1996).

Note that the symmetric transition density q(β^i,β^C) in step 2) must be specified. As in O'Donnell et al., the following multivariate normal distribution is adopted

$$q(\boldsymbol{\beta}^i, \boldsymbol{\beta}^C) = \text{MVN}\big[\boldsymbol{\beta}^i, h[X'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_T)X]^{-1}\big] \tag{A3.7}$$

where $[X'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_T)X]^{-1}$ is the estimated covariance matrix obtained in the SUR estimation, and h is a scalar used to control the size of the 'step' in the iteration (i.e., h gives the rate at which the parameter space B is investigated).

Iteration of this procedure m times gives a sequence of parameter vectors β^1,…, $\beta^m \in$ B. For some s < m, the following holds: β^{s+1}**, …,** $\beta^m \sim f(\boldsymbol{\beta}/\mathbf{y})$. In words, after a large enough sequence of size s (the 'burn-in' period), the m-s final drawings converge, in the sense that they are drawn from the distribution $f(\boldsymbol{\beta}/\mathbf{y})$. Finally, the estimated β is simply the sample mean of β^{s+1}**, …,** β^m.[22]

For estimation purposes, the starting values, the value of the scalar h, and the sizes of the 'burn-in' period (s) and the whole sample (m) had been established. The value of h was set equal to 0.025 after trying different values. The selection was based on the maximum rate at which candidate vectors were accepted as the next value in the sequence. For many alternative values of h, there was no candidate accepted as next value in the sequence.

The total number of iterations was m = 110,000. The first 10,000 were used for the 'burn-in' period (s). In terms of the starting values, although the SUR estimates do not satisfy the constraints (i.e., the estimated vector β does not belong to B), those

[22] The 'burn-in' period guarantees two characteristics of the sub-sample of size m-s used to estimate β. First, the last m-s observations are effectively drawn from f(β/**y**). Second, those observations are independent from the starting value (Chib and Greenberg, 1996).

estimates were used to draw an initial value β^0. However, after trying 100,000 draws, no vector satisfied the restrictions, implying that the SUR estimates are 'far' from the required parameter space B.

An arbitrary vector β^0 was then chosen such that these starting values satisfy the constraints. The existence of candidate vectors β^C such that $\beta^C \in$ B and $\alpha(\beta^i,\beta^C) > U$ was, however, very sensitive to changes in the starting values. That is, similar to what happened with alternative values of h, the case B = $\{\beta^0\}$ has been the result in many runs that tried different starting values. Additionally, when some restrictions on the set B were relaxed,[23] only 34 out of 100,000 iterations satisfy both $\beta^C \in$ B and $\alpha(\beta^i,\beta^C) > U$, implying that $\beta^{i+1} = \beta^i$ in almost all the iterations. Given this available data set (the sample information), it seems that the set B is very restricted and narrow. Consequently, it is difficult to get a reliable posterior distribution.

Finally, convergence of the distribution was checked by taking two sub-samples, s_1 and s_2, of the last m-s iterations and comparing their means. The sub-sample s_1 is composed by the first 10,000 observations, while s_2 is composed by the last 50,000 observations. A likelihood ratio test to compare the mean vectors of the two sub-samples was performed. Results indicate that equality of the mean vectors is rejected, meaning that the iterations have not converged. Increasing the size of the burn-in period and the total number of iterations did not change the result of the convergence test. The MH parameter estimates, then, cannot be said to characterize the marginal distribution $f(\boldsymbol{\beta} / \mathbf{y})$.

---

[23] In particular, the Euler equations and the adjustment cost conditions. See page ??.

Table 1A
Parameter Estimates

Parameter	Estimate	t Value	Parameter	Estimate	t Value
h_k	-0.05576	-0.66	a_g	-0.00003	-0.41
b_{pkk}	0.142348	3.66	a_r	-0.02087	-0.40
b_{plk}	0.16427	1.32	a_y	481.825	1.52
b_{pkL}	-0.04882	-3.76	b_{kk}	-6.57309	-0.60
b_{plL}	0.244829	5.25	b_{kL}	-15.7573	-2.78
b_{pkpk}	-2.13166	-2.21	b_{kg}	-6.87E-06	-0.80
b_{pkpl}	4.332687	1.89	b_{kr}	0.015036	1.65
b_{plpl}	-26.9784	-2.85	b_{ky}	-95.7818	-3.13
n_{kg}	8.10E-08	3.76	b_{LL}	-6.57994	-1.26
n_{kr}	-0.0001	-6.01	b_{Lg}	3.48E-06	0.27
n_{ky}	0.205884	3.10	b_{Lr}	-0.00568	-0.54
h_l	0.346815	1.39	b_{Ly}	-18.3036	-0.66
n_{lg}	-1.66E-08	-0.28	b_{gg}	1.29E-11	0.37
n_{lr}	9.21E-06	0.18	b_{gr}	-1.37E-08	-0.61
n_{ly}	-0.0178	-0.10	b_{gy}	0.000049	0.51
a_0	-268.873	-1.76	b_{rr}	5.01E-06	0.28
a_k	105.181	2.67	b_{ry}	0.019253	0.33
a_L	47.76673	1.51	b_{yy}	482.802	1.30

