The Economic Returns to U.S. Public Agricultural Research

Julian M. Alston, Matthew A. Andersen, Jennifer S. James, and Philip G. Pardey

We use newly constructed state-specific data to explore the implications of common modeling choices for measures of research returns. Our results indicate that state-to-state spillover effects are important, that the research and development lag is longer than many studies have allowed, and that misspecification can give rise to significant biases. Across states, the average of the own-state benefit-cost ratios is 21:1, or 32:1 when the spillover benefits to other states are included. These ratios correspond to real internal rates of return of 9% or 10% per annum, much smaller than those typically reported in the literature, partly because we have corrected for a methodological flaw in computing rates of return.

In the United States, public support for investments in agricultural R&D continues to wane in spite of consistently high reported rates of return to agricultural R&D. This apparent paradox could simply reflect government failure, but it might also reflect skepticism about the evidence. Certainly some public policymakers and some economists—ourselves among them—are skeptical about the very high rates of return reported by some studies, and "gilding the lily" might have damaged the case for public support (Alston et al. 2000).

Data limitations require the imposition of restrictive assumptions that have unknown implications for estimation bias, but upward biases may also have resulted from particular modeling choices that were not made necessary by data constraints (Alston and Pardey 2001). This paper reports the main results from a long-running project in which we set out to obtain new and improved estimates of the returns to U.S. public agricultural research and development (R&D), to evaluate the role of modeling choices versus fundamental factors in influencing the findings and thus to provide a clearer understanding of the confidence that can be placed in the estimates.

To explore the consequences of common modeling choices and their implications for measures of research returns, we make use of an uncommonly rich and detailed panel of state-level data, which we developed for this purpose. It includes annual state-specific data on agricultural productivity for each of the 48 contiguous U.S. states over the years 1949–2002 and on agricultural research and extension expenditures by the federal and state governments over the years 1890–2002. The indexes of multifactor productivity (MFP) are Fisher ideal discrete approximations of Divisia indexes that reflect a careful effort to account for variation over time and among states in the composition of the aggregates of inputs and outputs and thereby minimize the role of index number problems. In our econometric models, we pay particular attention to the specification of the research lag structure and models of spatial spillovers, but to illustrate the role of fundamental factors, we compare the resulting estimates with simple approximations that abstract from the detail of the spatial and temporal aspects.

A more complete description of the data, models, and many of the results discussed here can be found in the study by Alston et al. (2010). Beyond presenting a succinct synthesis of the main results found by Alston and colleagues, we here extend that work in two important ways. First, we present new evidence on the time-series properties of the models, which provides additional support regarding the robustness of the results. Second, we present alternative measures of the rate of return to the investments, demonstrating why many of the previous results in the literature should be treated with skepticism.

Modeling Agricultural Research and Productivity

At the center of our empirical work is a model of state-specific productivity growth as a function of investments in agricultural research, built on foundations laid by Griliches (1964, 1979) and Evenson (1967), among others, as reviewed by Alston, Norton, and Pardey (1998) and Alston et al. (2010). Underlying the productivity patterns are changes in aggregate measures of inputs and outputs. In 2002, U.S. agriculture produced 2.6 times the quantity of output produced in 1949. It did this with marginally less aggregate inputs such that MFP grew faster than output. Our estimates indicate that output from agriculture increased on average by 1.68% per year over the period 1949–2002, while inputs used by agriculture declined by 0.11% per year; so, measured MFP grew by 1.78% per year. These averages reflect patterns of input and output growth that varied dramatically among the 48 contiguous states. Some states had both inputs and outputs growing, some had both falling, but the majority had output growing against a declining input quantity, and all had positive rates of MFP growth over the years 1949–2002, which ranged from 0.84% per year in Wyoming to 2.48% per year in North Carolina.

Investments in agricultural research and extension also evolved dramatically over the period of our analysis, with important changes both in the emphasis among federal, state, and local government and private sources of funding and in the balance of effort among performing agencies. We use state-specific panel data on investments in publicly performed research since 1890 and in extension since 1915, to develop research and extension knowledge stocks to be used in models of productivity over the years 1949–2002. Over that period total expenditure on public research and extension grew dramatically in total but unevenly. The intensities of spending on research and extension conducted by state government institutions have become quite varied, reflecting differences among states in growth in agricultural production, as well as in their investments in the creation and diffusion of knowledge.

Model Structure

We begin with a model in which agricultural productivity in every U.S. state (excluding Alaska and Hawaii) depends on past agricultural research and extension conducted by itself and every other U.S. state (a total of 48 states) and intramural research conducted by the USDA.[1] We can express this model in general terms, mathematically, as

$$(1) \qquad MFP_{i,t} = \mathrm{f}_i(\mathbf{R}_t, \mathbf{E}_t)$$

where $MFP_{i,t}$ is multifactor productivity in state i in year t, and $\mathbf{R}_t$ is a $49 \times (L_R + 1)$ matrix in which the typical element, $R_{j,t-k}$, is the investment in public agricultural R&D made by state j (for $j = 1, \ldots, 48$) or the USDA (for $j = 49$) in year t–k; similarly, $\mathbf{E}_t$ is a $48 \times (L_R + 1)$ matrix in which the typical element, $E_{j,t-k}$, is the investment in public agricultural extension made by state j in year t–k; L_R denotes the maximum number of years over which a given investment can affect MFP; and k varies between zero and L_R.

To implement this model we have to define the knowledge stock variables, which requires

[1] There are various ways to account for R&D activity, but with an eye to the policy implications of the results, our intent here is to evaluate the impacts of agricultural R&D on a "by performer" basis (as distinct from a "by funder" or other basis). The Organisation for Economic Co-operation and Development (2002) provides details on the internationally accepted standards for measuring R&D spending and performance, which we followed in compiling our R&D data.

jointly defining the spillover relationships (which allow for research conducted in one state to affect productivity in another) and the research lag distributions (which summarize the temporal relationship between spending and productivity). Even if the research and extension lag lengths were modest, they would imply an impossibly large number of research effects to estimate, so some restrictions must be imposed, as has been long recognized in studies of returns to R&D. Griliches (1979) suggested that

> it is probably best *to assume* a functional form for the lag distribution on the basis of prior knowledge and general considerations and not to expect the data to answer such fine questions. That is, a "solution" to the multicollinearity problem is a moderation of our demands on the data—our desires have to be kept within the bounds of our means. (p. 106, emphasis in original)

In our particular setting, the potential problems of multicollinearity and identification are many times greater than in the typical study using a single time series, since we have allowed for 48 states with interstate spillovers in every direction.

Previous econometric studies of effects of agricultural research on productivity, or rates of return to research, have almost invariably imposed some structure (often implicitly) to reduce the number of lag weights to be estimated and to impose other prior beliefs on the shape or length of the lag (see Alston et al. 2000 for details). Like most previous studies, and as advocated by Huffman and Evenson (2006a), we impose some restrictions on the lag distribution, to reduce the number of parameters to be estimated. First, we assume that as a baseline model, agricultural research and extension expenditures (at least at the margin) are fungible and can be combined into a single aggregate research and extension variable to which the same lag distribution would apply.[2] Second, we assume that the same-shaped lag distribution applies to a state's research and extension, regardless of who is adopting the results. Thus, productivity in each of the 48 states depends on 49 state-specific knowledge stocks (one own-state research stock, 47 other-state research stocks, and one federal research stock). Third, we assume that same lag shape applies to all the states within a given model. However, we do estimate the parameters that define the shape and effective length of the lag, and in that sense our approach is less restrictive than others that simply imposed a specific distribution *a priori* (such as the trapezoidal lag that was introduced by Huffman and Evenson [1993] and has been applied by many others since). In addition, we explore the implications of relaxing several of the baseline modeling assumptions.

Consequently, in the baseline model, the relationship between spending on research and extension and the knowledge stock produced within state i in year, t, $SK_{i,t}$, can be characterized using a single lag distribution, defined in terms of (a) an overall lag length, (b) a gestation lag, (c) a functional form (we used a gamma distribution), and, (d) within the functional form, parameters that determine the shape of the distribution, as follows:[3]

$$SK_{i,t} = \sum_{k=0}^{L_R} b_k (R_{i,t-k} + E_{i,t-k}), \tag{2}$$

where L_R is the total lag length, and the b_k parameters are the lag weights that are defined by the alternative lag distributions, and these weights sum to one:

$$\sum_{k=0}^{L_R} b_k = 1. \tag{3}$$

[2] This assumption may appear rather strong, but some aggregation assumptions are necessary and are always made in work of this nature. The research and extension variables themselves represent aggregates over different types of activities having different lagged impacts on productivity (ranging from relatively basic research to applied outreach and advisory services, a significant share of which may have no relationship to production agriculture, and across different fields of science that may have more or less relevance to production agriculture). We examine the empirical implications of this assumption later in the paper when testing the robustness of our preferred model.

[3] Our analysis is not confined to this baseline model and its particular assumptions. Importantly, unlike previous studies that simply impose assumptions about the research and extension lag, we examine the implications of alternatives, including the arbitrary and untested imposition of a particular, short lag distribution shape for extension combined with a specific trapezoidal lag distribution for research, as used by Huffman and Evenson (1993) and others. The results, discussed in the paper, did not especially favor the use of separate lags for extension, while serving to illustrate the implications of that choice for the estimated rates of return and the difficulty of discriminating among such alternatives using the kinds of data that are available.

