The Australian Journal of Agricultural and Resource Economics, 44:2, pp. 185–215

Research returns redux: a meta-analysis of the returns to agricultural R&D†

Julian M. Alston, Michele C. Marra, Philip G. Pardey, and T. J. Wyatt*

A total of 289 studies of returns to agricultural R&D were compiled and these provide 1821 estimates of rates of return. After removing statistical outliers and incomplete observations, across the remaining 1128 observations the estimated annual rates of return averaged 65 per cent overall — 80 per cent for research only, 80 per cent for extension only, and 47 per cent for research and extension combined. These averages reveal little meaningful information from a large body of literature, which provides rate-of-return estimates that are often not directly comparable. This study was aimed at trying to account for the differences. Several features of the methods used by research evaluators matter, in particular assumptions about lag lengths and the nature of the research-induced supply shift.

1. Introduction

Agricultural science administrators and those to whom they answer have been interested in measures of the economic benefits from agricultural R&D for a long time. McMillen's (1929, p. 141) account of the first-known attempt to evaluate US agricultural R&D illustrates some issues that have continued to plague the endeavour:

> During the last of his three notable terms as Secretary of Agriculture, 'Tama Jim' Wilson directed his bureau chiefs to compile a report that

† The authors wish to thank Connie Chan-Kang for her outstanding research assistance, and V. Kerry Smith and Jerry Carlson, anonymous reviewers, and participants at various workshops and seminars for their helpful suggestions. Partial funding support for this work was obtained from the Technical Advisory Committee of the Consultative Group on International Agricultural Research, from the Global Bureau of the US Agency for International Development, and from the University of California's Pacific Rim Project.

* The authors are employed, respectively, in the Department of Agricultural and Resource Economics (and Giannini Foundation), University of California, Davis, the Department of Agricultural and Resource Economics, North Carolina State University, the International Food Policy Research Institute, and the International Center for Research in the Semi-Arid Tropics, Sahelian Center, Niger. They are listed alphabetically, and senior authorship is not implied.

> would provide a picture of what, if any, profit could be shown to the country on the expenditures for research through the Department of Agriculture.
>
> Careful studies accompanied the compilation of the report. Numerous interests and industries were asked to estimate conservatively the value of such of the department's findings as affected their operations. Finally the expenditures were totaled in one column, the estimates of the returns in another, and the sheets placed before the venerable secretary.
>
> 'This will never do!' he protested. 'No one will swallow these figures!'
>
> The report revealed that for every single dollar that had been spent for scientific research in the Department of Agriculture, the nation was reaping an annual increase of nearly a thousand dollars in new wealth.
>
> 'Cut it down to $500,' insisted Wilson. 'That's as much as we can expect the public, or Congress to believe.'

The more recent literature has its roots in work by Schultz (1953) and Griliches (1957). Since then, hundreds of studies have reported measures of the returns to agricultural R&D. Although a great deal of effort and money has been spent on assessing the impacts of agricultural R&D, questions persist about what the resulting evidence means, its accuracy, and how it can be used.

Most agricultural economists and other agricultural scientists appear to believe that, in general, public-sector agricultural R&D has paid handsome dividends for society. In any event, that is the position most frequently stated and one rarely sees or hears a counter-view posited (exceptions include Pasour and Johnson 1982 and Kealey 1996); critics are more often concerned about distributional effects of socially profitable research. Nevertheless, even among agricultural scientists, who have a vested interest in the view that what they do for a living is good for the world, there is a range of subjective views about just how profitable the investment in agricultural R&D has been, or will be, for society as a whole. The rate-of-return evidence has no doubt played a part in defining the distribution of opinion, and refining what that evidence means can lead to a shift in general perceptions.

The past studies potentially provide a rich source of information, but limited advantage has been taken of this potential. Only partial periodic tabulations (e.g., Evenson, Waggoner and Ruttan 1979; Echeverría 1990; Alston and Pardey 1996; Fuglie *et al.* 1996) have been made. The previous reviews have typically considered a selected subset of the data, and the same core selection of studies has been common among such reviews. As a result, the conventional wisdom has been based on much less than the full amount of information that has been generated by economists on the rate of return to agricultural R&D, its variation among different types of research, and the consequences of other factors such as the evaluation methods used. These selections are only a small fraction of the 292 studies considered in the present study. They imply a much smaller range of rates of return than the full set of literature contains, a distorted perception of the evidence. For instance, Fuglie *et al.* (1996, p. 28), like many before them, concluded that 'Most studies that have estimated the aggregate social rate of return to research consistently found rates of return between 40 and 60 percent.' While these authors discussed some exceptions, the clear impression is one of an empirical consensus whereas the more complete set of literature contains a much greater range of results.

Pulling together this body of work and subjecting it to systematic, quantitative scrutiny can help us to develop a clearer sense of the distribution(s) of the rate-of-return estimates and to answer a range of more specific questions that are of direct importance to national and international decision-makers concerned with agricultural R&D. Common questions include:

1 Has the rate of return to agricultural R&D declined over time?
2 Do the rates of return to agricultural R&D differ (a) between less-developed and more-developed countries, or (b) between national agricultural research systems and international centres?
3 Does the rate of return to research vary according to its problematic focus (e.g., between crop and livestock research, among different crops, or between natural resources and commodity-related R&D)?
4 Does the rate of return vary between basic and more-applied research, or between research and extension?
5 Is systematic bias built into the estimates from particular evaluation techniques and estimation details, from other aspects of the analysis, or according to who does it (e.g., self-analysis versus external evaluation)?

Our aim has been to analyse the returns to agricultural R&D literature systematically and provide insights to these questions. A comprehensive review of the evidence is needed both to minimise the risk of the selection bias inherent in partial, qualitative summaries, and to allow a comparative assessment of the relative returns among alternatives within agricultural R&D. In addition, a comprehensive analysis of the literature can provide a basis for understanding *why* rates of return differ among studies, over time, and among research fields, and so on. Any such comprehensive analysis should be based on a methodology that seeks to ensure unbiased, clearly understood evidence. The appropriate methodology is meta-analysis (Hedges and Olkin 1985). Meta-analysis is, essentially, an analysis of analyses. The idea is to amass research findings statistically and elicit from them the 'weight of the evidence' of the past studies. The array of statistical procedures used to analyse any type of data can be applied in a meta-analysis, although usually some modifications are required for statistical inference with a meta-dataset.

