# Preliminary Version 

## Not for Citation

# The Effects of Public R\&D on U.S. Agriculture: A State-Level Analysis 

Alejandro S. Plastina and Lilyan E. Fulginiti

## The Effects of Public R\&D on

## Cost of U.S. Agriculture: A State-Level Analysis

## Introduction

Annual average growth estimates of U.S. agricultural total factor productivity (TFP) have been in the range of 1.34 to 2.31 percent for the second half of the last century ${ }^{1}$. Public investment in agricultural research and development (R\&D) has been singled out as a major contributor to agricultural productivity growth, in the order of 50 percent (Shane, Roe, and Gopinath 1998, p.9). The patterns of agricultural TFP growth at state-level, however, vary significantly (Ball et al 1999; Ball, Butault, and Nehring 2001; Acquaye, Alston and Pardey 2003), and so does the impact of public R\&D on them.

Several studies attempt to estimate the contribution of R\&D to agricultural growth at state-level by regressing state-level agricultural TFP (a "measure of our ignorance") against different proxies for the stock of own public R\&D and the stock of R\&D spill-ins from other states. The estimates are usually positive, even after controlling for the effects of private investment in R\&D and public investment in infrastructure (e.g. Huffman and Evenson; Alston et al; Yee et al). All those studies highlight the relevance of public investment in R\&D in the development of the agricultural sector. Huffman et al, using a cost function framework, estimate the (average) marginal internal rate of return (IRR) on own-state public investment in R\&D for five Midwestern states (Minnesota, Iowa, Illinois, Missouri and Indiana) to be $11 \%$ per annum, and the (average) marginal IRR due to public R\&D spillover effects to be $43 \%$ per annum.

The purpose of this study is to estimate the effects of public R\&D on cost of U.S. agriculture at state level. The theoretical approach consists of a dual dynamic model in a cost of adjustment framework (Onofri and Fulginiti 2005), where agricultural firms minimize costs of production intertemporally subject to the provision of local and regional public inputs. The resulting system of structural equations is estimated using production data from Acquaye, Alston and Pardey and a set of R\&D stock variables constructed by the authors for the present study.

[^0]
## The model

According to the literature, public goods might have positive effects on agricultural productivity, or, equivalently, they might reduce the cost of production of a given the level of agricultural output. The technology of the firm is represented by a cost function with only one perfectly variable input and a vector of quasi-fixed inputs. As in Onofri and Fulginiti (2005), the firm chooses gross investment so as to minimize intertemporal costs of production in a cost of adjustment framework. The basic assumption of the model is that at any point in time, the firm takes factor prices and the level of output to be exogenous and minimizes the discounted sum of future costs over an infinite horizon (Epstein and Denny 1983). More formally, the firm solves the following problem:
(1) $\operatorname{Min} \sum_{I(t)>0}^{\infty} \int_{o}^{-\rho t}\left[C(y ; Z, I ; G)+W^{\prime} Z\right] d t$ subject to: $\dot{Z}=I-\delta Z ; Z(0)=Z_{o} ; Z(t)>0 \quad \forall t$
where $C(y ; Z, I ; G)$ is the variable cost function, $y$ is output, $Z$ is the vector of quasi-fixed inputs, $W$ is the vector of relative rental prices corresponding to $Z, I$ is the vector of gross changes in quasi-fixed inputs, $\delta$ is a diagonal matrix of fixed rates of depreciation of the quasi-fixed factors, $G$ is the vector of public inputs, exogenous to the firm, and $\rho>0$ is the firm's real rate of discount. The price of the perfectly variable input is used as the numeraire. The apostrophe indicates the transposition operator.
Define now $J(y ; Z, W ; G)$ as the value function that solves problem (1). Assuming that the functions $C(y ; Z, I ; G)$ and $J(y ; Z, W ; G)$ satisfy the regularity conditions cited in Onofri and Fulginiti (2005), and making use of the Hamilton-Jacobi-Bellman equation, duality between $C(y ; Z, I, G)$ and $J(y ; Z, W ; G)$ can be established:
(2) $C(y ; Z, I ; G)=\underset{W}{\operatorname{Max}}\left[\rho J(y ; Z, W ; G)-W^{\prime} Z-J_{Z}^{\prime}(y ; Z, W ; G)(I-\delta Z)\right]$
or
(3) $\rho J(y ; Z, W ; G)=\operatorname{Min}_{I}\left[C(y ; Z, I ; G)+W^{\prime} Z+J_{z}{ }^{\prime}(y ; Z, W ; G)(I-\delta Z)\right]$

The optimal net investment demand functions are obtained by differentiating equation (3) with respect to the rental rates:
(4) $\dot{\mathrm{Z}}^{*}(y ; \mathrm{W}, \mathrm{Z} ; G)=J_{W Z}^{-1}(y ; \mathrm{Z}, W ; G)\left[\rho J_{W}(y ; \mathrm{Z}, W ; G)-\mathrm{Z}\right]$

The optimal variable factor demand, $X$, can be solved for by rearranging equation (4) in terms of $Z$ and introducing it into equation (3)
(5) $X^{*}(\mathrm{y}, \mathrm{Z}, W ; G)=\rho\left[J-\mathrm{W}^{\prime} J_{W}\right]-\left[J_{z}^{\prime}-\mathrm{W}^{\prime} J_{W Z}\right] \dot{\mathrm{Z}}{ }^{*}$

Equation (2) can also be used to evaluate the effects of $G$ on the firm's technology in steady state. Taking the total differential of $C$ with respect to $G$ and letting $\dot{Z}=0$ :
(6) $C_{G}(y, Z, I ; G)=\rho J_{G}(y, Z, W ; G)$

If public inputs effectively reduce costs of production, equation (6) should have a negative value.

## Econometric specification

The system of equations determined by (4) and (5) is estimated using state level data on labor $(L)$, purchased inputs ( $M$ ), land $(T)$, capital ( $K$ ), and agricultural output ( $y$ ) from Acquaye, Alston and Pardey (2003). Purchased inputs are assumed to be the only perfectly variable inputs, while capital, labor, and land are assumed to be quasi-fixed inputs. Input price and quantity indexes are Fisher indexes, while output quantity and price indexes are Tornquist indexes. The base year for all indexes is $1949=100$. Inputs and output quantities are obtained by multiplying the quantity indexes by the corresponding expenditure in 1949. Inputs and output prices are obtained by dividing each price index by the price index for purchased inputs.
The public input variables comprise the own-state's stock of public agricultural R\&D, $G$, and three spill-in variables from neighboring states: $S 1$ represents the stock of agricultural R\&D of the states that share common borders with the state under analysis (band 1); $S 2$ represents the stock of agricultural R\&D of the states that share common borders with the states in band 1 (band2), and $S 3$ represents the stock of agricultural R\&D of the states that share common borders with the states in band 2 (band 3). The logic for having separate public stock variables rests on the potentially different degrees of appropriability of $R \& D$ conducted in states with different degrees of similarity to the state under
analysis. The stocks of agricultural R\&D for each state are constructed using Chavas and Cox's method (1992). See Appendix 1 for a complete description of the variables. In order to econometrically estimate and test for the results of this model, a specific functional form must be adopted for the value function $J$. The following normalized quadratic value function is hypothesized:
(7) $J(\mathrm{~W}, \mathrm{Z}, \mathrm{Q})=a_{0}+\left[\begin{array}{lll}\mathrm{A}_{\mathrm{W}}^{\prime} & \mathrm{A}_{\mathrm{Z}}^{\prime} & \mathrm{A}_{\mathrm{Q}}^{\prime}\end{array}\right]\left[\begin{array}{l}\mathrm{W} \\ \mathrm{Z} \\ \mathrm{Q}\end{array}\right]+\frac{1}{2}\left[\begin{array}{ll}\mathrm{W}^{\prime} & \mathrm{Z}^{\prime} \mathrm{Q}^{\prime}\end{array}\right]\left[\begin{array}{lll}\mathrm{B}_{\mathrm{ww}} & \mathrm{B}_{\mathrm{WZ}}^{-1} & \mathrm{~B}_{\mathrm{wQ}} \\ \mathrm{B}_{\mathrm{WZ}}^{-1} & \mathrm{~B}_{\mathrm{ZZ}} & \mathrm{B}_{\mathrm{ZQ}} \\ \mathrm{B}_{\mathrm{WQ}}^{\prime} & \mathrm{B}_{\mathrm{ZQ}}^{\prime} & \mathrm{B}_{\mathrm{QQ}}\end{array}\right]\left[\begin{array}{l}\mathrm{W} \\ \mathrm{Z} \\ \mathrm{Q}\end{array}\right]$
where

$$
W=\left[\begin{array}{c}
P_{L} \\
P_{K} \\
P_{T}
\end{array}\right] \quad Z=\left[\begin{array}{c}
L \\
K \\
T
\end{array}\right] \quad Q=\left[\begin{array}{c}
y \\
G \\
S 1 \\
S 2 \\
S 3
\end{array}\right]
$$