Gamma Lag Distribution Model

The research lag weights (b_k) implied by the gamma distribution are:

$$(4)\qquad b_k = \frac{(k-g+1)^{(\delta/1-\delta)}\lambda^{(k-g)}}{\sum_{k=0}^{L_R}[(k-g+1)^{(\delta/1-\delta)}\lambda^{(k-g)}]}$$

$$\text{for } L_R \geq k > g;\ \text{otherwise } b_k = 0$$

where g is the gestation lag before research begins to affect productivity, and δ and λ are parameters that define the shape of the distribution ($0 \leq \delta < 1$ and $0 \leq \lambda < 1$). Here, we assume a gestation lag of $g = 0$ years, but several distributions defined by combinations of δ and λ that we use imply weights very close to zero for small values of k, resulting in a longer effective gestation lag. In addition, based on our own previous experience with similar data and models (see, e.g., Pardey and Craig 1989) and some limited pretesting as a part of the present study, as well as a predisposition to allow for generously long lags, we allow for $L_R = 50$ years. The resulting lag distribution allows for positive contributions to the current stock from up to 50 years of past expenditures on research and extension, but particular values of λ and δ can correspond to a pattern of very low b_k parameters, after a time, that imply a much shorter effective maximum lag. Hence, the research knowledge stocks are defined as

$$(5)\qquad SK_{i,t} = \sum_{k=0}^{L_R} b_k(R_{i,t-k} + E_{i,t-k})$$

$$b_k = \frac{(k+1)^{(\delta/1-\delta)}\lambda^{(k)}}{\sum_{k=0}^{L_R}\left[(k+1)^{(\delta/1-\delta)}\lambda^{(k)}\right]}.$$

Spillover Weights Based on Similarity of Commodity Composition

Previous studies have imposed various (largely untested) assumptions to define the interstate spillover impacts of agricultural research and extension investments. As discussed by Alston (2002) and Alston et al. (2010), many studies simply ignored spatial spillovers, attributing all state-specific impacts to own-state investments, while those studies that have allowed for interstate spillovers have generally defined spillover potential based on physical proximity. Here, as a departure from those previous approaches, we use a measure of spillover potential based on the similarity of the commodity composition of output between pairs of states, and we evaluate the implications of this assumption for results compared with the main alternatives.[4]

We assume a linear state-to-state spillover relationship, and define

$$(6)\qquad SS_{i,t} = \sum_{j\neq i}^{49} \omega_{ij} SK_{j,t}$$

where ω_{ij} is a spillover coefficient, a weight that measures the contribution of a unit of the knowledge stock created in state j to the knowledge stock used in state i. To define the spillover coefficients, which measure the state-to-state spillover potential of agricultural research and extension, we borrow and adapt an approach introduced by Jaffe (1986) to measure interfirm or interindustry spillover effects. The variant used by Jaffe (1989) is closest to what we use here. Jaffe (1989) used characteristics of the patents obtained by firms to define a measure of technological closeness among them. We use the output characteristics of agriculture in the different states—representing agro-ecological and other relevant economic factors—to define the technological "closeness" of states to one another. The vector of output (value) shares $f_i = (f_{i1}, \ldots, f_{iM})$ locates state i in M-dimensional technological space. The corresponding measure of technological spillover potential is defined as:

$$(7)\qquad \omega_{ij} = \frac{\sum_{m=1}^{M} f_{im}f_{jm}}{\left(\sum_{m=1}^{M} f_{im}^2\right)^{1/2}\left(\sum_{m=1}^{M} f_{jm}^2\right)^{1/2}}$$

where f_{im} is the value of production of output m as a share of the total value of agricultural output in state i such that these shares fall between zero and one and sum to one (i.e., there are a total of M different outputs across the 48 states, and $0 \leq f_{im} \leq 1$ and Σ_m

[4] The notion here is that research spillovers among states producing similar or identical commodity portfolios are likely to be more pronounced than among states producing dissimilar or distinct sets of agricultural outputs. Thus, two predominantly dairy production states are more likely to be doing research of relevance to each other than if one state produced only milk and the other only oranges. To be sure, dairy (and other) production details vary from state to state for a host of reasons, but it is unlikely that the dairy research in New York has no application to dairy production in Minnesota or California, as would be implied by the geographical proximity restriction incorporated in the approach used by Huffman and Evenson (1993, 2006b), for example.

$f_{im} = 1$).[5] To define corresponding "spillover coefficients" for measuring the state-specific impacts of USDA research stocks (i.e., $\omega_{iF} = \omega_{i49}$, for $i = 1, \ldots, 48$), we apply equation (7) to index the similarity of each state's vector of output shares and the national vector of output shares.[6] Then, given this specification of the state-to-state spillover relationships, instead of 47 individual other-state knowledge stocks and a federal knowledge stock in the regression model, we use a single research spillover stock tailored for each state, to represent the aggregation of those 48 spill-in effects in that state.

Econometric Estimation and Results

Assuming a logarithmic functional form and augmenting the model to include state-specific fixed effects and a variable to reflect the effect of weather, the model becomes

$$\text{(8)} \qquad \ln MFP_{i,t} = \alpha_{0i} + \alpha_R \ln SK_{i,t} + \alpha_S \ln SS_{i,t} + \gamma \ln PRC_{i,t} + e_{i,t}.$$

The variables are defined as follows:

1. $MFP_{i,t}$ is a Fisher ideal index (i.e., a discrete approximation of a Divisia index) of multifactor agricultural productivity in state i in year t;
2. $SK_{i,t}$ is the own-state stock of knowledge in state i in year t from own-state spending on publicly performed agricultural research and extension over the previous 50 years, in real terms;
3. $SS_{i,t}$ is the state-specific spillover stock of knowledge in state i in year t from spending on agricultural research and extension conducted by federal and other-state public institutions over the previous 50 years, in real terms, constructed using the same lag distribution parameters as for $SK_{i,t}$;
4. $PRC_{i,t}$ is a state-specific pasture and rangeland condition index, measured in September for each year and published by the Economics, Statistics, and Market Information System of the USDA; and
5. $e_{i,t}$ is a residual, with an independent and identically distributed structure.

   Simple summary statistics are presented in table 1.

Notably, the specification in equation (8) does not include any variables to represent the stocks of knowledge from private agricultural research conducted in the United States or internationally, public agricultural research conducted in other countries, or nonagricultural research. The reason for excluding these variables is that appropriate data in suitably long time series simply are not available. The omission of these variables could lead to biases in the estimated effects of the included knowledge stocks if the omitted stocks are correlated with the included stocks. However, private research effects are embodied largely in inputs, and to the extent that the benefits are captured through royalties or the equivalent, they might not have much impact on measured productivity compared with an equivalent public research achievement provided to farmers and others for free. In addition, our adjustments for changes in input and output quality will have dealt with some of these impacts. This view is supported to some extent by some recent work by Huffman and Evenson (2006a).[7] Even so, we are conscious of the potentially biasing effects of omitting private agricultural R&D (as well as omitting U.S. nonagricultural research and

---

[5] A referee raised several questions about the structure and interpretation of these coefficients related to whether they may vary over time and, as a related point, the direction of causality (if agricultural R&D affects production patterns, then the weights are endogenous). The relevant issues are too many to deal with in the space available here, but many of them are discussed in some detail by Alston et al. (2010). In the present application, we use the average value of ω_{ij} for the sample period, which is simply a measure of the overall similarity of the agricultural output mix between states, as a proxy for the state-to-state spillover potential of agricultural research and extension. We compare the model using this specification with alternative specifications similar to those typically used in the literature.

[6] Paraphrasing Jaffe (1989, p. 88), in a sense ω_{ij} measures the degree of overlap of f_i and f_j. The numerator will be large when states i and j have very similar output mixes. The denominator normalizes the measure to be one when f_i and f_j are identical. Hence, ω_{ij} will be zero for pairs of states with no overlap in their output mix and one for pairs of states with an identical output mix; and for the in-between cases, $0 < \omega_{ij} < 1$. It is conceptually similar to a correlation coefficient. Like a correlation coefficient, it is completely symmetric: $\omega_{ij} = \omega_{ji}$, and $\omega_{ii} = 1$.

[7] Most studies of the effects of public agricultural research on productivity have not incorporated an explicit measure of private research. In a significant and rare exception, Huffman and Evenson (2006a) attempted to account for private research effects in an analysis using U.S. state-level data (see also Huffman and Evenson 1993, 2006b). In the absence of suitably long time series of private research expenditures, they used state-specific production weights applied to four classes of commodity-specific patent data to define state-specific annual flows of private research outputs, which they aggregated into state-specific stocks by applying trapezoidal lag weights over a 19-year period and summing. The resulting measure of "private agricultural research capital" did not make a statistically significant contribution to either of the productivity models that Huffman and Evenson (2006a) reported.