Statistical research synthesis, or meta-analysis, is a relatively young methodology. Prior to its inception, an accumulation of what was known about a particular research area depended upon narrative reviews and tabular compilations of results from a selection of studies. The selection usually was made by a researcher writing a new article in the area or an expert asked to provide a review for a journal or book, with few attempts at being exhaustive or subjecting the evidence to statistical analysis. This practice is still the norm in many disciplines. In the economic disciplines, meta-analysis has been used consistently only in the area of market research to analyse consumer response to various external stimuli such as advertising (Farley and Lehmann 1986). In agricultural and resource economics, meta-analyses have been limited so far to syntheses of studies measuring the value of a natural resource (Boyle *et al.* 1994; Smith and Kaoru 1990; Smith and Osborne 1993; Smith and Huang 1995) and the effect of farm size on measures of crop yield risk (Marra and Schurle 1994). All of these studies used multiple regression techniques to meta-analyse the effect of several factors on the study outcomes. The same approach is employed here.

2. Measurement issues and problems in rate-of-return studies

Some economists suspect that some of the estimated rates of return to R&D in the literature may have been systematically biased upwards by the procedures used (e.g., Alston and Pardey 1996, chapter 6). In assessing the rate-of-return evidence, it is useful to distinguish between two types of error, systematic error or bias that we can attribute to a decision in the analysis, and unavoidable, random error that we cannot account for explicitly and that varies in unpredictable ways from one analysis to another or from one project to another.

2.1 Conceptualising bias and precision

To see this distinction more clearly, let us define the measured rate of return for a particular project or program, p, (m_p) as being equal to the true rate of return (m_p^*) plus a measurement error (v_p). That is:

$$m_p = m_p^* + v_p.$$

An ideal measure is one that has a very small error. Different estimation approaches will imply different characteristics of the distribution of errors, which we can think of in terms of bias (the expected value of v_p, which is zero for an unbiased measure) and precision (the variance of v_p, which is zero for an exact estimate of m_p^*). We would expect m_i^* for project *i* to differ from m_j^* for project *j*, according to the different characteristics of the projects. The idea is to identify and account for those systematic differences. At the same time, characteristics of the program or the evaluation study will also affect the measurement errors, v_p, and it is important to account for these effects as much as possible to get meaningful information on the determinants of the rate of return.

2.2 Mis-measured costs and benefits

A number of factors might cause an estimate to depart systematically from the true rate of return. Some problems relate to the measurement of the streams of benefits and costs in ways such that the measures match up to the concepts they are meant to represent. Some of these issues are straightforward. For instance, many studies attribute all of the growth in productivity in an industry producing a particular commodity, in a particular place, to local public-sector expenditures on agricultural R&D specific to that commodity. This approach ignores the contribution of private-sector R&D (including the cost of development work to allow the results from public- and private-sector R&D to be adopted), fails to count the costs of basic R&D that may underpin the commodity-specific applied work, does not count the costs of extension, and assumes that the gains resulted from local commodity-specific R&D rather than as a result of spillovers from the same industry in other places or other industries.

A comprehensive evaluation would take into account all of the relevant costs and all of the relevant benefits. This can be very tricky to do. For instance, it is hard to know in many cases what is the source of a particular idea that led to an innovation. Apportioning overhead costs among projects or programs is not straightforward, especially when individual scientists are engaged in multiple activities (e.g., research and teaching). Studies that evaluate entire institutions can avoid the problem of apportioning costs but run into different problems. For instance, in *ex ante* assessments different scientists may be working on different projects that are mutually exclusive (e.g., different varieties of the same crop that cannot both be adopted in the same place), and the total benefits are not simply the sum of the anticipated benefits of all the projects (actually this is a problem with the evaluation of the individual projects that is often revealed only when we consider them together). Further, an institution-level evaluation avoids the problem of selection bias, in which only the successful projects are evaluated (i.e., counting all of the benefits against a fraction of the costs).

Another set of problems arises in institutions that have multiple roles —

such as US land grant colleges, which are engaged in teaching, research, and extension, or the centres of the Consultative Group on International Agricultural Research (CGIAR) with their roles in technology creation, scientist training, germplasm preservation, and institution building. When measuring the returns to the R&D activities, we should count an appropriate part, but not all, of the total costs, and some of the costs are hard to apportion appropriately. On the other hand, if we are assessing the entire set of the institution's investments, how do we measure the benefits from institution-building programs, say? In principle, what to do is clear. In practice, the benefits and attributable costs are diffuse and difficult to measure.

2.3 Selection bias

It is likely that, within any large portfolio of research projects, there will be a wide range of rates of return, including some failures. In *ex post* evaluation, it is natural for some to focus on the successful projects or programs. This is only a problem if the rate of return to the 'winners' is misinterpreted as representing the overall rate of return. The problem of selection bias can be perceived as the converse of the problem of apportioning costs, and avoiding double-counting benefits, so that the streams of benefits and costs are appropriately matched.

However, in a meta-analysis we would like to be able to make use of the fact that some studies may have deliberately selected 'winners' for evaluation. Many of the studies are based on production function analyses of aggregate data, including many research evaluation studies that evaluated not just selected projects but all of the research over specified time periods in particular research institutions or in particular industries. If selection bias matters, we may expect to find systematically lower rates of return for these more aggregative studies.

2.4 Other sources of error

As Alston, Norton and Pardey (1995) discuss in detail, the critical determinants of the measured benefits from a particular research program can be distilled into (1) the size of the industry affected; (2) the nature of the research-induced supply shift; (3) k, the percentage research-induced reduction in costs of production when the results are adopted; and (4) the timing of the flows of benefits (i.e., the research lags). Errors are likely to be small in estimating the value of production in the industry, at least for developed countries. The analyst chooses how to model the research-induced supply shift, and if a pivotal supply shift is assumed, the size of the estimated benefits might be half the estimate obtained when a parallel supply shift is assumed. How k is determined (e.g., measured directly using econometric methods, measured indirectly using industry or experimental yields, or given by expert opinion) matters, and problems in the estimation of k are sometimes related to other aspects of the model specification. In conducting any benefit-cost analysis of research, we must compare the differences in outcomes under two alternative scenarios. In ex post analyses we compare what actually happened with a hypothetical alternative — a *counterfactual* scenario — under a different research program. Problems can arise if the relevant counterfactual, given by the cost side of the analysis, is not properly reflected in the measurement of k (Alston, Norton and Pardey 1995); in other words, the same things must be held constant in measuring the benefits and the costs. The choice of lag length, especially the inappropriate truncation of lags in econometric studies, also can have serious implications for the results (Alston, Craig and Pardey 1999).