This is a second order Taylor approximation of $J$ in (W, Z, Q) as defined in equation (3). $\mathrm{A}_{\mathrm{i}}$ and $\mathrm{B}_{\mathrm{ij}}$ are parameter matrices of appropriate order reflecting the first order effect of each $i$ the former, and the second order or interaction terms between $i$ and $j$ the former. In a similar fashion to Paul (2001) and Cohen and Paul (2004), $a_{0}=\sum_{i=1}^{48} \delta_{i} F_{i}$ accounts for state fixed effects ( $F_{i}$ is a dummy variable for state $i$ ).

Equations (5) and (6) can be expressed in terms of the parameters of this value function in the following way:
(8) $\dot{\mathrm{Z}}^{*}=\rho \mathrm{B}_{\mathrm{wz}} A_{W}+\left(\rho u-\mathrm{B}_{\mathrm{wz}}\right) \mathrm{Z}+\rho \mathrm{B}_{\mathrm{Wz}} \mathrm{B}_{\mathrm{ww}} \mathrm{W}+\rho \mathrm{B}_{\mathrm{wz}} \mathrm{B}_{\mathrm{wQ}} \mathrm{Q}$

$$
\begin{align*}
X^{*} & =\rho\left[a_{0}-\frac{1}{2} \mathrm{~W}^{\prime} \mathrm{B}_{\mathrm{WW}} \mathrm{~W}+\mathrm{A}_{\mathrm{z}}^{\prime} \mathrm{Z}+\mathrm{A}_{\mathrm{Q}}^{\prime} \mathrm{Q}+\frac{1}{2} \mathrm{Z}^{\prime} \mathrm{B}_{\mathrm{ZZ}} \mathrm{Z}+\mathrm{Z}^{\prime} \mathrm{B}_{\mathrm{ZQ}} \mathrm{Q}+\right.  \tag{9}\\
& \left.+\frac{1}{2} \mathrm{Q}^{\prime} \mathrm{B}_{\mathrm{QQ}} \mathrm{Q}\right]-\left[\mathrm{A}_{\mathrm{Z}}^{\prime}+\mathrm{Z}^{\prime} \mathrm{B}_{\mathrm{ZZ}}+\mathrm{Q}^{\prime} \mathrm{B}_{\mathrm{ZQ}}^{\prime}\right] \dot{\mathrm{Z}}^{*}
\end{align*}
$$

where $u$ is the identity matrix. Note that equation (8) can be expressed as a multivariate flexible accelerator corrected for the existence of externalities:

$$
\begin{equation*}
\dot{\mathrm{Z}}^{*}(\mathrm{Z}, \mathrm{~W}, \mathrm{Q})=\mathrm{M}[\mathrm{Z}-\overline{\mathrm{Z}}(\mathrm{~W}, \mathrm{Q})] \tag{10}
\end{equation*}
$$

where $M=\left(\rho u-\mathrm{B}_{\mathrm{wz}}\right)$ and $\overline{\mathrm{Z}}=-\left(\rho u-\mathrm{B}_{\mathrm{wz}}\right)^{-1} \rho\left[\mathrm{~B}_{\mathrm{wz}} A_{W}+\mathrm{B}_{\mathrm{wz}} \mathrm{B}_{\mathrm{ww}} \mathrm{W}+\mathrm{B}_{\mathrm{WZ}} \mathrm{B}_{\mathrm{wQ}} \mathrm{Q}\right]$ which is the steady state value of the quasi-fixed inputs, and it depends on the own-state's stock of agricultural R\&D and the spill-ins from neighboring states.
So far, the model is based on production theory, while the data is aggregated at state level. Consistent linear aggregation requires the aggregate value function at state level to equal the unweighted sum of the value functions of all firms within a state, and the stock of private quasi-fixed inputs at state level to equal the unweighted sum of the stocks of private quasi-fixed inputs for all firms within a state, as well as the aggregate output to equal the unweighted sum of the output of all firms in the state (Onofri and Fulginiti 2005). More formally, consistent linear aggregation requires:

$$
\begin{align*}
& J(y, Z, W, G)=\sum_{i} J\left(y_{i}, Z_{i}, W, G\right) \\
& Z=\sum_{i} Z_{i}  \tag{11}\\
& y=\sum_{i} y_{i}
\end{align*}
$$

Note that the aggregation conditions apply only to private quasi-fixed inputs and output, but not to the stocks of public R\&D, since they are non-rival by definition. These conditions require the value function J to be linear in Z and y , implying that $J_{Z Z}=B_{Z Z}=0, J_{y y}=0$, and $J_{Z y}=0$.

In order to be able to estimate the model with annual data, a discrete time approximation must be undertaken for $\dot{Z} \cong Z_{t}-Z_{t-1}$, and the stocks of the private quasi-fixed inputs for each period are considered to be the stocks at the beginning of that period, or what it is the same, the stocks at the end of last period. Then, equations (8) and (9) can be reexpressed as:

$$
\begin{align*}
Z_{t}= & \rho B_{W Z} A_{W}+\left(u+\rho u-B_{W Z}\right) Z_{t-1}+\rho B_{W Z} B_{W W} W_{t}+\rho B_{W Z} B_{W Q} Q_{t}  \tag{12}\\
X_{t}^{*}= & \rho\left[a_{0}-\frac{1}{2} W_{t}{ }^{\prime} B_{W W} W_{t}+\left(A_{Q}^{\prime}+\frac{1}{2} Q_{t}{ }^{\prime} B_{Q Q}\right) Q_{t}\right]+  \tag{13}\\
& \left(A_{Z}^{\prime}+Q_{t}^{\prime} B_{Z Q}^{\prime}\right) B_{W Z}\left[Z_{t-1}-\rho\left(A_{W}+B_{W W} W_{t}+B_{W Q} Q_{t}\right)\right]
\end{align*}
$$

Equations (12) and (13) constitute the system of equations to be jointly estimated by SUR. The error structure is specified so as to allow for first order autocorrelation and a
spatial autoregressive (SAR) structure, to accommodate temporal and spatial lags in the system of equations (Cohen and Paul 2004):

$$
\begin{array}{ll}
u_{Z, i, t}=\rho_{s, Z} \sum_{j=1}^{48} w_{i, j} u_{Z, j, t}+\psi_{Z, i, t} \quad ; \quad \psi_{Z, i, t}=\rho_{\theta, Z} \psi_{Z, i, t-1}+\phi_{Z, i, t} \\
u_{X, i, t}=\rho_{s, X} \sum_{j=1}^{48} w_{i, j} u_{X, j, t}+\psi_{X, i, t} \quad \psi_{X, i, t}=\rho_{\theta, X} \psi_{X, i, t-1}+\phi_{X, i, t} \tag{15}
\end{array}
$$