Equation	Adj R-Sq.	D-W
Qk	0.975	1.98
QL	0.996	2.05
Qm	0.798	1.96

Table 2A
Shadow Prices based on ITSUR Estimates
Average By Decade

Decade	P*g	P*r
1949-1959	-0.00000004	0.00086800
1960-1969	-0.00000034	0.00071460
1970-1979	-0.00000082	0.00080810
1980-1989	-0.00000129	0.00135480
1990-1994	-0.00000115	0.00141570

Table 3A
Short-Run Elasticities of Demand for Private Inputs
w/ Respect to Public Goods - ITSUR Estimates
Average By Decade

Decade	E_{KG_SR}	E_{KR_SR}	E_{LG_SR}	E_{LR_SR}	E_{MG_SR}	E_{MR_SR}
1949-1959	-0.00320	0.00163	0.00573	-0.00678	-0.19251	0.34572
1960-1969	-0.00593	0.00229	0.01569	-0.01409	-0.34502	0.49375
1970-1979	-0.01534	0.00492	0.05827	-0.04354	-3.07294	5.80969
1980-1989	-0.03504	0.01594	0.16666	-0.17651	-1.70414	5.23461
1990-1994	-0.05688	0.03242	0.26842	-0.35618	-1.49635	3.71132

Table 4A
Long-Run Elasticities of Demand for Private Inputs
w/ Respect to Public Goods - ITSUR Estimates
Average By Decade

Decade	E_{KG_LR}	E_{KR_LR}	E_{LG_LR}	E_{LR_LR}	E_{MG_LR}	E_{MR_lR}
1949-1959	-0.0052	0.0015	-0.0494	0.0582	-0.0542	-0.3770
1960-1969	-0.0093	0.0021	-0.1214	0.1086	0.0562	-0.1892
1970-1979	-0.0227	0.0042	-0.4690	0.3491	0.7331	-0.3488
1980-1989	-0.0543	0.0143	-1.0949	1.1555	1.7980	-1.3611
1990-1994	-0.0874	0.0289	-1.7906	2.3675	2.4069	-2.9955

Table 5A
Adjusted Elsticity of Cost with Respect to Output
ITSUR Estimates
Average By Decade

Decade	ε_{cy}	ε_{cg}	ε_{gy}	ε_{cr}	ε_{ry}	$\varepsilon^{A}{}_{cy}$
1949-1959	6.0191	0.0162	-11.6249	-0.3326	-12.7337	10.0654
1960-1969	3.3421	0.1955	-7.1423	-0.2877	-10.2991	4.9092
1970-1979	1.6886	0.9398	-3.0628	-0.5372	-5.3083	1.6617
1980-1989	2.9251	2.7906	-1.6174	-2.4054	-1.9774	3.1680
1990-1994	5.0124	4.2069	-1.3134	-5.3131	-1.2816	6.2965

Table 1B
MH Parameter Estimates

Parameter	Estimate	t Value	Parameter	Estimate	t Value
hk	-0.05068	-2.06	ag	0.000032	1.76
bpkk	0.252834	13.34	ar	-0.01231	-1.19
bplk	-0.05284	-16.27	ay	-259.881	-1.70
bpkL	0.415566	2.89	bkk	-0.8161	-0.37
bplL	0.178916	6.93	bkL	2.563959	1.84
bpkpk	-1.49199	-2.12	bkg	2.41E-06	0.92
bpkpl	4.119652	5.28	bkr	-0.00177	-0.80
bplpl	-38.4536	-5.33	bky	15.13473	1.68
nkg	-1.27E-09	-0.22	bLL	1.751193	0.90
nkr	2.63E-06	0.59	bLg	-0.00001	-1.83
nky	-0.0221	-2.75	bLr	0.009502	1.90
hl	0.612025	3.33	bLy	20.56857	1.31
nlg	3.80E-09	0.19	bgg	2.77E-11	1.79
nlr	-3.5E-05	-1.41	bgr	-8.65E-08	-9.44
nly	0.155178	2.19	bgy	-0.00002	-1.55
a0	134.2401	1.59	brr	8.61E-06	1.97
ak	-19.6675	-1.88	bry	-0.00926	-1.02
aL	-22.1636	-1.30	byy	1274.933	9.22