3. A model of the determinants of estimated rates of return to agricultural R&D

The factors that might account for the variation in measured returns to agricultural R&D can be grouped into five broad categories: (1) a vector of characteristics of the rate of return *measure*, itself (**m**); (2) a vector of characteristics of the *analysts* performing the evaluation (**a**); (3) a vector of characteristics of the *research* being evaluated (**r**); (4) a vector of features of the *evaluation* (**e**); and (5) random measurement errors, u. The general hypothesised functional relationship (f) between the rate of return measure (m) and the explanatory variables is:

$$m = m^*(\mathbf{r}) + v(\mathbf{a}, \mathbf{r}, \mathbf{e}, u) = \mathrm{f}(\mathbf{a}, \mathbf{r}, \mathbf{e}) + u.$$

In other words, the measure, m, is equal to the true rate of return, m^*, plus the measurement error, v. The true measure depends only on the characteristics of the research being evaluated, while the measurement error, v, depends on the same characteristics of the research but also on various other explanatory factors, as well as the purely random component, u. In some instances a particular explanatory variable is associated only with the true part, or only with the error part, of the measure, but, in many cases, a particular explanatory variable can be expected to play multiple roles.

3.1 Characteristics of the rate of return measure

Studies vary in how they define and measure the internal rate of return, so certain characteristics of the rate of return are relevant as explanatory variables to account for variation in rates of return among studies. These include whether the rate of return was *real* or *nominal*, *marginal* or *average*, *social* or *private*. Also, we distinguish between whether the rate of return was synthesised by us or computed in the original study.[1]

3.2 Analyst characteristics

Whether the work represents a self-evaluation or not is an aspect that may tend to bias results favourably or unfavourably (as we saw in the opening quote). On the other hand, it may merely mean that the analyst is comparatively well informed. Thus, the characteristics of the analyst may provide information on possible biases, or greater precision, arising from the person or group who measures a rate of return either having an interest in certain results from the study, or having access to relatively good information about the research being evaluated. This set of effects can be captured by a dummy variable to represent the particular individual or group, but this treatment will not allow us to identify the two separate elements.

A related issue is whether the work was published or not and, if published, in what type of publication. These aspects will reflect the types of reviewer scrutiny to which the work was subjected, but the publication process may also discriminate against studies that either generate rates of return that fall outside the range of 'conventional wisdom' prevailing in the profession at the time, or may not be desirable to publish for some other reason. That is, there may be a type of selection bias involved here — the so-called 'file-drawer problem'. An objective in meta-analysis is to ensure that all studies (both published and unpublished) have an equal likelihood of being selected for the analysis.

3.3 Research characteristics

The rate of return is likely to vary systematically with changes in the characteristics of the research itself. These characteristics include (1) whether it is specific to a particular field of science (e.g., basic, applied, extension, all fields); (2) whether it relates to a particular commodity class (e.g., crops, livestock, all commodities); (3) the geographic region; (4) the type of institution that conducted the R&D (e.g., university or research institute); and (5) the scope of the research being evaluated (e.g., an entire national agricultural research system, the entire portfolio for an institute, a particular program, or a single project).

[1] Some studies do not report a rate of return but do report a benefit-cost ratio (*BC*) from which we computed an approximately equivalent internal rate of return (*IRR*) based on the formula for pricing an annuity and a discount factor (i) as: $IRR = BC \times i$ (see Alston *et al.* (2000) for derivations).

3.4 Evaluation characteristics

As discussed above, several characteristics of the evaluation have implications for the measure of the research-induced change in yield, productivity, or the supply shift; others for the size of measured benefits and costs of R&D, for a given research-induced supply shift. A primary distinction concerns whether the study involves an explicit economic surplus analysis, with a formal supply and demand model, or leaves the model implicit and uses an approximation based on a percentage research-induced supply shift multiplied by the initial value of production. Studies that use explicit surplus measures involve choices about the functional forms of supply and demand (e.g., linear or constant elasticity) and the nature of the research-induced supply shift (e.g., whether it was pivotal or parallel). Other market characteristics defined in such studies include whether the relevant commodity market is open or closed to trade, and, relatedly, whether prices are endogenous or exogenous, undistorted or subject to government programs.

A related, primary distinction is between *ex post* studies, typically concerned with the effects of actual past research, and *ex ante* studies, typically concerned with the returns to hypothetical future research investments. As well as having different types of data available, and potential application of different types of evaluation tools, there might be differences in purposes between *ex post* and *ex ante* studies. To the extent that *ex post* studies are done more often to justify past investments, while *ex ante* studies are done more often with a view to allocating resources, there may be different propensities for bias. Econometric studies of research benefits are always *ex post*. In these studies, measures of productivity (or costs or profit) are regressed against measures of past research investments, and then the results are used to deduce the effects of (flows of) expenditures on research on output (costs, or profit) which can be translated into corresponding benefit streams and subjected to benefit-cost analysis. Economic surplus analysis might be *ex ante* or *ex post*, and in some cases might draw on the results of econometric estimation.

A further set of specification choices relate to the research lag distribution, including its structure, shape, and length. These choices are often determined jointly with the size of the k shift, especially in econometric studies (the lag structure defines the pattern of the shifts over time and these are estimated jointly, econometrically; in other studies the k shift may refer to a maximum shift, which is combined with adoption percentages in the lag profile to determine the entire distribution of supply shifts over time). A key choice is whether to allow for a gestation lag between the commencement of research spending and the commencement of flows of resulting benefits.