Table 1. Simple Summary Statistics and Data for the Productivity Model

Symbol	Variable Name	Definition	Value Description	Value
$MFP_{i,t}$	Multifactor Agricultural Productivity	Fisher ideal index agricultural output in state i and year t	Minimum across all years and states	74.74
			Maximum across all years and states	481.83
			Average across years	
			All states	181.30
			California	176.14
			Minnesota	173.40
			Wyoming	142.24
$SK_{i,t}$	Own-State Stock of Knowledge	Constructed using 50 years of own-state government spending on agricultural research and extension (in real 2000 dollars) and specification of gamma lag distribution	Preferred lag distribution ($\lambda = 0.70, \delta = 0.90$)	
			Minimum across all years and states	\$5.0 million
			Maximum across all years and states	\$104.8 million
			Average across all years and states	\$33.1 million
$SS_{i,t}$	State-Specific Spill-in Stock of Knowledge	Constructed using federal and other-state government spending on agricultural research and extension (in real 2000 dollars), specification of lag distribution, and ω_{ij}-values used as weights	Preferred lag distribution ($\lambda = 0.70, \delta = 0.90$)	
			Minimum across all years and states	\$548.2 million
			Maximum across all years and states	\$1,436.0 million
			Average across all years and states	\$1,050.9 million
$PRC_{i,t}$	Pasture and Rangeland Condition Index	Measured in September for each year (published by the Economics, Statistics, and Market Information System of the USDA)	Minimum across all years and states	8
			Maximum across all years and states	107
			Average across years	
			All states	74.33
			California	73.27
			Minnesota	73.39
			Wyoming	78.02

Table 2. Summary of Results for the Base Model, Top-Ranked Models

Model Details	Model Results							
Model rank by SSE	1	2	3	4	5	6	7	8
Lag Distribution Characteristics								
λ	0.70	0.65	0.80	0.75	0.85	0.90	0.60	0.80
δ	0.90	0.90	0.85	0.85	0.80	0.75	0.90	0.80
Peak lag year	24	20	24	19	24	27	17	17
Elasticities with Respect to								
Own-state SAES	0.15	0.13	0.15	0.13	0.15	0.16	0.12	0.14
Own-state extension	0.18	0.15	0.18	0.16	0.18	0.19	0.13	0.15
All own-state combined	**0.32**	**0.28**	**0.33**	**0.29**	**0.33**	**0.35**	**0.25**	**0.29**
SAES spill-ins	0.07	0.09	0.07	0.10	0.07	0.06	0.12	0.10
Intramural spill-ins	0.07	0.10	0.07	0.10	0.07	0.06	0.11	0.10
Extension spill-ins	0.09	0.11	0.08	0.11	0.08	0.07	0.13	0.12
All spill-ins combined	**0.24**	**0.31**	**0.22**	**0.30**	**0.22**	**0.19**	**0.36**	**0.31**

Notes: SSE (sum of squared errors) indicates the goodness of fit of the model. The elasticities are of multifactor productivity with respect to the knowledge stock specified.

international research), and we explore this issue in later sections.

The models were estimated using various estimation procedures in Stata 10.0 with the International Science and Technology Practice and Policy Center (InSTePP) Production Accounts, version 4, as described by Pardey et al. (2009).[8] We used a type of grid-search procedure, in which we assigned values for the parameters of the gamma lag distribution (λ and δ), then constructed the knowledge stock variables using these parameters along with the expenditures on research and extension and the spillover coefficients (ω_{ij}), and then estimated the model using these constructed stocks.[9] By repeating this procedure using different values for λ and δ, we were able to search for the values of these parameters that, jointly with the estimated values for the other parameters, would best fit the data. Combining the following eight possible values for both λ and δ (0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95), a fixed maximum lag (50 years in most cases), and no gestation lag, yields a total of 64 possible combinations. A very wide range of shapes and effective lag lengths are encompassed by the range of parameter values tried. We initially thought we might conduct a further search over a finer grid, but upon review of our econometric results with the 64 lag distributions, we concluded that it would not be informative to do so because the top-ranked models were nearly indistinguishable.

The base model treats state-specific research and extension symmetrically, such that the same lag weights and spillover coefficients apply to both research and extension. This model was estimated using ordinary least squares with state-specific intercepts, which is a fixed-effects panel data estimator, for each of the 64 lag distributions.[10] Table 2 summarizes the main results for the highest-ranked eight models, arranged in order according to criteria of goodness-of-fit (sum of squared errors), highest to lowest from left to right. The best-fitting model was obtained with values for $\lambda = 0.70$ and $\delta = 0.90$, implying a peak lag weight at year 24, as seen in figure 1.[11] Among the models in table 2, the shape of the lag distribution was fairly similar across the top-ranked models compared with other models that did not fit as well. The peak lag varied somewhat, but the implied values for the elasticities of *MFP* with respect to the various knowledge stocks were very similar across the eight models—about 0.32 for own-state research and about 0.24 for spill-ins.

[8] Version 4 of the InSTePP Production Accounts represents a revised and updated version of the data used by Acquaye, Alston, and Pardey (2003) and originally developed and used by Craig and Pardey (1996). Pardey et al. (2009) provide further details on the construction of these data, which are presented and discussed by Alston et al. (2010).

[9] This approach of estimating productivity models with preconstructed research knowledge stocks is standard in much of the relevant previous work. Our important departure is to search across the range of possibilities for the lag distribution used to construct that stock and test among them, rather than simply impose one.

[10] We established that a fixed-effects estimator was preferred to a random-effects estimator using Hausman's (1978) specification test for fixed or random effects. The results are presented, along with additional diagnostic tests, by Alston et al. (2010, table 10-4).

[11] Figure 1 also shows the trapezoidal lag structure used by Huffman and Evenson (1993) and many others, plus a parameterization of our gamma distribution that closely approximates this specific trapezoidal form.

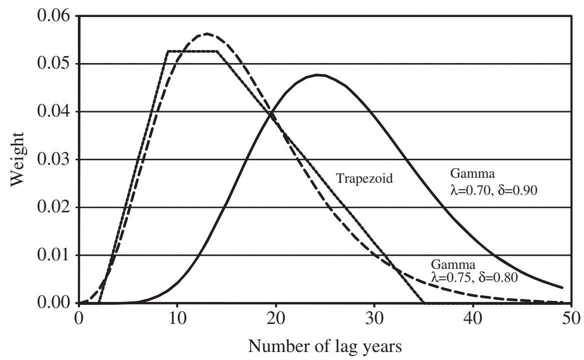

Figure 1. Gamma and Trapezoidal lag distributions.

Model Diagnostics

The analysis generally resulted in highly significant coefficient estimates; however, we also sought to verify the consistency of the estimates after controlling for some additional econometric issues. Specifically, we were concerned with autocorrelation of the residuals and unit roots in the state-specific *MFP* series, which can result in spurious parameter estimates. To address this concern, we reestimated the models in first-difference form (i.e., with all of the variables specified in logarithmic differences and absent an intercept term). In most cases first-differencing resulted in similarly shaped preferred lag distributions, as well as similar estimates of elasticities for a given lag distribution shape compared with the base models.

Heteroskedasticity and contemporaneous correlation were not of primary concern given the large sample size, the asymptotic properties of the estimators, and the statistical significance of the coefficient estimates. However, we also estimated the full grid of lag distributions using a feasible generalized least squares (FGLS) procedure that corrected for heteroskedasticity within states, contemporaneous correlation among states, and first-order autocorrelation of the residuals. The estimated elasticities from the FGLS estimation procedures were very similar to those from the fixed-effects regressions presented in table 2. For example, in the base model with the preferred lag distribution, the elasticity of *MFP* with respect to the own-state knowledge stock was 0.311 using an FGLS regression procedure and 0.322 using a fixed-effects regression procedure, while the corresponding elasticities with respect to spillover knowledge stocks were 0.241 and 0.235, respectively. Given the pattern of similarities of the estimates in the FGLS models to those in the base fixed-effects models, we were fairly certain that any autocorrelation of the residuals and issues of unit roots related to the state-specific *MFP* series were not substantially affecting the consistency of the fixed-effects estimates.[12]

Time-Series Properties of the Data

We also performed a formal statistical analysis of the time-series properties of the variables and a check of the consistency of the parameter estimates. A visual inspection of the data reveals that the series appear to be highly nonstationary. However, if the variables in the analysis are nonstationary but share the same order of integration (e.g., $I(1)$ behavior), then they might approach the same stochastic trend. In this case a linear combination of the variables could form a long-run equilibrium relationship that is stationary, and a linear regression of *MFP* on the research stock variables would result in consistent parameter estimates. Indeed, if the series were $I(1)$ and cointegrated, then the estimates would be superconsistent. Stock (1987) showed that such estimates converge to their probability limits faster than least squares estimates in stationary time-series models.

To establish whether our models produce consistent parameter estimates, we proceeded by first examining the data for the presence of a unit root and then establishing whether a cointegrating relationship exists between the variables. As a starting point, we applied augmented Dickey–Fuller tests on a state-by-state basis for the three variables of interest, $\ln MFP_{i,t}$, $\ln SK_{i,t}$, and $\ln SS_{i,t}$. Most of the state-specific data indicated the presence of a unit root, with the $\ln MFP_{i,t}$ series indicating a unit root in 41 of the 48 states, $\ln SK_{i,t}$ in 36, and $\ln SS_{i,t}$ in all 48.[13]

Given the evidence of unit roots in these data, it is possible that a long-run equilibrium relationship exists between the series; therefore, a test of a cointegrating relationship between the variables is warranted. Westerlund (2007) developed a test of cointegration in panel data that produces four test

[12] Additional comparisons of alternative estimators and models, as well as detailed diagnostic testing of the econometric models, are provided by Alston et al. (2010, chapter 10).