Table 2B
Shadow Prices of Public Inputs
MCMC Estimates
Average By Decade

Decade	P*g	P*r
1949-1959	0.0000008	0.0004382
	(0.0000002)	(0.0002293)
1960-1969	0.0000005	0.001426
	(0.0000003)	(0.0002053)
1970-1979	0.0000010	0.0043971
	(0.0000006)	(0.0004762)
1980-1989	0.0000048	0.0094812
	(0.0000011)	(0.0009774)
1990-1994	0.0000098	0.0131529
	(0.0000012)	(0.0013057)

Table 3B
Short-Run Elasticities of Demand for Private Inputs
w/ Respect to Public Goods - MCMC Estimation
Average By Decade

Decade	E_{KG_SR}	E_{KR_SR}	E_{LG_SR}	E_{LR_SR}	E_{MG_SR}	E_{MR_SR}
1949-1959	0.001024	-0.00314	-0.000064	0.000209	-0.055511	0.0165787
	(0.004361)	(0.004948)	(0.000521)	(0.000369)	(1.911529)	(1.6393535)
1960-1969	0.001571	-0.010085	-0.000274	0.000395	0.227909	-0.201544
	(0.008291)	(0.007133)	(0.001304)	(0.000703)	(0.665811)	(0.4454666)
1970-1979	0.004234	-0.022613	-0.000945	0.001136	0.351909	-0.398083
	(0.022389)	(0.016082)	(0.004381)	(0.001967)	(0.437095)	(0.2141283)
1980-1989	0.009688	-0.073346	-0.002552	0.004348	0.319634	-2.875702
	(0.051173)	(0.052028)	(0.011634)	(0.007419)	(1.412777)	(0.7075312)
1990-1994	0.01524	-0.144549	-0.004074	0.008698	1.854532	7.7162574
	(0.080443)	(0.102555)	(0.018737)	(0.014966)	(2.526301)	(2.3555209)

Table 4B
Long-Run Elasticities of Demand for Private Inputs
w/ Respect to Public Goods
Average By Decade

Decade	E_{KG_LR}	E_{KR_LR}	E_{LG_LR}	E_{LR_LR}	E_{MG_LR}	E_{MR_lR}
1949-1959	0.002857	-0.00876	0.000498	-0.00163	-0.01863	-0.01004
	(0.00027)	(0.000829)	(1.65E-05)	(5.38E-05)	(0.004214)	(0.004425)
1960-1969	0.002366	-0.01567	0.001421	-0.00252	-0.01692	-0.03368
	(0.000284)	(0.001882)	(3.88E-05)	(6.87E-05)	(0.009149)	(0.004084)
1970-1979	0.006648	-0.03662	0.00399	-0.00588	-0.06431	-0.17719
	(0.000803)	(0.004426)	(0.000134)	(0.000198)	(0.043828)	(0.016638)
1980-1989	0.016232	-0.12677	0.00912	-0.01906	-1.1995	-2.01438
	(0.002087)	(0.016297)	(0.000295)	(0.000617)	(0.34993)	(0.200925)
1990-1994	0.022502	-0.22018	0.015937	-0.04173	4.580564	6.403403
	(0.001775)	(0.017371)	(0.000547)	(0.001433)	(1.311022)	(1.783446)

Table 5B
Adjusted Elsticity of Cost with Respect to Output - MCMC Estimation
Average By Decade

Decade	ε_{cy}	ε_{cg}	ε_{gy}	ε_{cr}	ε_{ry}	ε^{A}_{cy}
1949-1959	2.08066	-0.01931	17.12097	-0.00999	0.74827	1.74257
	(0.088874)	(0.005226)	(11.76344)	(0.005263)	(39.3168)	(0.444036)
1960-1969	2.278679	-0.02102	1.315998	-0.03838	2.884776	2.140291
	(0.112834)	(0.009346)	(7.227383)	(0.005559)	(31.79968)	(1.553354)
1970-1979	2.581448	-0.07948	0.564339	-0.20971	1.486848	2.224795
	(0.155223)	(0.048143)	(3.099318)	(0.028902)	(16.38993)	(3.945232)
1980-1989	10.46272	-2.54554	0.298008	-4.15066	0.553867	7.405213
	(40.29378)	(13.01153)	(1.63664)	(15.19929)	(6.105425)	(93.66995)
1990-1994	-5.87366	3.100688	0.241997	4.283043	0.358983	-3.58576
	(0.954618)	(0.679007)	(1.329035)	(0.609656)	(3.95717)	(17.70768)