Some studies allow for spillover effects of research. Research conducted in one place, say California, may yield results that are adopted in other states or internationally (i.e., *spillouts*), which will increase global benefits but will reduce California's benefits if California is an exporter of the affected commodity and will increase California's benefits for an imported good. Thus the theoretical effects on the rate of return of the consideration of spillouts in the analysis are ambiguous. Conversely, California agriculture benefits from *spillins* of agricultural research results from other states and internationally, as well as non-agricultural research results, and an evaluation of the local returns to California's research may be biased up if these spillins are inappropriately attributed to California's research.

A final set of choices concerns what allowance is made for the effects of market distortions on the measures of benefits and costs. One such choice is whether to assume a dollar of public expenditure on research costs society one dollar or, alternatively, following Fox (1985), to allow for the deadweight costs of taxation (δ cents per dollar of revenue raised) and charge $1 + \delta$ dollars of marginal social cost per dollar of government spending.[2] By attaching a higher cost to the stream of research investments, studies that allow for $\delta > 0$ would be expected to find lower rates of return to research, everything else being equal (so long as they are measuring the social rate of return to public investments). In addition, some studies of research benefits allow for the effects of distorted exchange rates, government commodity programs, or environmental externalities. Allowing for the deadweight losses from taxation will reduce the rate of return, other factors held constant, while the effects of allowing for commodity programs, exchange rate distortions, or other distortions, are less clear and will depend on other aspects of the analysis.

4. Overview of the literature: the dataset

We compiled a comprehensive collection and listing of the empirical literature on rates of return to agricultural R&D (including both published articles and reports, and unpublished, 'grey' literature). Alston *et al.* (2000) provide details on the 292 studies reporting estimates of the returns to research. Many of the studies provide more than one estimate, so the data base for analysis includes 1 886 observations; an average of 6.5 estimates per published study.

About one-third of the publications compiled for our study are refereed journal articles. Over 60 per cent of the publications are discussion papers,

[2] See Fullerton (1991) and Ballard and Fullerton (1992) for views on the appropriate value for δ.

Table 1 Publication patterns over time

	Number of publications		Number of observations		Observations per publication	
	Journal	Other	Journal	Other	Journal	Other
1958–69	3	3	8	23	2.7	7.7
1970–79	24	14	187	118	7.8	8.4
1980–89	38	46	264	383	6.9	8.3
1990–98	34	130	166	737	4.9	5.7
All observations	99	193	625	1261	6.3	6.5

Note: This table is based on the full sample of 292 publications reporting 1 886 observations. See notes to table 5

working papers, reports, and various other grey literature. As can be seen in table 1, the pace of publishing rate-of-return studies has picked up considerably over the years: each decade published twice as much as the previous one, a classic pattern for early-stage diffusion. The balance of publication outlets has shifted, along with the rate of publications, with what appears to be faster growth in the grey literature. Much of the early literature was published in relatively formal outlets, reflecting the fact that the first studies were breaking methodological ground or that early grey literature was eventually published or lost.

We reviewed all of the relevant papers and scored each estimate according to (1) characteristics of the *measure* of the rate of return (e.g., real versus nominal, marginal versus average, private versus social, reported in the study versus deduced by us); (2) characteristics of the *analyst*, sometimes defined by the characteristics of the author(s) of the study (e.g., authors' name(s), institutional affiliation and whether it was a self-evaluation); (3) aspects of the *research* being evaluated, including its focus (commodity orientation, natural resource focus), period during which the research was performed, nature of technology (e.g., biological, chemical, mechanical), nature of R&D (e.g., basic, applied, extension), the sector to which it applies (e.g., input supply, on-farm, post-harvest), its country/regional focus, and the institutional details of the agency doing the research being evaluated (e.g., national government, near government, international, private); and (4) characteristics of the *evaluation*, including technical estimation details (nature of lag structure, overall lag length, length of gestation lag if any, method of estimation, and treatment of price and other market distortions), as well as when and where the study was published.[3]

[3] Initially two coders scored an identical subset of studies and their results were compared. The degree of consistency between the two led us to conclude that coder bias would not be a problem, and one of the two went on to score the entire dataset.

4.1 Analyst (first-author) characteristics

Sometimes it is instructive to know who is doing the evaluating. Evaluators having certain institutional affiliations may approach the research evaluation differently or may tend to have strong prior views (i.e., pre-set biases) about the rate of return to a particular type of research. First-author employment is one measure of the general institutional bent of an evaluation. Just over one-half of the first-author evaluators were employed by universities, with about one-half of those being US land-grant universities. Government evaluators made up almost one-quarter of the first authors, and international researchers, almost 10 per cent. The rest of the first authors were either affiliated with international funding institutions or private corporations, or their affiliation was not identified. Almost 28 per cent of the evaluations were self-evaluations, while over one-half were not (the rest could not be categorised).

4.2 Research characteristics

Table 2 reports the numbers of publications and numbers of rate-of-return estimates according to the nature of the research being evaluated. In the meta-analysis, the unit of observation is the estimate and, unless otherwise noted, the categories are mutually exclusive. At the publication level, however, few categories are mutually exclusive since a single publication might estimate separate returns to, say, basic and applied research, or for research related to different commodities.

It can be seen in table 2 that the distribution of estimates of rates of return is concentrated in certain categories. Few studies evaluated basic research or extension; most computed returns to either all types of research, or research and extension.[4] The lion's share concerned yield-enhancing R&D, followed by crop and livestock management, and pest and disease management. Farming technology is the main focus, with the few studies of off-farm R&D evenly divided between pre- and post-farm technology. Research evaluations are mostly multi-institutional, although significant numbers of studies concerned a specific project, program, or organization. Government is the dominant category of research performer represented in evaluation studies; only 25 studies explicitly evaluated privately performed research. Overwhelmingly, evaluations relate to research into crops (rice, wheat, and maize research together account for almost one quarter of the data base). These patterns among the data have implications for the potential to draw precise conclusions about

[4] The distinctions between basic and applied research are not always clear. Rates of return were identified as applying to 'basic' or 'applied (or maintenance)' research only if reported as such by the authors of the evaluation studies.