[13] In each augmented Dickey–Fuller test, we set $\alpha = 5\%$ and included an intercept, trend, and three lags of the dependent variable.

statistics, $G_\alpha, G_\tau, P_\alpha$, and P_τ.[14] Consider the following error-correction model with one lag of the dependent variable and one covariate:

$$\Delta y_{it} = \alpha_i + \beta_{i1}\Delta y_{it-1} + \delta_{i0}\Delta x_{it} + \delta_{i1}\Delta x_{it-1} + \beta_i(y_{it-1} - \delta_i x_{it-1}) + u_{it}. \tag{9}$$

The parameter β_i provides an estimate of the speed of adjustment toward the long-run equilibrium, and if $\beta_i = 0$, then there is no error correction, and thus no cointegrating relationship between the variables. The G_α and G_τ test statistics begin with a weighted average of the state-specific β_i parameters and their t-ratios and test the null hypothesis that $\beta_i = 0$ for all i versus the alternative that $\beta_i < 0$ for at least one i. The P_α and P_τ test statistics pool the sample over all the states and test the null hypothesis that $\beta_i = 0$ for all i versus the alternative that $\beta_i = \beta < 0$ for all i. If the observations are correlated between cross-sectional units (states), robust critical values can be obtained through a bootstrapping procedure. Table 3 reproduces the panel data tests of cointegration developed by Westerlund (2007, table 7), including the group mean, G_α and G_τ, as well as the pooled test statistics, P_α and P_τ. Panel (a) of table 3 shows the test results between $\ln MFP_{i,t}$ and $\ln SK_{i,t}$, and panel (b) shows the test results between $\ln MFP_{i,t}$ and $\ln SS_{i,t}$.

The null hypothesis of these tests is that there is no cointegration between the variables. The results of the Westerlund tests generally indicate that a cointegrating relationship exists between *MFP* and the research stock variables. The only results indicating rejection of a cointegrating relation are the bootstrapped versions of G_α and P_α between $\ln \mathrm{MFP}_{i,t}$ and $\ln \mathrm{SS}_{i,t}$, but these results are not as robust as the *tau* versions of the tests, which indicate that a cointegrating relationship does exist.[15] Furthermore, the 48-state average value of the estimated speed-of-adjustment parameters (i.e., the $\hat{\beta}_i$ parameters) is equal to -0.71 in the panel (a) results, and -0.83 in the panel (b) results.[16] The test results in table 3 provide convincing evidence that the models specified in logarithms produce superconsistent parameter estimates.

Table 3. Results from Panel Data Tests of Cointegration

Statistic	Value	Z-value	P-value	Robust P-value
Panel (a): Tests between $\ln MFP_{i,t}$ and $\ln SK_{i,t}$ (average AIC selected lag length is 2.04 years)				
G_τ	−3.93	−16.64	0.00	0.00
G_α	−12.32	−6.55	0.00	0.00
P_τ	−21.60	−11.24	0.00	0.00
P_α	−11.11	−10.20	0.00	0.00
Panel (b): Tests between $\ln MFP_{i,t}$ and $\ln SS_{i,t}$ (average AIC selected lag length is 2.25 years)				
G_τ	−4.32	−19.67	0.00	0.00
G_α	−10.26	−3.91	0.00	0.26
P_τ	−22.45	−12.08	0.00	0.00
P_α	−8.85	−6.78	0.00	0.16

Notes: AIC = Akaike information criterion. All calculations were done using Stata and the "xtwest" command. The null hypothesis is that there is no cointegration. In each test we included an intercept term and set the lag length according to the average AIC. The robust P-values correct for cross-sectional dependence among the states using a bootstrapping procedure. One hundred replications were used for the bootstrapping procedure, and the lag length was set to one.

Finally, the formal statistical test results are supported by the empirical observation that our models, specified either in logarithmic or in growth-rate form, produce similar results in terms of the estimated elasticities.

Marginal Benefit-Cost Ratios—Base-Model Results

We used the estimated productivity model to compute the *marginal* benefit associated with various hypothetical (counterfactual) changes in research investments. Specifically, we computed the state-specific and national benefits from a small ($1,000) change in 1950 in expenditures (a) on research by a particular state, (b) on extension by a particular state, or (c) on USDA intramural research by the federal government. In computing the national benefits, we took into account that both federal and state-specific research investments have effects on all the states.[17]

The gross annual research benefits ($GARB$) to state i in year t were computed using the following approximation:

$$GARB_{i,t} = \Delta \ln MFP_{i,t} V_{i,t} \tag{10}$$

[14] See also Persyn and Westerlund (2008) for additional details about this test and its implementation in Stata.

[15] Based on Monte Carlo simulations assuming cross-sectional dependence, Westerlund (2007, p. 730) reported that: "At [a]t At not [a]t one end of the scale, we have the G_τ and P_τ tests, which actually appear to be quite robust to the cross-sectional correlation."

[16] These results pertain to the models that do not correct for contemporaneous correlation among the states.

[17] This explicit simulation approach is less prone to error or misinterpretation than is an analytically derived approximation to a rate of return, as some studies have used.

where $V_{i,t}$ is the real value (in year 2000 dollars) of agricultural production in state i in year t, and $\Delta \ln MFP_{i,t}$ is the proportional change in estimated agricultural productivity in state i in year t, associated with the simulated \$1,000 increase in spending in 1950.[18] Since the variables are in logarithms, the simulated proportional change in MFP is simply equal to $\Delta \ln MFP = \ln MFP^1 - \ln MFP^0$, where the superscript 0 denotes the predicted $\ln MFP$ given the actual research expenditure and the superscript 1 denotes the predicted $\ln MFP$ with the increased (counterfactual) expenditure. Then, the present value in the year 2002 of benefits accruing to state i (PVB_i) was computed using a (correspondingly real) discount rate of $r = 3\%$ per year:

$$(11) \quad PVB_i = \sum_{t=1950}^{2002} GARB_{i,t}(1+r)^{2002-t}$$
$$= \sum_{t=1950}^{2002} \Delta \ln MFP_{i,t} V_{i,t} (1+r)^{2002-t}.$$

The benefit-cost ratio for that \$1,000 investment is given by dividing the present value of benefits by the present value of the costs: $PVC = \$1,000(1+r)^{53}(= \$4,650 \text{ for } r = 3\%)$. Hence, marginal benefit-cost ratios (or benefits per dollar of additional expenditure) were computed as $B_i^h/C_i^h = PVB_i^h/\$4,650$, where the superscript h denotes which of (a) one of the 48 state-specific research expenditures, (b) one of the 48 state-specific extension expenditures, or (c) federal research expenditures was increased by \$1,000 in 1950 to generate the stream of benefits being evaluated.

We computed marginal benefit-cost ratios for increases in investments in research conducted by any of the 48 State Agricultural Experiment Stations (SAESs) or by the USDA itself.[19] Table 4 summarizes the benefit-cost ratios in terms of the regional averages (representing the simple average of the entries for the states within each region), the minimum, maximum, and simple average across the 48 states, and state-specific entries for some selected states (California, Minnesota, and Wyoming) for the base model (the top-ranked model 1 in table 2). In table 4, for each state and for each region, the entries in columns (1) and (2) are measures of the benefits, per dollar of expenditure, accruing to that state (column (1), the own-state benefits) and the nation as a whole (column (2), both the own-state benefits and the spillover benefits to the other 47 states) from an increase in state-specific research spending; column (3) shows the benefits accruing to each state, or region, or the country as a whole from an increase in USDA intramural research expenditures.[20] All of these figures are in common terms, expressing real, marginal benefits per dollar invested (associated with a small change in expenditure in 1950).

Within a row in table 4, comparing columns (1) and (2), we can compare the own-state payoff to that state from investing in research and extension versus the payoff to the nation as a whole; the difference between these two is the spillover benefit per dollar. This comparison indicates the magnitude of the distortion in incentives for a state to conduct the quantity and mixture of agricultural research that

---

[18] This approximation is likely to be reasonably valid as a measure of the total benefits for a small research-induced change in production, as a result of a comparatively small change in research investment. However, to the extent that these marginal changes in research spending induce price changes, the benefits will be distributed between producers and consumers, depending on the elasticities of supply and demand, and this might imply differences in the spatial distribution of benefits, compared with our analysis that implicitly presumes that all of the benefits are enjoyed within the innovating state (i.e., accruing to producers or assuming an absence of interstate and international trade). This distortion will itself be unevenly distributed. Some states produce commodities for which the United States as a whole does not appreciably influence the world price, let alone an individual state; but California, for instance, significantly influences world prices for a substantial share of its production, and a sizable share of its production is consumed in other states. This means that there are greater spillovers of California's research benefits than for most other states, driven by price changes, which we have not accounted for here.

[19] Here SAES research is used as a shorthand for the funds (from all sources) spent on research undertaken by the SAESs and selected other cooperating institutions in the same state. Thus, for each state, we included research spending by the SAESs (including the 1890 colleges) and the state-specific veterinary medicine schools. We excluded from our intramural USDA series research spending by the state-specific forestry schools (for symmetry with the coverage of our agricultural productivity series) and likewise omitted forestry-related research spending. The cooperating state institutions report their expenditures to USDA's Current Research Information System (CRIS) on a voluntary basis, and so the readily obtainable data are neither complete nor reported in a consistent fashion (from year to year and among states). We did a considerable amount of work to clean up erroneous and sometimes large reporting problems with the CRIS data for these state-specific cooperating institutions. The extension series is an estimate of total funding (from all sources) for state cooperative extension work obtained from a variety of published and unpublished sources. See Alston et al. (2010, pp. 229–236) for more details.

[20] Because the base model treats state-specific research and extension symmetrically, the own-state benefit-cost ratio for SAES research in any state is the same as the own-state benefit-cost ratio for extension; the same is true for the national benefit-cost ratios for state government expenditures on research and extension.