$\phi_{Z, i, t} \sim$ i.i.d. $N\left(0, \sigma_{Z}^{2}\right) ; \phi_{X, i, t} \sim$ i.i.d. $N\left(0, \sigma_{X}^{2}\right) ; w_{i, j}$ is the weight that state $j$ has on state $i$; $w_{i, i}=0 ;-1 \leq \rho_{s, Z} \leq 1 ;-1 \leq \rho_{\theta, Z} \leq 1 ;-1 \leq \rho_{s, X} \leq 1 ;-1 \leq \rho_{\theta, X} \leq 1$. The elements of $\psi_{Z, i, t}$ are assumed to be i.i.d., and $\phi_{Z, i, t}$ and $\psi_{Z, i, t}$ are assumed to be independent. Similarly, the elements of $\psi_{X, i, t}$ are assumed to be i.i.d., and $\phi_{X, i, t}$ and $\psi_{X, i, t}$ are assumed to be independent. Furthermore, we assume no cross-equation correlation among the "well behaved" errors $\phi_{X, i, t}$ and $\phi_{Z, i, t}$, i.e. we assume $\phi_{X, i, t}$ and $\phi_{L, i, t}, \phi_{K, i, t} \phi_{T, i, t}$, and $\phi_{L, i, t}$ and $\phi_{K, i, t}, \phi_{T, i, t}$, and $\phi_{K, i, t}$ and $\phi_{T, i, t}$ are independent.

The temporal dimension of the panel data is recognized by incorporating an autoregressive structure as in Holtz-Eakin and Schwartz (1995), Keleijian and Robinson (1997) and Cohen and Paul (2004). The spatial autoregressive error structure accounts for shocks in the derived demands that disseminate through neighboring states. Although for simplicity of exposition the SAR lag structure in equations (14) and (15) is of order 1 , we have no a priori information on the proper SAR lag structure, i.e. we do not know the extent to which shocks in the derived demands on one state disseminate to other neighboring states. Applying an extension of the Keleijian and Robinson (1992) spatial autocorrelation test for a system of equations (Cohen and Paul 2005), we are able to test econometrically the extent of the spatial dissemination of the shocks. The test requires an a priori specification of the bands of states that might be spatially correlated, but it does not require knowledge of the weights $w_{i, j}$. In particular, we estimate the system of equations (14) and (15) with the autoregressive part of the error structure by iterative seemingly unrelated regression (ITSUR), with no spatial autocorrelation, and save the estimates of the errors $\varepsilon_{l, i, t}=\rho_{s, l} \sum_{j=1}^{48} w_{i, j}\left(u_{l, j, t}-\rho_{\theta, l} u_{l, j, t-1}\right)+\ldots+\phi_{l, i, t}$, where $l=K, M, L, T$; $i=1, \ldots, 48$; and $j=1950, \ldots, 1991$. Note that the three dots indicate potentially significant
higher order SAR lags, for which we are testing. Then, for each equation, the hypothesis of first order SAR lags among state $i$ and its neighboring states $j$ from band 1 is tested by running the following ordinary least squares (OLS) regression:

$$
\begin{equation*}
\varepsilon_{l, i, t} \varepsilon_{l, j, t}=c_{1}+\vartheta_{1, l, i, j, t} \tag{16}
\end{equation*}
$$

where $\vartheta_{1, l, i, j, t}$ is assumed to satisfy the classical assumptions in the OLS model. If the parameter estimate for the constant term $c_{1}$ is significantly different from zero, the null hypothesis "Ho: $\varepsilon_{l, i, t}$ and $\varepsilon_{l, j, t}$ are not spatially correlated" is rejected in favor of the alternative hypothesis "Ha: $\varepsilon_{l, i, t}$ and $\varepsilon_{l, j, t}$ are spatially correlated".

If Ho is rejected, then the hypothesis of second order SAR lags among state $i$ and its neighboring states $j$ ' from band 2 is tested by running the following OLS regression:

$$
\begin{equation*}
\varepsilon_{l, i, t} \varepsilon_{l, j^{\prime}, t}=c_{2}+\vartheta_{2, l, l, j^{\prime}, t} \tag{17}
\end{equation*}
$$

where $\vartheta_{2, l, i, j, t}$ is assumed to satisfy the classical assumptions in the OLS model. If the parameter estimate for the constant term $c_{2}$ is significantly different from zero, the null hypothesis "Ho: $\varepsilon_{l, i, t}$ and $\varepsilon_{l, j, t}$ are not spatially correlated" is rejected in favor of the alternative hypothesis "Ha: $\varepsilon_{l, i, t}$ and $\varepsilon_{l, j^{\prime}, t}$ are spatially correlated". If Ho is rejected, then the hypothesis of third order SAR lags among state $i$ and its neighboring states $j$ ' ' from band 3 is applied. And the test is applied successively in a similar manner for higher order spatial lags.

For our model, the Keleijian and Robinson (1992) tests suggest that the spatial lag structure in the demand for land $(\mathrm{T})$ is of order three, in the demand for purchased inputs $(\mathrm{M})$ is of order one, and in the demands for labor $(\mathrm{L})$ and capital $(\mathrm{K})$ is of order five ${ }^{2}$. More formally, the error structure is:

$$
\begin{align*}
& u_{T, i, t}=\rho_{s 1, T} \sum_{j=1}^{48} w_{1, i, j} u_{T, j, t}+\rho_{s 2, T} \sum_{j=1}^{48} w_{2, i, j} u_{T, j, t}+\rho_{s 3, T} \sum_{j=1}^{48} w_{3, i, j} u_{T, j, t}+\psi_{T, i, t}  \tag{18}\\
& u_{M, i, t}=\rho_{s, M} \sum_{j=1}^{48} w_{i, j} u_{M, j, t}+\psi_{M, i, t} \tag{19}
\end{align*}
$$

[^1]\[

$$
\begin{align*}
u_{h, i, t} & =\rho_{s 1, h} \sum_{j=1}^{48} w_{1, i, j} u_{h, j, t}+\rho_{s 2, h} \sum_{j=1}^{48} w_{2, i, j} u_{h, j, t}+\rho_{s 3, h} \sum_{j=1}^{48} w_{3, i, j} u_{h, j, t} \\
& +\rho_{s 4, h} \sum_{j=1}^{48} w_{4, i, j} u_{h, j, t}+\rho_{s 5, h} \sum_{j=1}^{48} w_{5, i, j} u_{h, j, t}+\psi_{h, i, t} ;  \tag{20}\\
h & =L, K \\
\psi_{f, i, t} & =\rho_{\theta, f} \psi_{f, i, t-1}+\phi_{f, i, t} ; f=L, K, T, M \tag{21}
\end{align*}
$$
\]

For each equation, after stacking up all observations, the error structure in matrix form is represented as:

$$
\begin{align*}
& u_{T}=\rho_{s 1, T} W_{1} u_{T}+\rho_{s 2, T} W_{2} u_{T}+\rho_{s 3, T} W_{3} u_{T}+\psi_{T}  \tag{22}\\
& u_{M}=\rho_{s, M} W_{1} u_{M}+\psi_{M}  \tag{23}\\
& u_{h}=\rho_{s 1, h} W_{1} u_{h}+\rho_{s 2, h} W_{2} u_{h}+\rho_{s 3, h} W_{3} u_{h}+\rho_{s 4, h} W u_{h}+\rho_{s 5, h} W_{5} u_{h}+\psi_{h} ;  \tag{24}\\
& h=L, K \\
& \psi_{f, t}=\rho_{\theta, f} \psi_{f, t-1}+\phi_{f, t} ; f=L, K, T, M \tag{25}
\end{align*}
$$

where $W_{i}=I \otimes \omega_{i}, i=1,2,3,4,5$; and $\omega_{i}$ is a 48x48 matrix of neighbors with each element taking the value of $1 / \mathrm{N}_{\mathrm{i}}$ if the corresponding state belongs to band $i$ and 0 otherwise. $\mathrm{N}_{\mathrm{i}}$ is the number of neighboring states in band $i$, and $I$ is a $42 \times 42$ identity matrix. The estimation of the SAR lags is performed through the Generalized Method of Moments (GMM) approach derived by Keleijian and Prucha (1999), and applied in a system of equation framework by Cohen and Paul (2005). In terms of the stacked errors $\varepsilon_{l, i, t}$, the relevant moments for the estimation of the first order SAR lag for equation M are:

$$
\begin{align*}
& E\left[(1 / N T) \varepsilon_{M}^{\prime} \varepsilon_{M}\right]=\sigma_{M}^{2}  \tag{26}\\
& E\left[(1 / N T) \varepsilon_{M}^{\prime} W_{1}^{\prime} W_{1} \varepsilon_{M}\right]=\sigma_{M}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{1}^{\prime} W_{1}\right)  \tag{27}\\
& E\left[(1 / N T) \varepsilon_{M}^{\prime} W_{1}^{\prime} \varepsilon_{M}\right]=0 \tag{28}
\end{align*}
$$