Table 2 Profile of research characteristics

	Number		Share of total [a]	
	Publications	Estimates	Publications (%)	Estimates (%)
Research orientation				
Basic research	10	32	3.4	1.7
Applied research	44	194	15.1	10.3
All research	155	929	53.1	49.3
Research and extension	118	646	40.4	34.3
Extension	24	82	8.2	4.3
Unspecified	3	3	1.0	0.2
Research focus				
Yield enhancing	142	805	48.6	42.7
Crop & livestock management	95	585	32.5	31.0
Pest & disease management	76	479	26.0	25.4
Information	7	26	2.4	1.4
Post farm	15	89	5.1	4.7
Other	39	175	13.4	9.3
Unspecified	91	678	31.2	35.9
Economic sector				
Farming	179	1 054	61.3	55.9
Processing	12	34	4.1	1.8
Inputs	15	59	5.1	3.1
General agriculture	89	671	30.5	35.6
Other	15	68	5.1	3.6
Institutional orientation				
Project	57	293	19.5	15.5
Program	68	315	23.3	16.7
Agency	25	166	8.6	8.8
Multi-institutional	149	1 112	51.0	59.0
Research performer				
Government	229	1 323	78.4	70.1
University (except US land grants)	28	175	9.6	9.3
US land grants	44	438	15.1	23.2
International	27	62	9.2	3.3
Private	25	167	8.6	8.9
Other	10	40	3.4	2.1
Unspecified	29	250	9.9	13.3
Commodity focus				
Field crops[b]	165	985	56.5	52.2
Maize	37	184	12.7	9.8
Wheat	42	163	14.4	8.6
Rice	30	88	10.3	4.7
Livestock[c]	42	242	14.4	12.8
Crops & Livestock	15	84	5.1	4.5
Tree crops	21	117	7.2	6.2
Natural resources[d]	15	79	5.1	4.2
All agriculture	57	355	19.5	18.8
Unclear	8	24	2.7	1.3

Notes: This table is based on the full sample of 292 publications reporting 1 886 observations.
[a] Percentages in each section may not total 100 because categories are not always mutually exclusive. In particular, a single publication may provide multiple estimates from different categories.
[b] Includes all crops, barley, beans, cassava, groundnuts, maize, millet, other crops, pigeon pea/chickpea, potato, rice, sesame, sorghum, wheat.
[c] Includes beef, swine, poultry, sheep/goat, all livestock, dairy, other livestock, pasture, 'dairy and beef'.
[d] Includes fishery and forestry.

aspects that have not been represented very extensively in the past evaluations (such as private research or research into natural resources).

4.3 Evaluation characteristics

As discussed above, method matters. Table 3 documents some primary modelling choices. It documents the distribution of the evaluation evidence according to one set of characteristics of the evaluation, those related to model specification. A primary distinction is between rates of return derived from econometric models, especially where the lag structure has been estimated econometrically, and those derived from economic surplus models in which the lag structure was assumed and imposed, along with other aspects. These are not mutually exclusive categories since some studies have used both methods. A total of 100 studies used econometric estimates, but only eight of these simulated counterfactual research programs to generate rates of return; almost all deduced a rate of return analytically, as an algebraic transformation of estimated parameters.[5] As shown by Alston, Norton and Pardey (1995, pp. 193–206), the analytical approach is hard to get right. Among the 209 studies that used some form of economic surplus, almost half (90) used a simple approximation originally proposed by Griliches (1957), Gross Annual Research Benefits (GARB) equal to k times the value of production — an *implicit* economic surplus measure. Furthermore, most used closed-economy models or a simple small-country model. Only 17 studies allowed for an effect of research on world prices.

A key determinant of the estimate of the annual benefits from the adoption of a new technology is the measure of the research-induced shift in supply (or increase in productivity), sometimes referred to as k, as above. Table 4 shows the distribution among studies of methods for estimating this shift. Among the 130 studies using econometric methods, most used production functions or productivity functions. Among the 175 studies using non-econometric methods, about half used experimental yields, and a further quarter used industry yields. Only a handful of studies allowed for spillins and spillouts of research effects.

[5] This procedure solves the estimated model for a rate of return, rather than conducting an explicit simulation of streams of flows of benefits and costs. For instance, Alston, Craig and Pardey (1999) approximated the rate of return, r, using

$$0 = \frac{V}{MFP}\sum_{s=0}^{\infty} b_s(1+r)^{-s} - VR$$

where V is the value of output, MFP is multifactor productivity, VR is the expenditure on research, and b_s is the coefficient linking productivity today to the logarithm of research in the year s years past.

Table 3 Specifications used to evaluate benefits

	Number		Share of total	
	Publications	Estimates	Publications %	Estimates %
Modelling approach				
Econometric	99	733	33.9	38.9
Analytical	91	699	31.2	37.1
Simulated	8	34	2.7	1.8
Economic surplus	209	1147	71.6	60.8
Implicit	90	497	30.8	26.4
Explicit	119	650	40.8	34.5
Unspecified	3	12	1.0	0.6
Number of markets, explicit				
Single	113	624	38.7	33.1
Multihorizontal	6	16	2.1	0.8
Multivertical	5	21	1.7	1.1
Unclear	1	1	0.3	0.1
Trade structure, explicit surplus model				
Closed	68	386	23.3	20.5
Open				
Large	17	53	5.8	2.8
Small	52	222	17.8	11.8
Unclear	1	1	0.3	0.1

Note: This table is based on the full sample of 292 publications reporting 1 886 observations.

Table 4 Estimation of research-induced supply shifts

	Number		Share of total	
	Publications	Estimates	Publications %	Estimates %
Econometric approach	130	969	43.2	48.4
Production	55	413	18.8	21.9
Productivity	44	335	15.1	17.8
Cost	7	51	2.4	2.7
Supply	18	110	6.2	5.8
Non-parametric	2	4	0.7	0.2
Other [a]	11	60	3.8	3.2
Non-econometric approach	175	917	59.9	48.6
Experimental yields	93	460	31.8	24.4
Industry yields	47	204	16.1	10.8
Experimental productivity	5	89	1.7	4.7
Other [a]	46	203	15.8	10.8
Incremental costs included	81	487	27.7	25.8
Spillovers [b]				
Spillins	41	324	14.0	17.2
Spillouts	11	94	3.8	5.0
No spillovers	257	1486	88.0	78.8

Notes: This table is based on the full sample of 292 publications reporting 1 886 observations. The number of publications and estimates corresponding to the methods listed under econometric and non-econometric approach will not add to the total of each section because categories are not always mutually exclusive. In particular, a single publication may have used multiple methods.
[a] Supply shift calculated by other means (e.g., direct measurement) or by cost reduction.
[b] Some estimates have spillover effects both ways.