Table 4. Marginal Benefit-Cost Ratios for Research and Extension

State or Region	Benefit-Cost Ratios: State Research and Extension: Own-State (1)	National (2)	USDA Intramural Research (3)
		Ratio	
Total			17.5
48 States			
Average	21.0	32.1	0.4
Minimum	2.4	9.9	0.0
Maximum	57.8	69.2	1.6
Selected States			
California	33.3	43.4	1.4
Minnesota	40.6	55.4	0.8
Wyoming	12.7	23.6	0.1
Regions			
Pacific	21.8	32.9	0.6
Mountain	20.0	31.6	0.1
N Plains	42.4	54.5	0.5
S Plains	20.2	31.0	0.5
Central	33.7	46.8	0.8
Southeast	15.1	26.7	0.3
Northeast	9.4	18.4	0.1

Notes: Based on model ranked 1 in table 2.

will generate the greatest national payoff, if it attaches no value to interstate spillouts of its research results. In California, for instance, the marginal own-state payoff is $33.3 per dollar and the marginal national payoff is $43.4 per dollar; for Minnesota, the corresponding figures are $40.6 and $55.4 per dollar; for Wyoming, $12.7 and $23.6 per dollar.[21]

The own-state and national benefit-cost ratios vary considerably among states. The own-state benefit-cost ratio for state research and extension ranges from 2.4:1 to 57.8:1, around an average of 21.0:1. Similarly, the national benefit-cost ratios range from 9.9:1 to 69.2:1, around an average of 32.1:1. The spillover benefits are relatively constant across the states, and thus variation in the own-state benefits drives most of the interstate differences in national benefits from SAES research and extension. Hence, spillovers typically represent a smaller share of the total benefits in those states where own-state benefits are comparatively large. Spillover benefits to other states are worth $6–$16 per dollar spent on research; and in some states—especially states having small agricultural sectors—the spillover benefits account for the majority of the national benefits. USDA intramural research yielded a national benefit-cost ratio of 17.5:1, generally lower than the national benefit-cost ratio for research and extension conducted by states.

Effects of Alternative Specification Choices

We tried a range of variations in the model specification, including (*a*) different functional forms for the model (linear rather than logarithmic, and estimated in first-differences or growth rates rather than in levels), (*b*) differential treatment of the lag structure for extension compared with research (including a 4-year geometric lag for extension rather than the 50-year gamma lag as used for research, with or without allowing interstate spillovers of extension effects), (*c*) a different lag distribution shape (a 35-year trapezoidal lag model, as used by Huffman and Evenson [1993] rather than a 50-year gamma lag model), (*d*) different specifications of the spillover relationship (including models with no spillovers or spillovers based on proximity according to USDA regions rather than our model based on the similarity of commodity composition), and (*e*) alternative restrictions on the maximum lag length for research. Combining all of these variations implied a large number of alternative specifications to be estimated and compared. Based on an evaluation of the statistical performance of the models and other implications, we generally favor the base model, in logarithms, over all the alternatives. The specification choice that had the most profound implications for the estimates was the choice of a linear versus logarithmic functional form, and our statistical tests clearly favored the logarithmic specification, which has been the standard choice in published work.

In table 5 we present the results from a selection of alternative models, summarized in terms of the own-state and national benefits from research and extension conducted by individual states, as well as the national benefit from USDA intramural research. In each case, as appropriate, we present the results using the best-fitting gamma distribution for the particular model specification. Across all the models summarized in table 5, some consistent

[21] Rounding to the nearest whole dollar arguably conveys the appropriate degree of precision of these estimates, but we report them here and elsewhere in the text to one decimal place to facilitate cross-referencing with the relevant table.

Table 5. Effects of Specification Choices on Marginal Benefit-Cost Ratios

Model	Own-State Benefit-Cost Ratio for SAES Research			National Benefit-Cost Ratio for SAES Research			National Benefit-Cost Ratio for USDA Research
	Min.	Max.	Average	Min.	Max.	Average	Total
				Ratio			
Functional Form							
Logarithmic	2.4	57.8	21.0	9.9	69.2	32.1	17.5
Growth (first-difference logarithmic)	1.4	29.4	10.7	15.9	52.1	32.2	33.7
Linear	0.2	43.9	10.0	14.2	74.4	39.7	47.0
First-difference linear	0.2	51.3	11.2	16.7	84.5	46.3	55.4
Extension Treatment							
50 gamma lags with spillovers	2.4	57.8	21.0	9.9	69.2	32.1	17.5
50 gamma lags without spillovers	2.3	55.9	20.3	14.1	73.6	37.7	27.3
4 geom. lags with spillovers	1.3	27.8	9.3	16.2	50.5	31.4	34.7
4 geom. lags without spillovers	1.3	28.5	9.5	23.9	64.5	42.7	52.0
Lag Distribution for Research (R) and Extension (E)							
50-year gamma, R&E	2.4	57.8	21.0	9.9	69.2	32.1	17.5
35-year trap., R&E	3.4	53.5	19.8	18.0	75.4	41.2	33.6
50-year gamma, R; 4-year geom., E	1.3	27.8	9.3	16.2	50.5	31.4	34.7
35-year trap., R; 4-year geom., E	2.2	34.0	11.8	20.2	61.0	38.2	41.3
Spillovers							
Based on output mix	2.4	57.8	21.0	9.9	69.2	32.1	17.5
Based on USDA regions	2.3	48.5	17.6	6.6	62.4	24.8	60.5
No spillovers	4.5	90.0	33.7	4.5	90.0	33.7	n/a
Research Lag Length							
50-year gamma, R&E	2.4	57.8	21.0	9.9	69.2	32.1	17.5
35-year gamma, R&E	2.4	56.7	20.4	11.7	71.0	20.4	21.9
20-year gamma, R&E	2.5	39.5	14.8	17.2	63.0	36.3	33.7
50-year gamma, R; 4-year geom., E	1.3	27.8	9.3	16.2	50.5	31.4	34.7
35-year gamma, R; 4-year geom., E	1.1	27.0	8.8	15.3	48.9	30.1	33.4
20-year gamma, R; 4-year geom., E	1.7	29.4	10.4	17.6	55.0	33.7	36.6

Note: trap. = trapezoid; geom. = geometric.

patterns emerge in the estimates. First, the different models imply a range of estimates for the social benefit-cost ratio for USDA intramural research, but in every model the ratio is much greater than 1.0. Second, in every case the national benefit-cost ratios for SAES research are much greater than 1.0 in every state, and the average values across the 48 states are quite large. Third, the national benefit-cost ratios for SAES research are large relative to the own-state benefit-cost ratios. In all but two of the models (the linear model in levels or in first-difference form), the marginal own-state benefit-cost ratio was greater than 1.0 in every state, and the average value across the 48 states was much greater than 1.0 in every model. The different models do imply very different ranges of estimates of benefit-cost ratios for SAES research among the states, but the overall range is much smaller when we leave out the clearly misspecified models that were included for illustrative purposes.

The implication of these results is that specification choices do influence the results, but not in ways that change the primary messages. The marginal social benefits from agricultural research and extension are generally very large relative to the costs, though the benefit-cost ratios vary among states systematically depending on the characteristics of the states; and the spillover benefits are an important component of the total benefits, such that the national benefits are much greater than the own-state benefits from SAES research and extension, with the implication that individual states can be expected to underinvest in these activities from a national perspective. If accurate, these high own-state and even higher national benefit-cost ratios represent evidence of past underinvestment by both the state and federal governments in public agricultural research.

Credibility of Results

Our measures of benefit-cost ratios are large, and it is natural to be skeptical. One way to address that skepticism is to set aside the complex models and simply compute the value of the growth in agricultural productivity and compare it with the cost of agricultural research. In this section we present simple measures of this nature, which abstract from the issues of spatial spillovers and R&D lags that were central to our econometric analysis.

The Value of Productivity Growth

Over the period 1949–2002, our index of MFP more than doubled, from 100 in 1949 to about 257 in 2002, and if aggregate input had been held constant at the 1949 quantities, output would have increased by a factor of 2.6:1. Of the actual output in 2002, only 39% (i.e., $100/257 = 0.39$) could be accounted for by conventional inputs using 1949 technology, holding productivity constant. The remaining 61% is accounted for by economies of scale along with improvements in infrastructure and inputs and other technological changes. Hence, of the total production value, worth \$173.3 billion in 2002, only 39%, or \$67.3 billion, could be accounted for by conventional inputs using 1949 technology, and the remaining \$106.0 billion is attributable to the factors that gave rise to improved productivity. Among these factors is new technology, developed and adopted as a result of agricultural research and extension.

The actual value of agricultural output (AV_t) can be divided into two parts: (*a*) one representing what the value of output would have been, given the actual input quantities, if productivity had not grown since 1949—i.e., the hypothetical value, $HV_t = AV_t \times (100/MFP_t)$; and (*b*) another, a residual, representing the value of additional output attributable to productivity growth—i.e., residual value, $RV_t = AV_t - HV_t = AV_t \times (MFP_t - 100)/MFP_t$. As productivity increases over time, the share of the value of production attributable to productivity growth increases. Among the 48 states, the share of the total value of agricultural output in 2002 attributable to growth in productivity since 1949 averaged 58% but ranged from as low as 36% (Wyoming) to as high as 79% (Mississippi).