For equation T, the three SAR lags are determined by the system of equations consistent of the equivalent of equations (26)-(28) and the following equations:

$$
\begin{align*}
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{2}^{\prime} W_{2} \varepsilon_{T}\right]=\sigma_{T}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{2}^{\prime} W_{2}\right)  \tag{29}\\
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{1}^{\prime} W_{2} \varepsilon_{T}\right]=\sigma_{T}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{1}^{\prime} W_{2}\right)  \tag{30}\\
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{2}^{\prime} \varepsilon_{T}\right]=0  \tag{31}\\
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{3}^{\prime} W_{3} \varepsilon_{T}\right]=\sigma_{T}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{3}^{\prime} W_{3}\right) \tag{32}
\end{align*}
$$

$$
\begin{align*}
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{1}^{\prime} W_{3} \varepsilon_{T}\right]=\sigma_{T}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{1}^{\prime} W_{3}\right)  \tag{33}\\
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{2}^{\prime} W_{3} \varepsilon_{T}\right]=\sigma_{T}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{2}^{\prime} W_{3}\right)  \tag{34}\\
& E\left[(1 / N T) \varepsilon_{T}^{\prime} W_{3}^{\prime} \varepsilon_{T}\right]=0 \tag{35}
\end{align*}
$$

To obtain the five SAR lags for equations L and K , the GMM approach is applied to the equivalent of equations (26)-(35) and the following equations:

$$
\begin{align*}
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{4}^{\prime} W_{4} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{4}^{\prime} W_{4}\right)  \tag{36}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{1}^{\prime} W_{4} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{1}^{\prime} W_{4}\right)  \tag{37}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{2}^{\prime} W_{4} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{2}^{\prime} W_{4}\right)  \tag{38}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{3}^{\prime} W_{4} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{3}^{\prime} W_{4}\right)  \tag{39}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{4}^{\prime} \varepsilon_{h}\right]=0  \tag{40}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{5}^{\prime} W_{5} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{5}^{\prime} W_{5}\right)  \tag{41}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{1}^{\prime} W_{5} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{1}^{\prime} W_{5}\right)  \tag{42}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{2}^{\prime} W_{5} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{2}^{\prime} W_{5}\right)  \tag{43}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{3}^{\prime} W_{5} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{3}^{\prime} W_{5}\right)  \tag{44}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{4}^{\prime} W_{5} \varepsilon_{h}\right]=\sigma_{h}^{2}(N T)^{-1} \operatorname{Tr}\left(W_{4}^{\prime} W_{5}\right)  \tag{45}\\
& E\left[(1 / N T) \varepsilon_{h}^{\prime} W_{5}^{\prime} \varepsilon_{h}\right]=0 \tag{46}
\end{align*}
$$

where $h=L, K$.
The estimates of the spatial lags are: $\rho_{s 1, M}=0.43358, \rho_{s 1, T}=0.29855$,

$$
\begin{aligned}
& \rho_{s 2, T}=0.29073, \rho_{s 3, T}=-0.07143, \rho_{s 1, L}=0.21995, \rho_{s 2, L}=0.10796, \\
& \rho_{s 3, L}=0.14481, \rho_{s 4, L}=-0.01758, \rho_{s 5, L}=0.10844, \rho_{s 1, K}=0.45758, \rho_{s 2, K}=0.03746,
\end{aligned}
$$

$$
\rho_{s 3, K}=0.07029, \rho_{s 4, K}=0.08152, \rho_{s 5, K}=-0.03164 \text {. The coefficients of the first order }
$$

SAR lags are all positive, and tend to decrease in value for higher order SAR lags. Using the estimated spatial lags, we perform a Cochrane-Orcutt transformation of the variables in each of the four demand equations, and estimate the transformed system using ITSUR.

## Results

The estimation was conducted using PROC MODEL in SAS 9.1. The Adjusted R-square for each equation exceeds 0.98 (Appendix 2) and most of the parameter estimates are statistically significant (Appendix 3). But in order to assess the effects of public R\&D in agricultural productivity, equation (6) is estimated at the mean of the data for each state (Table 1). The variances are estimated applying the delta method, and the p-value corresponds to a 2-sided test. Surprisingly, the effect of the own-state R\&D stock is positive for most of the states, indicating that an increase in the stock of R\&D increases the cost of agricultural production, instead of decreasing it. The only states for which the estimated effect is negative and significant are Delaware, Rhode Island and Nevada.

Table 1. Effects of the Own-state Public R\&D Stock on the Cost of Agricultural

## Production

| State | Estimated <br> Effect | T-test | p-value |
| :---: | :---: | :---: | :---: |
| AL | 776.167 | 1.478 | 0.139 |
| AR | $1,002.980$ | 2.080 | 0.038 |
| AZ | -213.411 | -0.453 | 0.651 |
| CA | $4,641.320$ | 5.273 | 0.000 |
| CO | 846.311 | 1.877 | 0.061 |
| CT | -352.418 | -0.873 | 0.382 |
| DE | -860.067 | -1.833 | 0.067 |
| FL | -157.509 | -0.255 | 0.799 |
| GA | 261.138 | 0.468 | 0.640 |
| IA | $2,854.170$ | 4.626 | 0.000 |
| ID | 610.233 | 1.385 | 0.166 |
| IL | $2,649.040$ | 4.152 | 0.000 |
| IN | $1,748.630$ | 3.306 | 0.001 |
| KS | $1,946.560$ | 3.904 | 0.000 |
| KY | $2,483.800$ | 4.629 | 0.000 |
| LA | $1,028.970$ | 1.835 | 0.066 |
| MA | -373.898 | -0.943 | 0.346 |
| MD | 580.727 | 0.849 | 0.396 |
| ME | -321.665 | -0.759 | 0.448 |
| MI | $1,348.620$ | 2.768 | 0.006 |
| MN | $2,150.450$ | 4.103 | 0.000 |
| MO | $2,500.050$ | 4.650 | 0.000 |


| MS | $1,607.130$ | 2.780 | 0.005 |
| :---: | :---: | :---: | :---: |
| MT | 605.611 | 1.533 | 0.125 |
| NC | $1,895.070$ | 3.559 | 0.000 |
| ND | 831.959 | 2.028 | 0.043 |
| NE | $1,429.040$ | 2.953 | 0.003 |
| NH | -740.928 | -1.577 | 0.115 |
| NJ | -214.502 | -0.518 | 0.604 |
| NM | 16.366 | 0.039 | 0.969 |
| NV | -939.720 | -1.810 | 0.070 |
| NY | $1,886.740$ | 3.169 | 0.002 |
| OH | $2,030.310$ | 3.851 | 0.000 |
| OK | $1,387.390$ | 2.876 | 0.004 |
| OR | $1,818.680$ | 3.361 | 0.001 |
| PA | $1,400.150$ | 2.719 | 0.007 |
| RI | -698.774 | -1.747 | 0.081 |
| SC | 357.610 | 0.697 | 0.486 |
| SD | $1,018.710$ | 2.566 | 0.010 |
| TN | $2,099.880$ | 4.062 | 0.000 |
| TX | $5,389.540$ | 6.657 | 0.000 |
| UT | 283.983 | 0.619 | 0.536 |
| VA | 856.327 | 1.802 | 0.071 |
| VT | -412.604 | -1.045 | 0.296 |
| WA | $1,623.600$ | 3.124 | 0.002 |
| WI | $2,880.400$ | 5.349 | 0.000 |
| WV | 78.944 | 0.179 | 0.858 |
| WY | 31.487 | 0.082 | 0.934 |

Another measure of interest is the dual measure of the economies of size, i.e. the change in total cost when output changes. At mean values, the estimated values suggest diseconomies of size in all states, i.e. increases in output generated ceteris paribus increases in total cost.