In a non-econometric analysis, excessive truncation of the lag will reduce the rate of return because some future benefits will be ignored. In an econometric study, however, the opposite can (and indeed does) happen because larger short-term benefits are estimated when a truncated lag is used. Alston *et al.* (2000) report that more than half the estimates do not even clearly specify this element. Polynomial lags are the most frequent choice in those studies that do specify the lag structure. Of the 876 estimates with an explicit research lag structure, 338 did not include any gestation lag between the time when research expenditure is incurred and the time when the resulting benefits begin to flow. Perhaps the most important difference among the studies, however, is the lag length. Among the studies that used an explicit lag structure, most used research lag lengths of less than 20 years; extension lag lengths were mostly less than 10 years.

All the study characteristics discussed above can be expected to have some influence on the measured rate of return — either by affecting the true rate of return or the measurement error. In the next section we attempt to quantify some of the more important effects.

5. Meta-analysis of returns to agricultural R&D

5.1 Data for the analysis

One feature of the evidence on rates of return is the relatively small signal-to-noise ratio. The rates of return range from small negative numbers to an extreme and implausible rate of more than 700 000 per cent per annum.[6] This range might reflect differences in typical rates of return among different sets of studies — differences among groups such as applied versus basic research, or research on natural resources versus commodities. Unfortunately, however, the range of rates of return is similarly large within each of the primary groups of studies of interest here; the large range reflects variation *within*

[6] Investing $1 at an internal rate of return of 700 000 per cent per annum would generate $7 000 after one year, $49 million after two years, $343 billion after three years, and $2 401 trillion after four years. The GDP of the world in 1997 was $29 trillion. Suppose the investment of $1.21 billion in 1980 in US public agricultural R&D had earned an internal rate of return of 50 per cent per annum, the midpoint of the conventional wisdom (e.g., Fuglie *et al.* 1996) and close to the mean for aggregate US studies in the dataset used in our regression analysis (49.4 per cent per annum). The accumulated stream of benefits would be worth $6 trillion (1980 dollars) by the year 2000, 70 years' worth of agricultural GDP. The same amount invested at 8 per cent per annum (e.g., Alston, Craig, and Pardey 1999) would be worth $6 billion (1980 dollars) in 2000 — more plausible and still a good investment.

more than *among* groups. This large within-group variation makes it more difficult to discern statistically significant differences among groups. Figure 1 provides a graphical breakdown of the rate-of-return estimates grouped into several categories — specifically *ex post* and *ex ante* studies, nominal and real rates of return, and all the rate-of-return observations versus research-only.

In order to reduce the role of the extreme observations in masking the information content of the data, we discarded outliers using the method proposed by Belsey, Kuh and Welsch (1980) in which a number of statistical tests are used to assess changes in the predicted and residual values as a consequence of deleting observations. In all, 30 observations were discarded as being outliers having undue influence on the regression parameters. Table 5 provides some summary statistics on the distributions of rates of return to research, extension, and both research and extension for both the full dataset and the meta-dataset used in the regression analysis. Both table 5 and figure 1 illustrate the generally wide spread within each category, as well as the positively skewed nature of the distributions.

Table 5 Ranges of rates of return

	Number of observations	Rate of return %				
		Mean	Mode	Median	Minimum	Maximum
Full Sample [a]						
Research only	1144	99.6	46.0	48.0	−7.4	5645
Extension only	80	84.6	47.0	62.9	0	636
Research and extension	628	47.6	28.0	37.0	−100.0	430
All observations	1852	81.3	40.0	44.3	−100.0	5645
Regression Sample [b]						
Research only	598	79.6	26.0	49.0	−7.4	910
Extension only	18	80.1	91.0	58.4	1.3	350
Research and extension	512	46.6	28.0	36.0	−100.0	430
All observations	1128	64.6	28.0	42.0	−100.0	910

Notes:
[a] The original full sample included 292 publications reporting 1 886 observations. Of these, 9 publications were dropped because rather than specific rates of return they reported results such as '>100 per cent' or '< 0'. As a result of these exclusions, 32 observations were lost. Of the remaining 1 854, two observations were dropped as extreme (and influential) outliers. These two estimates were 724 323 per cent and 455 290 per cent per annum.
[b] Excludes outliers and observations that could not be used in the regression owing to incomplete information on explanatory variables. See text discussion.