To summarize the stream of values of agricultural output attributable to productivity improvements, the yearly residual values, RV_t (defined above), were expressed in constant (2000) dollars. The deflated values were compounded at a real interest rate of 3% per annum and evaluated in the year 2002. The resulting stream of values of agricultural output attributable to productivity improvements is equivalent to a onetime payment of more than \$7.4 trillion in 2002, an enormous benefit from improved agricultural productivity in the United States during the post-WWII period.

Approximate Benefit-Cost Ratios

We compared the value of productivity gains since 1949 compounded forward over 54 years to 2002 with the expenditures on agricultural research and extension during 1929–1982 compounded forward to 2002. Both costs and benefits were converted into real terms using the GDP price deflator and accumulated forward to 2002 using a real discount rate of 3% per annum.

The simple ratios of approximate benefits in 2002 to approximate costs in 2002 are biased estimates of the true benefit-cost ratios for several reasons. First, the existence of long R&D lags means that we have left out some of the relevant costs (research expenditures prior to 1929 will have contributed to productivity growth between 1949 and 2002) and some of the relevant benefits (research expenditures between 1949 and 1982 will generate benefits for many years after 2002). Depending on the pattern of benefits and costs over time and the effects of discounting, these two sources of bias could be offsetting. However, given the generally rising pattern of research expenditures and the annual flows of benefits from productivity gains, we would expect the effect of the understatement of benefits to outweigh the effect of the understatement of costs, biasing the benefit-cost ratios down on balance. Second, a significant share, perhaps as much as half of the total benefits, may be attributable to private and rest-of-world research. Third, spillover effects mean that some of a state's productivity growth will be attributable to expenditures by other states and the federal government; conversely, some of the national benefits from a state's research expenditures will accrue as productivity gains in other states. In estimates at the regional level, the distortions associated with omitting state-to-state spillovers will be much smaller, and in estimates at the national level they will be absent.

In table 6 we compare estimates of approximate *average* benefit-cost ratios with the preferred estimates of *marginal* social benefit-cost ratios derived from the econometric estimation. Columns (1) and (2) show the estimates of marginal private and social benefit-cost ratios from the econometric model (as in table 4), while columns (3) and (4) show the approximate measures, comparing benefits over 1949–2002 with costs over 1929–1982, and allowing for either 100% or only 50% of the total benefits to be attributed to the public agricultural research and extension expenditures included in the measure of cost. The approximate measure of the national benefit-cost ratio could be as high as 25.6 (the upper bound with 100% attribution, in column (3)) or as low as 12.8 (our lower bound with 50% attribution, in column (4)). In column (3), the corresponding upper-bound estimates of regional benefit-cost ratios range from 18.1 to 63.6; the state-specific benefit-cost ratios range from 5.4 to 77.7, and the simple average of these 48 estimates is 30.4. In column (4), the corresponding lower-bound estimates of regional benefit-cost ratios range from 9.0 to 31.8; the state-specific benefit-cost ratios range from 2.7 to 38.8, and the simple average of these 48 estimates is 15.2. The estimates of marginal social benefit-cost ratios in column (2) are remarkably similar to the estimates of approximate average benefit-cost ratios with 100% attribution in column

Table 6. Benefit-Cost Ratios: Approximations versus Econometric Estimates

	Econometric Model, Marginal Benefit-Cost Ratio		Approximate Average Benefit-Cost Ratio Costs 1929–1982, Benefits 1949–2002	
State or Region	State R&E (own-state) (1)	State R&E (national) (2)	100% Attribution (3)	50% Attribution (4)
		Ratio		
48 States				
Average	21.0	32.1	30.4	15.2
Minimum	2.4	9.9	5.4	2.7
Maximum	57.8	69.2	77.7	38.8
Selected States				
California	33.3	43.4	48.5	24.2
Minnesota	40.6	55.4	55.6	27.8
Wyoming	12.7	23.6	17.0	8.5
Regions				
Pacific	21.8	32.9	41.1	20.5
Mountain	20.0	31.6	30.5	15.3
N Plains	42.4	54.5	63.6	31.8
S Plains	20.2	31.0	27.3	13.6
Central	33.7	46.8	40.6	20.3
Southeast	15.1	26.7	28.6	14.3
Northeast	9.4	18.4	18.1	9.0
United States (includes USDA intramural)			**25.6**	**12.8**
USDA intramural		**17.5**		

Notes: SAES = State Agricultural Experiment Station; R&E = research and extension.

Table 7. Terminal Values Implied by Various Rates of Return

Number of Years from Initial Investment	Rate of Return, % per Annum 10	20	50
	Terminal Value, Dollars of Benefit per Dollar of Initial Investment		
20	7	38	3,325
35	28	591	1,456,110
40	45	1,470	11,057,332
50	117	9,100	637,621,500

Note: Sources developed by the authors.

(3), while the estimates of state-specific (private) marginal benefit-cost ratios in column (1) are more comparable to the approximate average benefit-cost ratios with 50% attribution in column (4).

Recalibrating Rates of Return to Agricultural R&D

We prefer to use benefit-cost ratios, but the preponderance of precedent literature reports internal rates of returns. Having conducted a meta-analysis of 292 studies that reported estimates of returns to agricultural R&D, Alston et al. (2000, p. 55, table 12) reported an overall mean internal rate of return for their sample of 1,852 estimates of 81.3% per annum, with a mode of 40% and a median of 44.3%. After dropping some outliers and incomplete observations, they conducted regression analysis using a sample of 1,128 estimates with a mean of 64.6%, a mode of 28%, and a median of 42.0%. The main mass of the distribution of internal rates of return reported in the literature is between 20% and 80% per annum.[22] Other reviews of the literature may not have covered the same studies or done so in the same ways but nevertheless reached similar general conclusions—for instance, Evenson (2002) and Fuglie and Heisey (2007). In a recent report distilling the evidence for the Council for Agricultural Science and Technology, Huffman, Norton and Tweeten (2011, p. 6) reiterated the typical finding but reported a point estimate as follows:

> Numerous in-depth studies at the University of Chicago, Yale University, Iowa State University, the University of Minnesota, and elsewhere have carefully calculated the rate of return to investing in public agricultural research. Focusing on the contribution of productivity-oriented agricultural research undertaken by the main U.S. public agricultural research institutions—SAESs, [veterinary medical centers], [Agricultural Research Service], and [Economic Research Service]—to agricultural productivity in the 48 contiguous states, including spillover effects to other states in the same geoclimatic region, during 1970–2004, the marginal real rate of return is approximately 50% (Huffman 2010; Huffman and Evenson 2006a,b).

It is easy to show that a 50% rate of return is implausible for a long-term investment yielding benefits that compound over 35 years (as in Huffman 2010 and Huffman and Evenson 2006a, 2006b) let alone over 50 years, which we have found is more appropriate for U.S. public agricultural R&D. Table 7 includes some sample calculations of the terminal values of investments of one dollar over various time periods using alternative real rates of return to illustrate this point. One dollar invested at 50% per annum would be worth more than $3,000 at the end of 20 years, nearly $1.5 million at the end of 35 years, and a whopping $637 million at the end of 50 years. To provide some perspective, if the roughly $4 billion invested in public agricultural R&D in 2005 earned a return of 50% per annum compounding over 35 years, by 2040 the accumulated benefits would be worth $5,824,000 billion (2000 prices)—more than 100 times the projected U.S. GDP in 2040

[22] When characterizing the evidence from the literature, economists often use a range like this, but more often it is a narrower one with a smaller mean (such as the 20%–60% range reported by Fuglie and Heisey [2007]). As discussed by Alston et al. (2000), such selective reporting of the literature may be misleading, giving a false impression of both the average and the size of the range around it.

Table 8. Conventional and Modified Internal Rates of Return

	Conventional Internal Rate of Return (IRR)		Modified Internal Rate of Return (MIRR)	
	State R&E (own-state) (1)	State R&E (national) (2)	State R&E (own-state) (3)	State R&E (national) (4)
	Percent per Year			
48 States				
Average	18.9	22.7	8.8	9.9
Minimum	7.4	15.3	4.8	7.7
Maximum	27.6	29.1	11.4	11.7
Selected States				
California	24.1	26.1	10.2	10.7
Minnesota	24.7	27.3	10.6	11.3
Wyoming	16.8	20.9	8.2	9.5
Regions				
Pacific	20.2	23.5	9.1	10.1
Mountain	19.0	22.7	8.9	10.0
N Plains	24.9	27.0	10.7	11.2
S Plains	19.5	22.7	9.0	10.0
Central	23.1	25.9	10.1	10.8
Southeast	17.6	22.0	8.3	9.7
Northeast	14.0	19.0	7.2	8.8

Notes: R&E = research and extension. The figures in columns (3) and (4) are modified internal rates of return assuming a 3% reinvestment rate.

and more than 10 times the projected global GDP in 2040.[23] Clearly, as these figures illustrate, a 50% rate of return compounding over a long period of time is implausible. Perhaps this fact may help account for why the very large estimates have been discounted and ignored by some policymakers. Even a 10% real rate of return yields a large terminal value when compounded over 35 or 50 years.

We computed conventional internal rates of return using the same streams of benefits and costs simulated by the base-model that we used to compute the benefit-cost ratios reported in table 4. The internal rate of return is by definition the discount rate that makes the present value of the benefits equal to the present value of the costs. Summary results are reported in table 8, in columns (1) and (2), while more detailed results for all the states are presented in table 9. Relative to the mainstream of the literature, our preferred logarithmic model yielded estimates at the lower end of the range for both social and private annual rates of return to state and federal agricultural R&D—around 20%. Specifically, our estimates of own-state rates of return ranged from 7.4% to 27.6%, with an average of 18.9% per annum across the states, and the estimates of national rates of return ranged from 15.3% to 29.1%, with an average of 22.7% per annum across the states.