Table 2. Estimated Economies of Size

| State | Estimated <br> Effect | T-test | p-value |
| :---: | :---: | :---: | :---: |
| AL | 0.169 | 3.561 | 0.000 |
| AR | 0.184 | 4.269 | 0.000 |
| AZ | 0.160 | 3.710 | 0.000 |
| CA | 0.041 | 0.772 | 0.440 |


| CO | 0.175 | 4.495 | 0.000 |
| :---: | :---: | :---: | :---: |
| CT | 0.188 | 5.048 | 0.000 |
| DE | 0.181 | 4.103 | 0.000 |
| FL | 0.117 | 2.172 | 0.030 |
| GA | 0.144 | 2.917 | 0.004 |
| IA | 0.138 | 3.140 | 0.002 |
| ID | 0.175 | 4.502 | 0.000 |
| IL | 0.110 | 2.365 | 0.018 |
| IN | 0.154 | 3.508 | 0.000 |
| KS | 0.172 | 4.400 | 0.000 |
| KY | 0.176 | 4.216 | 0.000 |
| LA | 0.142 | 2.990 | 0.003 |
| MA | 0.192 | 5.230 | 0.000 |
| MD | 0.049 | 0.911 | 0.362 |
| ME | 0.199 | 5.045 | 0.000 |
| MI | 0.157 | 3.724 | 0.000 |
| MN | 0.145 | 3.506 | 0.000 |
| MO | 0.167 | 3.881 | 0.000 |
| MS | 0.152 | 3.109 | 0.002 |
| MT | 0.177 | 5.170 | 0.000 |
| NC | 0.166 | 3.759 | 0.000 |
| ND | 0.162 | 4.579 | 0.000 |
| NE | 0.160 | 3.982 | 0.000 |
| NH | 0.170 | 3.855 | 0.000 |
| NJ | 0.172 | 4.556 | 0.000 |
| NM | 0.184 | 4.757 | 0.000 |
| NV | 0.129 | 2.716 | 0.007 |
| NY | 0.092 | 1.999 | 0.046 |
| OH | 0.150 | 3.457 | 0.001 |
| OK | 0.163 | 3.929 | 0.000 |
| OR | 0.159 | 3.948 | 0.000 |
| PA | 0.140 | 3.246 | 0.001 |
| RI | 0.198 | 5.337 | 0.000 |
| SC | 0.172 | 3.656 | 0.000 |
| SD | 0.185 | 5.407 | 0.000 |
| TN | 0.186 | 4.379 | 0.000 |
| TX | 0.113 | 2.482 | 0.013 |
| UT | 0.175 | 4.351 | 0.000 |
| VA | 0.180 | 4.185 | 0.000 |
| VT | 0.197 | 5.331 | 0.000 |
| WA | 0.157 | 3.816 | 0.000 |


| WI | 0.150 | 3.717 | 0.000 |
| :---: | :---: | :---: | :---: |
| WV | 0.192 | 4.687 | 0.000 |
| WY | 0.190 | 5.549 | 0.000 |

These results are unsatisfactory, and further research should be undertaken to find the causes of these unexpected results, that may be related to the time series dimension of the panel data.

## References:

Acquaye, A., Alston, J., and P. Pardey. "Post-war Productivity Patterns in U.S. Agriculture: Influences of Aggregation Procedures in a State-Level Analysis." American Journal of Agricultural Economics 85 (February 2003): pp. 59-80. http://www.apec.umn.edu/faculty/ppardey/data.html

Ahearn, M., Yee, J., Ball, E., and R. Nehring. "Agricultural Productivity in the United States." Economic Research Service, U.S. Department of Agriculture. Agriculture Information Bulletin No. 740, Washington, DC, 1998.

Alston, J., and P. Pardey, Making Science Pay, AEI Press, 1996.
Alston, J., Craig, B., and P. Pardey "Dynamics in the Creation and Depreciation of Knowledge and the Returns to Research" EPDT Discussion paper no. 35, IFPRI, Washington, DC, 1998.

Ball, E., Butault, J-P., and R. Nehring. "U.S. Agriculture, 1960-96: A Multilateral Comparison of Total Factor Productivity." Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No.1895, Washington, DC, 2001.

Ball, E., Gollop, F., Kelly-Hawke A., and G. Swinand. "Patterns of State Productivity Growth in the U.S. Farm Sector: Linking State and Aggregate Models." American Journal of Agricultural Economics 81 (February 1999): 164-179.

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

Huffman and Evenson. 1993. Science for Agriculture: A long term perspective. IOWA State University Press, Ames, Iowa.

Huffman, W., Ball, E., Gopinath, M., and A. Somwaru. "Public R\&D and Infrastructure Policies: Effects on Cost of Midwestern Agriculture." In Agricultural Productivity: Measurement and Sources of Growth. Ed. Ball E. and G. Norton, Kluwer Academic Publishers, 2002.

O’Donnel, C., Shumway, C., and E. Ball. "Input Demands and Inefficiencies in U.S. Agriculture." American Journal of Agricultural Economics 81 (Nov 1999): 865880.

Onofri, A., and L. Fulginiti. "Public Inputs and Dynamic Producer Behavior: Endogenous Growth in U.S. Agriculture." Paper presented at the IX EWEPA, Brussels, June 2005.

Shane, M., Roe, T., and M. Gopinath. "U.S. Agricultural Growth and Productivity: An Economywide Perspective." Economic Research Service, U.S. Department of Agriculture. Agricultural Economic Report No. 758, Washington, DC, 1998.

Yee, J., Huffman, W., Aheran, M., and D. Newton. "Sources of Agricultural Productivity Growth at the State Level, 1960-1993." In Agricultural Productivity:

Measurement and Sources of Growth. Ed. Ball E. and G. Norton, Kluwer Academic Publishers, 2002.

## Appendix 1. Description of the Variables

## I. Production Variables: Acquaye, Alston, and Pardey, 1949-1991, Base 1949 Variables:

- Output price and quantity: field crops, fruit and nuts, vegetables, livestock, Greenhouse \& Nursery ("horticulture"), National Conservation Reserve Program Benefits. Most output quantity and price data were obtained from various publications of Agricultural Statistics, and USDA. Most of the quantity data are the reported quantities produced per state, and the price data are the state-specific prices received on farms. Some quantity data were derived implicitly from value and price data.
- Purchased Inputs price and quantity: Fertilizers (nitrogen, phosphoric oxide, and potash), Pesticides, Seeds, Purchased Feed, Water Usage, Other Operating Expenditures (Electricity, Fuels and oils, Repair and Maintenance, Machine hire, Miscellaneous).
- Capital price and quantity: Biological and Physical capital stocks. Livestock with service lives in excess of one year are treated as biological capital (beef and milking cows and heifers that have calved, ewes over one year, sows that have farrowed, and chickens not counting broilers). Physical capital:

Automobiles, Trucks, Tractors, Combines, Forage equipment, and Buildings and Structures.