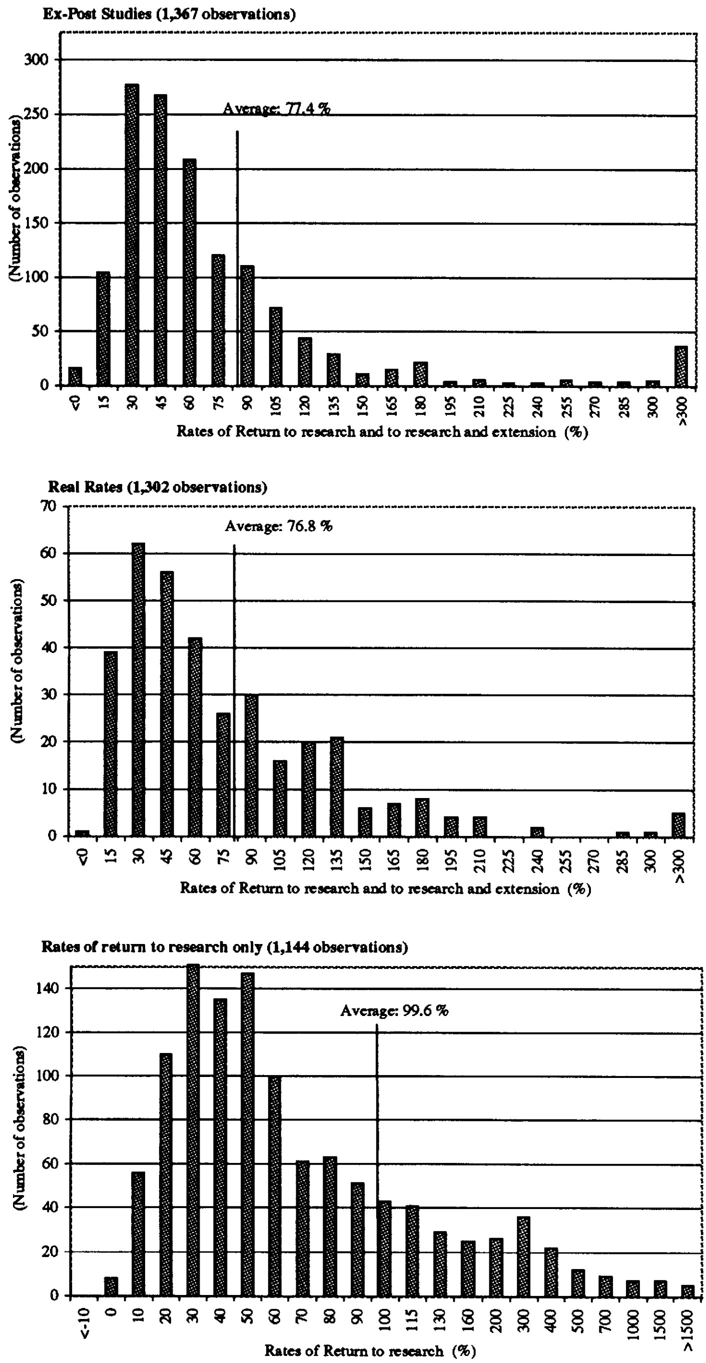
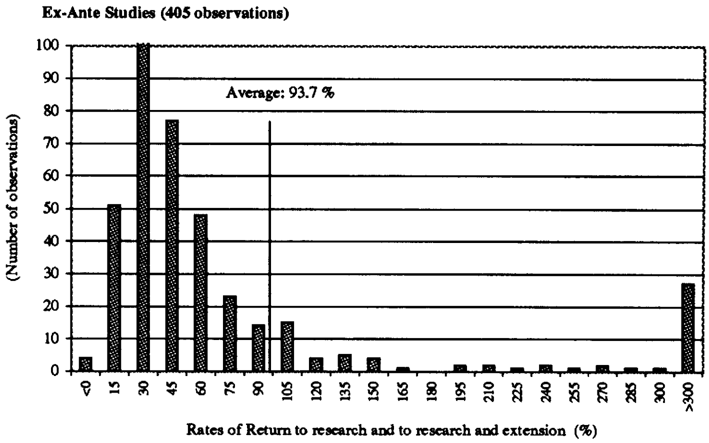
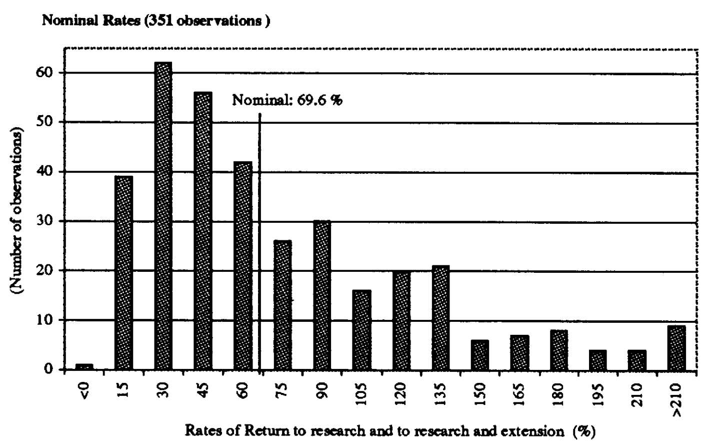
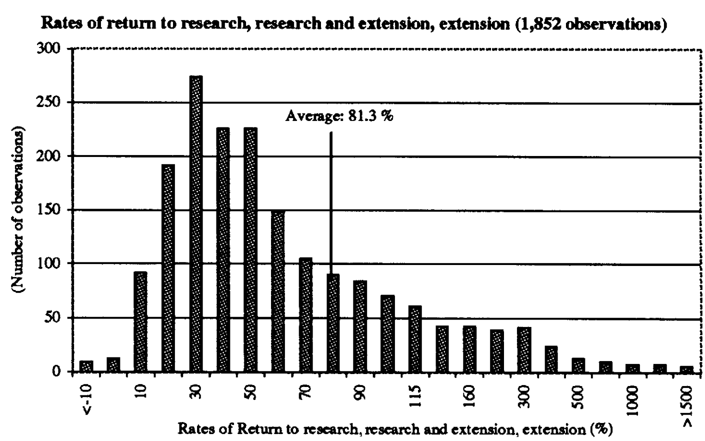


Figure 1 Distributions of rates of return to agricultural R&D

Note: [a] Excludes two extreme outliers


Ex-Ante Studies (405 observations)
(Number of observations)
Average: 93.7 %
<0 15 30 45 60 75 90 105 120 135 150 165 180 195 210 225 240 255 270 285 300 >300
Rates of Return to research and to research and extension (%)


Nominal Rates (351 observations)
(Number of observations)
Nominal: 69.6 %
<0 15 30 45 60 75 90 105 120 135 150 165 180 195 210 >210
Rates of Return to research and to research and extension (%)


Rates of return to research, research and extension, extension (1,852 observations)
(Number of observations)
Average: 81.3 %
<-10 10 30 50 70 90 115 160 300 500 1000 >1500
Rates of Return to research, research and extension, extension (%)

The overall average rate of return across all 1 128 observations used in the regression was 64.66 per cent per annum, with a standard deviation of 86.08 per cent. In this sample, the estimated annual rates of return averaged 80 per cent for research only, 80 per cent for extension only, and 47 per cent for research and extension combined. Conditional mean rates of return and standard deviations associated with each variable are shown in table 6. These conditional means provide some indication of how particular characteristics might influence the computed rate of return, but of course these univariate measures are only partial. The idea in the regression model is to take account of the different effects jointly, in a multivariate analysis.