Our estimates of conventional internal rates of return in columns (1) and (2) of table 8 are much smaller than those reported typically—for instance, the 50% annual rate of return reported by Huffman, Norton and Tweeten (2011). Even so, we think our own measures are unrealistically high—for a conceptual reason as well as because of their unrealistic implications as indicated in table 7. Specifically, the conventional internal rate of return implicitly assumes that the flows of benefits can be reinvested at the same rate as the investment being evaluated. It is suited for a situation where those entities that would pay the cost would also reap the returns, whereas in the present context the government pays the cost but the benefits accrue to producers and consumers of farm products. In our application, if a public research investment is to earn a rate of return of 50% per annum, the conventional calculation will be correct only if the farmers and consumers to whom the streams of benefits accrue can (and do) invest their net benefits at the same 50% rate of return.

[23] Fogel (2007) forecasted that U.S. GDP would reach $41,944 billion in 2040 and that global GDP would reach $307,857 billion (2000 purchasing power parity prices).

Table 9. Marginal Benefit-Cost Ratios and Internal Rates of Return

State or Region	Benefit-Cost Ratio: State R&E (own-state) (1)	Benefit-Cost Ratio: State R&E (national) (2)	Benefit-Cost Ratio: USDA intramural (3)	Internal Rate of Return: State R&E (own-state) (4)	Internal Rate of Return: State R&E (national) (5)	MIRR: State R&E (own-state) (6)	MIRR: State R&E (national) (7)
	Ratio			*Percent per Year*			
Pacific	**21.8**	**32.9**	**0.6**	**20.2**	**23.5**	**9.1**	**10.1**
California	33.3	43.4	1.4	24.1	26.1	10.2	10.7
Oregon	11.3	24.1	0.2	16.3	21.3	7.9	9.5
Washington	20.9	31.2	0.3	20.3	23.1	9.2	10.0
Mountain	**20.0**	**31.6**	**0.1**	**19.0**	**22.7**	**8.9**	**10.0**
Arizona	26.6	36.9	0.2	22.1	24.5	9.7	10.4
Colorado	31.1	43.8	0.3	22.5	25.2	10.0	10.8
Idaho	34.0	44.8	0.2	23.3	25.5	10.2	10.8
Montana	22.0	32.2	0.2	20.4	23.1	9.3	10.1
Nevada	7.3	19.2	0.0	13.0	19.1	7.0	9.0
New Mexico	15.6	28.2	0.1	18.1	22.2	8.6	9.8
Utah	11.0	24.5	0.1	16.0	21.3	7.9	9.5
Wyoming	12.7	23.6	0.1	16.8	20.9	8.2	9.5
N Plains	**42.4**	**54.5**	**0.5**	**24.9**	**27.0**	**10.7**	**11.2**
Kansas	33.6	45.3	0.7	23.3	25.6	10.2	10.8
Nebraska	51.3	64.9	0.8	26.3	28.4	11.1	11.6
North Dakota	37.3	46.0	0.3	23.8	25.5	10.4	10.9
South Dakota	47.4	61.7	0.4	26.3	28.4	10.9	11.5
S Plains	**20.2**	**31.0**	**0.5**	**19.5**	**22.7**	**9.0**	**10.0**
Arkansas	26.8	35.7	0.4	21.4	23.6	9.7	10.3
Louisiana	12.2	23.0	0.2	16.8	21.0	8.1	9.4
Mississippi	15.1	25.3	0.3	18.3	21.7	8.5	9.6
Oklahoma	19.0	31.4	0.3	19.1	22.8	9.0	10.1
Texas	28.2	39.4	1.1	21.9	24.5	9.8	10.5
Central	**33.7**	**46.8**	**0.8**	**23.1**	**25.9**	**10.1**	**10.8**
Illinois	43.0	53.8	1.3	25.1	27.0	10.7	11.2
Indiana	27.1	39.4	0.7	21.7	24.6	9.7	10.5
Iowa	57.8	69.2	1.6	27.6	29.1	11.4	11.7
Michigan	17.1	31.5	0.3	19.1	23.4	8.8	10.1
Minnesota	40.6	55.4	0.8	24.7	27.3	10.6	11.3
Missouri	34.7	49.9	0.6	24.3	27.1	10.3	11.0
Ohio	22.4	37.0	0.5	20.2	24.0	9.3	10.4
Wisconsin	26.7	38.3	0.6	22.1	24.8	9.7	10.5
Southeast	**15.1**	**26.7**	**0.3**	**17.6**	**22.0**	**8.3**	**9.7**
Alabama	13.4	24.8	0.3	17.1	21.3	8.3	9.6
Florida	21.6	28.2	0.4	20.5	22.4	9.3	9.8
Georgia	20.5	31.0	0.4	20.3	23.2	9.2	10.0
Kentucky	18.5	30.5	0.3	19.2	22.8	8.9	10.0
North Carolina	19.9	27.5	0.5	20.6	22.8	9.1	9.8
South Carolina	11.2	23.1	0.1	16.1	20.9	7.9	9.4
Tennessee	15.7	31.3	0.2	18.4	23.3	8.6	10.1
Virginia	11.8	26.3	0.2	16.7	22.0	8.0	9.7
West Virginia	3.8	17.6	0.0	9.7	18.9	5.7	8.8
Northeast	**9.4**	**18.4**	**0.1**	**14.0**	**19.0**	**7.2**	**8.8**
Connecticut	5.4	14.2	0.0	11.8	17.6	6.4	8.4
Delaware	15.8	21.5	0.0	17.9	20.0	8.6	9.3
Maine	13.5	20.1	0.1	17.5	20.1	8.3	9.1
Maryland	14.1	26.1	0.1	17.5	21.7	8.4	9.7
Massachusetts	4.7	13.3	0.0	10.8	17.0	6.1	8.2
New Hampshire	4.4	14.0	0.0	10.6	17.5	6.0	8.4
New Jersey	4.7	13.7	0.1	11.3	17.5	6.1	8.3

Continued

Table 9. Continued

State or Region	Benefit-Cost Ratio: State R&E (own-state) (1)	Benefit-Cost Ratio: State R&E (national) (2)	Benefit-Cost Ratio: USDA intramural (3)	Internal Rate of Return: State R&E (own-state) (4)	Internal Rate of Return: State R&E (national) (5)	MIRR: State R&E (own-state) (6)	MIRR: State R&E (national) (7)
New York	8.3	18.1	0.3	14.2	19.1	7.3	8.9
Pennsylvania	18.0	30.3	0.3	19.1	22.8	8.9	10.0
Rhode Island	2.4	9.9	0.0	7.4	15.3	4.8	7.7
Vermont	12.4	21.5	0.0	16.2	20.0	8.1	9.3
U.S. average[a]	**21.0**	**32.1**	**0.4**	**18.9**	**22.7**	**8.8**	**9.9**

Notes: R&E = research and extension; MIRR = modified internal rate of return. The figures in columns (6) and (7) are MIRRs assuming a 3% per annum reinvestment rate. [a]Average of 48 contiguous U.S. states.

Kierulff (2008) provides a recent discussion of conventional measures of internal rates of return, their shortcomings, and the reason a so-called modified version is preferred for financial analyses applied to investments that yield streams of revenue.[24]

Consider an investment of I_t dollars in time t that will yield a flow of benefits, B_{t+n}, over the following N years. The conventional internal rate of return, i, solves the equation

$$(12) \qquad \sum_{n=0}^{N} B_{t+n}(1+i)^{N-n} - I_t(1+i)^N = 0.$$

Alternatively, suppose the stream of benefits would be reinvested by the beneficiaries (say, farmers or food consumers) at some external rate of return, r, which could be different from the rate for the project being evaluated. Then we would want to solve for the modified internal rate of return, m, which solves the problem

$$(13) \qquad \sum_{n=0}^{N} B_{t+n}(1+r)^{N-n} - I_t(1+m)^N = 0.$$

Intuitively, m is the rate at which one could afford to borrow the amount to be invested, I_t, given that it would generate the flow of benefits, B_{t+n}, that would be reinvested at the external rate, r. It can be seen that the conventional calculation of the internal rate of return is a special case of equation (13), which assumes $r = i (= m)$, which is implausible for public projects yielding flows of benefits that imply very large conventional internal rates of return.

We computed the modified internal rates of return corresponding to the conventional internal rates of return in tables 8 and 9, assuming that benefits could be reinvested at a real rate of 3% per annum (the same rate we used to compute the benefit-cost ratios). Summary results are reported in table 8, in columns (3) and (4), while more detailed results for all the states are presented in table 9. Our estimates of the own-state modified internal rates of return ranged from 4.8% to 11.4%, with an average of 8.8% per annum across the states, while the estimates of national rates of return (including interstate spillovers) ranged from 7.7% to 11.7%, with an average of 9.9% per annum across the states. We also computed the conventional internal rate of return for USDA intramural research, which was 18.7% per annum, and the corresponding modified internal rate of return, which was 8.7%. All of these modified rates of return are plausible yet consistent with very high benefit-cost ratios.