- Land price and quantity: It includes three types of land: 1.Grassland, Pasture, Range, and Grazed Forest: forested pasture and range consisting of forest, brush grown pasture, arid woodlands and other areas within forested areas that have grass or other forage growth. It includes woodland pasture in farms and rough estimates of forested grazing land not in farms. It also includes all open land used primarily for pasture and grazing, including shrub and brush land types of pasture, grazing land with sagebrush and scattered mesquite, and all tame and native grasses, legumes, and other forage used for pasture or grazing. This data should not include cropland pastured but is not always clearly distinguished. It counts acres both in and out of farms but does not count acres in the Federal Conservation Research Program; 2. Cropland: Non-
irrigated cropland harvested, crop failure, cultivated summer fallow, cropland used only for pasture, and idle cropland; 3. Irrigated land: total irrigated acres. It includes both irrigated cropland and irrigated pastureland.
- Labor price and quantity: The quantity data for labor are the number of hours worked on-farm by the respective class of labor, and the price data are the wage received per hour of employment on the farm. Data for labor are made up of 30 farm operator classes (five age and six education characteristics), family labor, and hired labor.
- TFP: Indexes of total factor productivity for each State and year are formed as the ratio of the output index to the input index.


## Indexes:

The indexes of both quantities and prices of single inputs (land, labor, capital, purchased inputs) and single outputs (field crops, fruit and nuts, vegetables, livestock, Greenhouse \& Nursery, National Conservation Reserve Program Benefits) are Fisher indexes. Aggregate output and aggregate input are calculated using Tornqvist indexes. Total Factor Productivity is estimated as the ratio of aggregate output to aggregate input. The base year is 1949=100 for both price and quantity indexes.

## II. R\&D Stocks

(1) Procedure for Constructing the Agricultural R\&D Stocks at State Level In order to construct the series of Public Agricultural Research and Development (R\&D) stocks for each state, the following procedure has been followed:
a) Agricultural R\&D expenditures at state level conducted at the State Agricultural Experiment Stations (SAES) were calculated in current US\$ for the period 19191970 from different issues of the Report on the Agricultural Experiment Stations.
b) Total Public Agricultural R\&D expenditures at state level were calculated in current US\$ for the period 1970-1999 from the Current Research Information System Database.
c) The Agricultural R\&D Price Index was constructed for the period 1919-1999 from Huffman and Evenson (1993) and USDA data, and it was used to express the expenditure series in constant 1949 US\$.
d) Total Agricultural R\&D expenditures at state level for years 1919-1969 in constant 1949 US\$ were estimated as an expansion of the Ag. R\&D expenditures conducted at SAES in constant US\$ by the average ratio over 1970-1980 of Total Ag. R\&D expenditures to Ag. R\&D expenditures conducted at the SAES.
e) The stocks of Ag. R\&D at state level were constructed using Chavas and Cox's (1992) method.

## 1) Agricultural R\&D expenditures at state level conducted at the State Agricultural <br> Experiment Stations, 1919-1970

The calculations build upon the income of the SAES reported in the Report on the Agricultural Experiment Stations (USDA) for the years 1919-1970. Every time two or more SAESs reported R\&D expenditures for a state, the summation of the R\&D expenditures at each SAES is reported as the Agricultural R\&D expenditures at state level conducted at the SAES. Furthermore, since the Report on the Agricultural Experiment Stations reports data on income for each SAES, the expenditures of the SAES were calculated according to the general formula:

SAES Expenditures(t) =
Total Federal Funds(t) - Total Unobliged Balances from Federal Funds(t*)

- [Total Cooperative Forestry Research Act Funds (Mc Intire-Stennis ${ }^{3}$ )(t) Unobliged Mc Intire-Stennis Funds(t+1)]
- Carryovers from the Marketing Act ( $\mathrm{t}+1$ )
+ Total Non-Federal Funds(t) - Total Non-Federal Funds Balance from previous year(t+1)
where $(t)$ indicates the year when the data was reported, and ( $\mathrm{t}^{*}$ ) indicates that the source varied for different years, based on the availability of information: for 1919-1948, the

[^2]Unobliged Balances from Federal Funds were extracted from the Report on the Agricultural Experiment Stations of the year under analysis, (t); for 1949-1955, Unexpended Balances from the Hatch, Adams, Purnell, and Banhead-Jones Acts were obtained from the report of the year under analysis ( $t$ ), while Unexpended Balances for the Research and Marketing Federal Funds were obtained from the report of the following year, ( $\mathrm{t}+1$ ); for all years after 1955, Unexpended Balances from the Hatch Act as Amended and Regional Funds are obtained from the Report of the year under analysis, (t), while Unexpended Balances for other Federal Funds are obtained from the Report of the following year, $(t+1)$.

Note that the Marketing Act was first implemented in 1948, so there are no carryovers from that concept for years previous to 1949.

The 1942 Report on the Agricultural Experiment Stations does not include information on Total Non-Federal Funds Balance for year 1941, so the concept was calculated as the difference between the Total Income and the Total Expenditure reported in the 1941 Report. The same amount was added to the reported Non-Federal Funds for 1942.
2) Total Public Agricultural R\&D expenditures at state level, 1970-1999 Total Public Agricultural R\&D expenditures at state level were calculated using gross actual expenditures data from the Current Research Information System (USDA). The concept includes:

1. USDA Appropriations

### 1.1. Agricultural Research Service (ARS) Funds

### 1.2. Economic Research Service (ERS) Funds

### 1.3. Other USDA

2. Cooperative State Research, Education, and Extension Service (CSREES)

Administered Funds (for SAES and 1890 Institutions)
2.1. Hatch Act
2.2. Evans Allen Act
2.3. Animal Health
2.4. Grants and Agricultural Markets
2.5. National Research Initiative (NRI) Grants
2.6. Small Business Innovation Research (SBIR) Grants

### 2.7. Other CSREES Grants

3. Other USDA Funds (for SAES and 1890 Institutions)
4. Other Federal Funds (for SAES and 1890 Institutions)
5. State Appropriations (for SAES and 1890 Institutions)
6. Other Non-Federal Funds (for SAES and 1890 Institutions)

The series of Total Public Agricultural R\&D expenditures at state level ("Total RD" hereon) excludes the USDA appropriations for the Forest Service (FS), the Mc IntireStennis Act from the CSREES Administered Funds, and all funds for Forestry Schools.

## 3) Agricultural $R \& D$ Price Index

The Price Index was constructed using Price Index for Agricultural Research (1984=1) published by Huffman and Evenson (HE) and the Agricultural R\&D Deflator (2001=1) published by the ERS, USDA. The HF Price Index spans over 1888-1990, while the ERS R\&D Deflator spans over 1970-2001. The base of the Ag. R\&D Deflator was changed to 1984=1, and the correlation among the two series was measured to be almost perfect (0.9974) over the period 1970-1990 (Figure 1). Therefore, the Agricultural R\&D Price Index for 1919-1999 (1984=1) consists of the ERS R\&D Deflator for the period 19701999, and the HE Price Index for 1919-1969.
Finally, the base of the Agricultural R\&D Price Index was changed to 1949=1 to match the base year of the agricultural productivity variables in Acquaye et al (2003). The Ag. R\&D Price Index was used to express the SAES Expenditures and the Total RD series in constant 1949 US\$.


Figure 1
4) Total RD 1919-1969

In a similar fashion to Yee et al (2002), Total Ag. R\&D Expenditures at state level for years 1919-1969 in constant 1949 US\$ were calculated as an expansion of the SAES Expenditures in constant 1949 US\$ by the average ratio over 1970-1980 of Total RD to SAES Expenditures in constant 1949 US\$. SAES Expenditures for 1970-1980 were calculated from the Current Research Information System (USDA) as an aggregate of the following concepts:

1. CSREES Administered Funds (for SAES only)
1.1. Hatch Act
1.2. Evans Allen Act
1.3. Animal Health
1.4. Grants and Agricultural Markets
1.5. National Research Initiative (NRI) Grants
1.6. Small Business Innovation Research (SBIR) Grants
1.7. Other CSREES Grants
2. Other USDA Funds (for SAES only)
3. Other Federal Funds (for SAES only)
4. State Appropriations (for SAES only)
5. Other Non-Federal Funds (for SAES only)

For most of the states, the ratio of Total RD to SAES Expenditures showed low variability over 1970-1980. The only states for which the coefficient of variation is greater than $10 \%$ are Wyoming, Colorado and Delaware. Even in these cases, the coefficients of variation are always below $16 \%{ }^{4}$

## 5) The stocks of $A g . R \& D$ at state level

The stock of R\&D for each state is constructed as a weighted average of the previous 30 years of Total RD in constant 1949 US\$, using an inverted-V pattern of weights (Figure 2). The weights are the marginal effects of public research expenditures on U.S. agricultural productivity reported by Chavas and Cox (1992). The weighting scheme implies that public R\&D expenditures incurred at year t start having some effect on agricultural productivity eight years later, with an ever increasing marginal effect until 25 years after being incurred, when the marginal effect reaches it maximum. The marginal effects of public R\&D expenditures on agricultural productivity die off to zero from year 26 to year 31 after being incurred.


Figure 2

[^3]
## (2) Procedure for Constructing the Spill-in Variable for Each State

In an attempt to capture the effects of the "stock of knowledge" in "similar" states on the agricultural productivity of a given state, a set of spill-in variables is constructed, according to a measure of "similarity" among the states. As before, the stock of knowledge is proxied by the Ag. R\&D Stock. Given that both the own-state stock of R\&D and the stock of R\&D in "similar" states are public inputs, we rely on the assumption of imperfect appropriability to distinguish between the effects each of the stocks has on agricultural productivity. Huffman and Evenson (1989, 1992, 1993 and 2001) and McCunn and Huffman (2000) assume perfect appropriability among the states within the same region, so the stock of knowledge is proxied by a variable that contains both a state’s own agricultural R\&D stock and spill-ins from similar states.

In the present study, as in Khanna et al (1994), Yee and Huffman (2001), Huffman et al (2002), and Yee et al (2002), the measure of similarity is the geographical proximity among the states, intended to be a proxy for climatic conditions, production conditions, input-output mixes, etc. ${ }^{5}$

The main difference with previous studies is that while they use an ad-hoc regional grouping of states (climatic, geopolitical, etc.) to construct the stocks of spillovers, the present study relies only on geographical proximity among states, regardless of the region they belong to. In particular, geographical adjacency is the criteria used to construct the spill-in stocks. The first spill-in stock variable is constructed as the aggregate of the stock of agricultural R\&D conducted in states sharing a common border with the state under analysis. This variable will be referred to as the "band 1 spill-in stocks". For example, the band 1 spill-in stock for Nebraska consists of the sum of the stocks of R\&D in Wyoming, South Dakota, Iowa, Missouri, Kansas and Colorado.

The second spill-in stock variable is constructed as the aggregate of the stock of agricultural R\&D conducted in states sharing a common border with the states on the band 1 spill-in stock. This variable will be referred to as the "band 2 spill-in stocks". For example, the band 2 spill-in stock for Nebraska consists of the sum of the stocks of R\&D

[^4]in New Mexico, Arizona, Utah, Idaho, Montana, North Dakota, Minnesota, Wisconsin, Illinois, Kentucky, Tennessee, Arkansas and Oklahoma.

The third spill-in stock variable, the "band 3 spill-in stocks", is the sum of the stock of agricultural R\&D conducted in states sharing a common border with the states on the band 2 spill-in stock. For example, the band 3 spill-in stock for Nebraska consists of the sum of the stocks of R\&D in Texas, California, Nevada, Oregon, Washington, Michigan, Indiana, Ohio, West Virginia, Virginia, North Carolina, Louisiana, Mississippi, Alabama and Georgia.

## Appendix 2. Nonlinear ITSUR estimation on transformed data



## Appendix 3. Nonlinear ITSUR Parameter Estimates

| Parameter | Approx |  |  | Approx |
| :---: | :---: | :---: | :---: | :---: |
|  | Estimate | Std Err | t Value | $\mathrm{Pr}>\mid \mathrm{t\mid}$ |
| DELTAMAL | 23254071 | 21867561 | 1.06 | 0.2877 |
| DELTAMAZ | -7.736E7 | 18910674 | -4.09 | <. 0001 |
| DELTAMAR | -1.502E8 | 38812281 | -3.87 | 0.0001 |
| DELTAMCA | -6.193E8 | 1.275 E 8 | -4.86 | <. 0001 |
| DELTAMCO | -1.737E8 | 36579264 | -4.75 | <. 0001 |
| DELTAMCT | -2.702E7 | 10865760 | -2.49 | 0.0130 |
| DELTAMDE | -5.536E7 | 11373873 | -4.87 | <. 0001 |
| DELTAMFL | -4.329E7 | 25904782 | -1.67 | 0.0949 |
| DELTAMGA | 90993683 | 24947435 | 3.65 | 0.0003 |
| DELTAMID | -2.094E8 | 30943089 | -6.77 | <. 0001 |
| DELTAMIL | -7.439E8 | 82754692 | -8.99 | <. 0001 |
| DELTAMIN | -3.348E8 | 49160517 | -6.81 | <. 0001 |
| DELTAMIA | -7.184E8 | 99918687 | -7.19 | <. 0001 |
| DELTAMKS | -4.244E8 | 61683231 | -6.88 | <. 0001 |
| DELTAMKY | -3.206E8 | 43099973 | -7.44 | <. 0001 |
| DELTAMLA | -1.734E8 | 27441504 | -6.32 | <. 0001 |
| DELTAMME | -3.202E7 | 11783162 | -2.72 | 0.0066 |
| DELTAMMD | 6681165 | 17565108 | 0.38 | 0.7037 |
| DELTAMMA | -2.29E7 | 10164387 | -2.25 | 0.0244 |
| DELTAMMI | -6.839E7 | 24054426 | -2.84 | 0.0045 |
| DELTAMMN | -2.938E8 | 45854806 | -6.41 | <. 0001 |
| DELTAMMS | -1.207E8 | 30787327 | -3.92 | <. 0001 |
| DELTAMMO | -2.73E8 | 48072759 | -5.68 | <. 0001 |
| DELTAMMT | -2.493E8 | 26913389 | -9.26 | <. 0001 |
| DELTAMNE | -3.578E8 | 56255859 | -6.36 | <. 0001 |
| DELTAMNV | -1.061E8 | 11993563 | -8.85 | <. 0001 |
| DELTAMNH | -5.92E7 | 12446071 | -4.76 | <. 0001 |
| DELTAMNJ | 5284324 | 11308020 | 0.47 | 0.6403 |
| DELTAMNM | -1.186E8 | 18682426 | -6.35 | <. 0001 |
| DELTAMNY | -6.357E7 | 28858360 | -2.20 | 0.0277 |
| DELTAMNC | -5.18E7 | 33079679 | -1.57 | 0.1175 |
| DELTAMND | -3.446E8 | 32537799 | -10.59 | <. 0001 |
| DELTAMOH | -2.48E8 | 37982238 | -6.53 | <. 0001 |
| DELTAMOK | -1.441E8 | 28585076 | -5.04 | <. 0001 |
| DELTAMOR | -3.487E8 | 50235808 | -6.94 | <. 0001 |
| DELTAMPA | 10586165 | 27390566 | 0.39 | 0.6992 |
| DELTAMRI | -5.602E7 | 9045215 | -6.19 | <. 0001 |
| DELTAMSC | -3.501E7 | 14920373 | -2.35 | 0.0191 |
| DELTAMSD | -3.315E8 | 37482016 | -8.85 | <. 0001 |
| DELTAMTN | -2.268E8 | 37432504 | -6.06 | <. 0001 |
| DELTAMTX | -7.642E8 | 1.1006 E 8 | -6.94 | <. 0001 |