Conclusion

Measures of the payoff to public agricultural R&D are potentially useful for policy, and this usefulness will be greater if the measures are transparent, well understood, and credible. The overwhelming message from the extant literature on the returns to agricultural R&D is that it has paid handsome dividends and has been underfunded—yet the underfunding pattern persists. In the work reported in this paper, we set out to develop new evidence on the returns to agricultural research and extension and present it in a new light.

[24] Biondi (2006) suggests that the modified internal rate of return concept was first proposed by Duvillard in the late 19th century and was "reinvented" in the late 1950s by Solomon (1956), Hirshleifer (1958), and Baldwin (1959).

While our estimates apply to both research and extension, since they enter our model symmetrically, much of the previous literature has emphasized returns to agricultural research per se, and that is the benchmark for comparison.

The work reported here entails several contributions. The analysis is based on entirely new measures of both agricultural productivity and state and federal government investments in agricultural research and extension that were developed specifically for this purpose. The models used here are also new and different from those used previously in some ways that have implications for findings. In particular we tested for lag length in a flexible gamma lag distribution model and, compared with typically used models, our preferred model suggests a much longer lag length, which in turn has implications for measured rates of return. In addition, we used a new approach to model spatial spillovers, based on the similarity of commodity composition rather than spatial proximity, and evaluated the implications. Rather than simply impose a set of modeling assumptions, we evaluated the implications of our own modeling choices versus alternatives typically reported for findings with respect to returns to research. We found that some elements of specification choices had quite significant impacts on findings but that the main finding was consistent across models: a very high social payoff to the investment with very significant state-to-state spillover effects compounding incentive problems and justifying a significant federal role.

Nevertheless, the combination of specification choices in our preferred model resulted in a much lower conventionally measured internal rate of return to research than has been reported typically in previous studies. These comparatively low rates of return reflect our comparatively long lags and our greater attention to reducing other sources of misattribution bias that have contributed to very high rates of return found in some studies (as discussed by, e.g., Alston and Pardey [2001]), and the comparison lends credibility to our results. Moreover, we show that the conventional internal rate of return measures are implausible. Our modified internal rates of return are much lower than the very high rates that are still part of the mainstream in the literature and being presented to policymakers. These new findings regarding the prevalent use of a flawed metric provide some empirical justification for the skepticism sometimes expressed about very high estimated rates of return to research, which may have contributed in turn to skepticism about the value of the investment.

To address that skepticism, we developed simple, approximate measures of benefit-cost ratios. These measures are based on comparing the value of productivity growth with the cost of investments in agricultural research, without specifically modeling the statistical relationship between productivity and spending over space and time, thereby avoiding the problem of specification bias. They generate similar measures to those coming from the econometric analysis, illustrating the point that the econometric estimates reflect the same fundamental forces at work. Specifically, agricultural productivity growth is worth many times more than the annual spending on agricultural R&D (including extension). Even if only a fraction is attributed to R&D, and even if the lags are very long, the implied benefit-cost ratio will be very large.

Our specific empirical results are interesting, but in this work we have sought to emphasize the insights we can draw from the overall pattern and robustness of the evidence. Throughout we have emphasized two elements: (*a*) the spatial and temporal attribution problems associated with modeling R&D lags and (*b*) spatial spillovers. Our results show that R&D lags are very long, much longer than most previous studies have allowed, which has potential implications for problems for policy prescriptions, as well as econometric biases. Likewise, spatial spillovers are empirically important, contributing to important differences between state-specific and national benefits from SAES research. Studies that do not account appropriately for spillovers may suffer from econometric biases and could yield inappropriate policy prescriptions.

The finding of substantial interstate technology spillovers suggests that states would underinvest in agricultural R&D from a national perspective, even if they did not underinvest from a narrower state-specific perspective. Federal support for SAES research can be justified on these grounds. As well as providing a justification for federal support of SAES research, spatial technology spillovers provide a justification for intramural research by the USDA. Our results indicate that even with substantial support from the federal government, most states substantially underinvest in agricultural R&D, in the sense that both the in-state and national returns well exceed the costs of additional investments in agricultural R&D; they also indicate that these institutional failures continue to impose very large opportunity costs on individual states and the nation as a whole.

References

Acquaye, A. K., J. M. Alston, and P. G. Pardey. 2003. Post-War Productivity Patterns in U.S. Agriculture: Influences of Aggregation Procedures in a State-Level Analysis. *American Journal of Agricultural Economics* 85(1): 59–80.

Alston, J. M. Spillovers. 2002. *Australian Journal of Agricultural and Resource Economics* 46(3): 315–346.

Alston, J. M., M. A. Andersen, J. S. James, and P. G. Pardey. 2010. *Persistence Pays: U.S. Agricultural Productivity Growth and the Benefits from Public R&D Spending*. New York: Springer.

Alston, J. M., C. Chan-Kang, M. C. Marra, P. G. Pardey, and T. Wyatt. 2000. *A Meta-Analysis of Rates of Return to Agricultural R&D: Ex Pede Herculem?* Research Report No. 113, Washington DC: International Food Policy Research Institute.

Alston, J. M., G. W. Norton, and P. G. Pardey. 1995. *Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca, NY: Cornell University Press.

Alston, J. M., and P. G. Pardey. 2001. Attribution and Other Problems in Assessing the Returns to Agricultural Research. *Agricultural Economics* 25(2–3): 141–152.

Baldwin, R. H. 1959. How to Assess Investment Proposals. *Harvard Business Review* 37: 98–104.

Biondi, Y. 2006. The Double Emergence of the Modified Internal Rate of Return: The Neglected Financial Work of Duvillard (1755–1832) in a Comparative Perspective. *European Journal of the History of Economic Thought* 13(3): 311–335.

Craig, B. J., and P. G. Pardey. 1996. Productivity Measurement in the Presence of Quality Change. *American Journal of Agricultural Economics* 78(5): 1349–1354.

Evenson, R. E. 1967. The Contribution of Agricultural Research to Production. *Journal of Farm Economics* 49: 1415–1425.

Evenson, R. E. 2002. Economic Impacts of Agricultural Research and Extension. In *Handbook of Agricultural Economics, Volume 1: Agricultural Production*, ed. B. L. Gardner and G. C. Rausser. New York: Elsevier.

Fogel, R. W. 2007. *Capitalism and Democracy in 2040: Forecasts and Speculations*. Working Paper No. 3184. Washington DC: National Bureau of Economic Research.

Fuglie, K. O., and P. W. Heisey. 2007. *Economic Returns to Public Agricultural Research*. ERS Economic Brief No. 10. Washington DC: USDA.

Griliches, Z. 1964. Research Expenditures, Education, and the Aggregate Agricultural Production Function. *American Economic Review* 54(6): 961–974.

Griliches, Z. 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics* 10: 92–116.

Hausman, J. A. 1978. Specification Tests in Econometrics. *Econometrica* 46: 1251–1271.

Hirshleifer, J. 1958. On the Theory of Optimal Investment Decision. *Journal of Political Economy* 66(4): 329–352.

Huffman, W. E. 2010. *Measuring Public Agricultural Research Capital and Its Contribution to State Agricultural Productivity*. Working paper, Department of Economics. Ames: Iowa State University.

Huffman, W. E., and R. E. Evenson. 1993. *Science for Agriculture: A Long-Term Perspective*. Ames: Iowa State University Press.

Huffman, W. E., and R. E. Evenson. 2006a. Do Formula or Competitive Grant Funds Have Greater Impacts on State Agricultural Productivity? *American Journal of Agricultural Economics* 88(4): 783–798.

Huffman, W. E., and R. E. Evenson. 2006b. *Science for Agriculture: A Long-Term Perspective*, 2nd ed. Ames, Iowa: Blackwell Publishing.

Huffman, W. E., G. W. Norton, and L. G. Tweeten. 2011. *Investing in a Better Future through Public Agricultural Research*. Ames: Council for Agricultural Science and Technology (CAST). Available at http://www.cast-science.org/ http://www.cast-science.org/publications/index.cfm/investing_in_a_better_future_through_public_agricultural_research?show=product&productID=2963.

Jaffe, A. B. 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firm's Patents, Profits, and Market

Value. *American Economic Review* 76(5): 984–1001.

Jaffe, A. B. 1989. Characterizing the “Technological Position” of Firms, with Application to Quantifying Technological Opportunity and Research Spillovers. *Research Policy* 18(2): 87–97.

Kierulff, H. 2008. MIRR: A Better Measure. *Business Horizons* 51(4): 321–329.

Organisation for Economic Co-Operation and Development. 2002. *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development*. Paris: OECD.

Pardey, P. G., M. A. Andersen, B. J. Craig, and J. M. Alston. 2009. Primary Data Documentation: U.S. Agricultural Input, Output, and Productivity Series, 1949–2002 (version 4). InSTePP Data Documentation. International Science and Technology Practice and Policy Center, University of Minnesota.

Pardey, P. G., and B. Craig. 1989. Causal Relationships between Public Sector Agricultural Research Expenditures and Output. *American Journal of Agricultural Economics* 71: 9–19.

Persyn, D., and J. Westerlund. 2008. Error-Correction–Based Cointegration Tests for Panel Data. *Stata Journal* 8 (2): 232–241.

Solomon, E. 1956. The Arithmetic of Capital-Budgeting Decisions. *Journal of Business* 29(2): 124–129.

Stock, J. 1987. Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica* 55(5): 1035–1056.

Westerlund, J. 2007. Testing for Error Correction in Panel Data. *Oxford Bulletin of Economics and Statistics* 69: 709–748.