| Parameter | Approx |  |  | Approx |
| :---: | :---: | :---: | :---: | :---: |
|  | Estimate | Std Err | t Value | $\mathrm{Pr}>\|\mathrm{t}\|$ |
| Deltamut | -1.938E8 | 26646722 | -7.27 | <. 0001 |
| DELTAMVT | -3.252E7 | 9965178 | -3.26 | 0.0011 |
| deltamva | -6.638E7 | 18971590 | -3.50 | 0.0005 |
| DELTAMWA | -2.903E8 | 45659504 | -6.36 | <. 0001 |
| Deltamwv | -5.736E7 | 11999386 | -4.78 | <. 0001 |
| DELTAMWI | -2.555E8 | 43613064 | -5.86 | <. 0001 |
| Deltamwy | -1.461E8 | 18856360 | -7.75 | <. 0001 |
| DELTAL | -4.48068 | 0.5304 | -8.45 | <. 0001 |
| DELTAK | 2.475968 | 0.5798 | 4.27 | <. 0001 |
| DELTAT | 20.40223 | 3.4161 | 5.97 | <. 0001 |
| DELTAY | 3.001588 | 0.2106 | 14.26 | <. 0001 |
| DELTAG | -20822.2 | 3790.5 | -5.49 | <. 0001 |
| DELTAS1 | -2459.76 | 744.7 | -3.30 | 0.0010 |
| DELTAS2 | 691.9051 | 402.8 | 1.72 | 0.0860 |
| DELTAS3 | 205.66 | 304.0 | 0.68 | 0.4987 |
| AL | 899633.2 | 6123229 | 0.15 | 0.8832 |
| AK | 79314547 | 15824063 | 5.01 | <. 0001 |
| AT | 4755965 | 2770149 | 1.72 | 0.0862 |
| CLG | 3058.467 | 1398.1 | 2.19 | 0.0288 |
| CLS1 | -833.636 | 404.3 | -2.06 | 0.0393 |
| CLS2 | 31.06303 | 185.0 | 0.17 | 0.8667 |
| CLS3 | 621.7079 | 155.5 | 4.00 | <. 0001 |
| CLY | 0.144716 | 0.1195 | 1.21 | 0.2259 |
| CKG | 3313.025 | 2129.7 | 1.56 | 0.1199 |
| CKS1 | -2127.66 | 1094.3 | -1.94 | 0.0520 |
| CKS2 | -421.647 | 534.4 | -0.79 | 0.4302 |
| CKS3 | 1490.583 | 449.8 | 3.31 | 0.0009 |
| CKY | -0.20164 | 0.2005 | -1.01 | 0.3147 |
| CTG | -1163.52 | 461.7 | -2.52 | 0.0118 |
| CTS1 | -672.902 | 155.2 | -4.34 | <. 0001 |
| CTS2 | -74.0007 | 80.3903 | -0.92 | 0.3574 |
| CTS3 | 242.356 | 66.9532 | 3.62 | 0.0003 |
| CTY | -0.16069 | 0.0398 | -4.04 | <. 0001 |
| ALL | 0.119516 | 0.00375 | 31.85 | <. 0001 |
| ALK | -0.00693 | 0.00521 | -1.33 | 0.1839 |
| ALT | -0.01653 | 0.00339 | -4.88 | <. 0001 |
| AKL | -0.03357 | 0.00304 | -11.06 | <. 0001 |
| AKK | 0.100288 | 0.00540 | 18.56 | <. 0001 |
| AKT | -0.02167 | 0.00391 | -5.54 | <. 0001 |
| ATL | 0.007273 | 0.000655 | 11.10 | <. 0001 |
| ATK | -0.00509 | 0.000988 | -5.16 | <. 0001 |
| ATT | 0.064422 | 0.000669 | 96.30 | <. 0001 |
| GLL | -7507311 | 3705779 | -2.03 | 0.0429 |


| Parameter | Approx |  |  | Approx |
| :---: | :---: | :---: | :---: | :---: |
|  | Estimate | Std Err | t Value | Pr > \|t| |
| GLK | -1506489 | 7251075 | -0.21 | 0.8354 |
| GLT | -387061 | 1489049 | -0.26 | 0.7949 |
| GKK | -8.148E7 | 17380809 | -4.69 | <. 0001 |
| GKT | -4914931 | 2285697 | -2.15 | 0.0317 |
| GTT | 269105.9 | 158531 | 1.70 | 0.0898 |
| BYG | -0.00025 | 0.000037 | -6.70 | <. 0001 |
| BYS1 | 0.000012 | 8.271E-6 | 1.41 | 0.1595 |
| BYS2 | 8.368E-6 | 3.919E-6 | 2.14 | 0.0329 |
| BYS3 | -0.00001 | $3.334 \mathrm{E}-6$ | -3.21 | 0.0013 |
| BGG | 1.572199 | 0.6698 | 2.35 | 0.0190 |
| BGS1 | -0.0233 | 0.1594 | -0.15 | 0.8838 |
| BGS2 | -0.11341 | 0.0734 | -1.54 | 0.1227 |
| BGS3 | 0.173314 | 0.0562 | 3.08 | 0.0021 |
| BS1S1 | 0.063276 | 0.0433 | 1.46 | 0.1445 |
| BS2S1 | 0.02749 | 0.0199 | 1.38 | 0.1664 |
| BS3S1 | -0.00201 | 0.0149 | -0.13 | 0.8927 |
| BS2S2 | -0.04905 | 0.0149 | -3.29 | 0.0010 |
| BS3S2 | 0.006957 | 0.00861 | 0.81 | 0.4193 |
| BS3S3 | -0.00307 | 0.00798 | -0.38 | 0.7004 |
| BLG | 0.00103 | 0.000108 | 9.49 | <. 0001 |
| BLS1 | 9.205E-6 | 0.000020 | 0.45 | 0.6513 |
| BLS2 | 9.982E-7 | 0.000011 | 0.09 | 0.9253 |
| BLS3 | -0.00001 | 7.905E-6 | -1.41 | 0.1582 |
| BKG | 0.000557 | 0.000130 | 4.29 | <. 0001 |
| BKS1 | -0.00009 | 0.000034 | -2.70 | 0.0070 |
| BKS2 | -5.86E-6 | 0.000017 | -0.34 | 0.7358 |
| BKS3 | 0.000047 | 0.000014 | 3.44 | 0.0006 |
| BTG | 0.001176 | 0.000243 | 4.84 | <. 0001 |
| BTS1 | -0.00007 | 0.000043 | -1.71 | 0.0869 |
| BTS2 | 0.000018 | 0.000024 | 0.77 | 0.4420 |
| BTS3 | 0.000014 | 0.000017 | 0.84 | 0.3996 |
| AR_l1_1_1 | 0.254845 | 0.0268 | 9.49 | <. 0001 |
| AR_l1_2_2 | 0.698231 | 0.0218 | 32.08 | <. 0001 |
| AR_l1_4_4 | 0.470896 | 0.0389 | 12.11 | <. 0001 |
| AR_l1_3_3 | 0.366494 | 0.0271 | 13.55 | <. 0001 |


[^0]:    ${ }^{1}$ See Table A-4 in Acquaye et al.; Table 2 in Ball et al (1999); Table 4-5 in Alston and Pardey; Table 1 in Aheran et al; Table 2 in Shane et al; Table 7.a in Huffman and Evenson (the compound annual growth rate of agricultural TFP for $1950-1990$ is $2.06 \%$ ).

[^1]:    ${ }^{2}$ Although we could have continued testing for higher order spatial autocorrelation lags, we decided that most of the spatial effects would be captured with five spatial lags.

[^2]:    ${ }^{3}$ McIntire-Stennis Cooperative Forestry. 16 U.S.C. 582a, et seq. McIntire-Stennis Cooperative Forestry allocates funds on a formula basis for forestry research, which includes forests and related rangelands, at institutions offering graduate training in the sciences basic to forestry or having a forestry school. Eligible institutions are designated by the State. A 100 percent non-federal match is required. (USDA, 2005)

[^3]:    ${ }^{4}$ Time series methods were explored but the standard errors for earlier years were huge, making it as adhoc as the used method.

[^4]:    ${ }^{5}$ Alston et al (2002) use a different measure of similarity, based on technological proximity across states according to their output mixes rather than geographical proximity.