5.2 The regression model

The regression equation is a linear model of the form:

$$m = \mathrm{b}_0 + \mathbf{Xb} + \varepsilon,$$

where b_0 is the intercept, $\mathbf{b}$ is the vector of slope coefficients, $\mathbf{X}$ is the matrix of explanatory variables included in the model, and ε is the error term. Apart from some time-trend variables, all of the explanatory variables are dichotomous dummy variables, which indicate the presence or absence of particular characteristics.

The model was estimated by ordinary least squares (OLS) regression. We considered two types of potential problems with the regression errors that might affect the OLS estimates. First, the nature of the data may give rise to heteroskedasticity. Second, we might expect to find a common variance and some covariance among certain 'clusters' of errors, such as those coming from the same study or studies using the same or similar data, but still expect these errors to be independent of and have a different variance from other errors or clusters of errors. If both types of error problems are suspected, it is difficult to tell what the overall effect might be on the estimated parameters and their standard errors. Also, there is no proven way to correct for the second problem, although *ad hoc* methods have been suggested (e.g., Hall 1984; Hall, Horowitz and Jing 1995). Since our meta-dataset is relatively large, the potential distortions might be expected to be small so we did not correct the error-covariance matrix for either problem.

5.3 Estimation results

The results of the meta-analysis proper are given in table 7. The model includes all the variables that economic theory and experience led us to believe to be most important for explaining the variation in the rate-of-return estimates, as well as some that are the subject of some debate among research evaluators and a few suggested by reviewers. A high proportion of the estimated coefficients in the model have plausible magnitudes and signs. Next, we discuss these results.

Characteristics of the rate-of-return measure

Nominal rates of return should tend to be higher than real rates of return (the difference reflecting, approximately, the general rate of inflation for the same geographic location and period of analysis). This relationship is evident only in the subset of studies for which the benefit stream includes the inflationary 1970s, in which case nominal rates of return were, on average 25 per cent (26.12 − 1.54 per cent) higher than their real counterparts, and this effect is statistically significant at the 5 per cent level. The rate of return in *ex post* analyses was higher than those in *ex ante* analyses by 18 per cent, which is consistent with our conjecture that *ex post* analyses tended to pick 'winners'. Properly measured, social rates of return to research should be greater than private rates, because social rates take into account positive spillovers. The regression indicates that the social rates of return are, indeed, higher, by about 14 percentage points, but this coefficient is not statistically significantly different from zero.

Compared with measures of rates of return to research only, the results suggest that measures of the rate of return to extension only or to both research and extension were lower (by 60 and 34 percentage points, respectively). These effects are statistically significant. The cost of extension effort is not accounted for in the research-only measures, while extension effects are difficult to exclude from the benefits stream, and this would result in an upward bias in the research-only measures compared with measures of either extension alone or research and extension combined. In contrast, the conditional means suggest that the rates of return to research only and extension only were about equal, and both were much higher than for research and extension combined.

Finally, we imputed some 65 rate-of-return estimates from reported estimates of benefit-cost ratios, and the regression results indicate that these imputed rates of return were 163 per cent per annum higher than reported rates of return, other things being equal. This might have resulted from our assumption of an infinite stream of constant benefits for the imputed rates of return, while the directly reported rate-of-return measures contain a mixture of assumptions about lag lengths and the flows of benefits over time. In addition, however, it might be because absurdly high benefit-cost ratios are not as obvious as the absurdly high rates of return they imply, so that less effort might have been spent attempting to reduce the returns in studies that reported benefit-cost ratios instead of rates of return.

Table 6 Conditional mean rates of return for the variables in the regression data set

Default category	Summary statistics: Number[a]	Mean (%)	Standard deviation (%)	Explanatory variable	Summary statistics: Number[a]	Mean (%)	Standard deviation (%)
Real	856	63.64	87.14	Nominal	279	67.83	82.74
Ex ante	225	85.93	140.34	Ex post (or unknown)	903	59.36	65.00
Average	777	64.34	94.56	Marginal	351	65.35	63.52
Private	32	80.24	157.61	Social	1096	64.20	83.16
Research only	597	79.75	107.91	Extension only	18	80.06	85.59
				Both research and extension	513	46.55	43.70
Reported	1063	56.79	61.82	Imputed from a B-C ratio	65	193.22	221.92
First author affiliation — government	219	65.75	82.47	University	692	66.89	95.14
				International research centre	78	61.32	55.18
				International funding body	33	60.24	68.49
				Private sector (or unknown)	35	67.19	42.48
Independent assessment	891	67.12	89.20	Self-evaluation	237	55.37	72.59
Government (research performer)	801	57.09	59.76	University research performer	281	89.63	133.27
				International research organization	54	55.99	50.72
				Private sector	90	49.55	41.16
				Other	202	56.84	83.38
All agriculture	134	54.09	53.77	Tree crops	58	71.62	107.77
				Field crops	622	65.83	92.16
				Livestock	161	82.51	93.83
				Natural resources	61	40.96	71.35
				Unspecified research focus	92	52.20	46.30
Not specified as basic research	1115	64.24	85.41	Specified as basic research	13	100.01	131.37
Public research	1004	66.67	90.03	Private research	13	79.77	68.11
				Both private and public research	111	44.69	32.15

Developing-country performers	526	53.91	52.99	Developed-country performers	602	74.04	106.07
Single project evaluated	222	105.15	159.38	Research program evaluated	257	42.12	37.49
				Research institution evaluated	79	67.75	43.25
				Multiple research institutions evaluated	570	58.62	54.90
Evaluation published as a book or chapter, discussion paper, report, or other	704	59.29	68.26	Evaluation published as an article in a refereed journal	424	73.56	108.95
Non-econometrically estimated supply shift	649	66.04	101.83	Econometrically estimated supply shift	479	62.78	58.38
Benefits not calculated directly from an econometric model	788	65.26	93.67	Benefits calculated directly from an econometric model	340	63.25	65.28
Benefits calculated using an explicit surplus model with a parallel supply shift	190	51.76	74.62	Using an explicit surplus model with a pivotal supply shift	264	50.82	48.88
				Using an explicit surplus model with neither a pivotal nor a parallel supply shift	13	51.62	19.81
				Using an implicit surplus model	310	86.54	126.66
Industry data for supply shift	694	59.50	57.94	Experimental data for supply shift	434	72.90	117.48
No gestation lag	911	64.92	90.48	Gestation lag > 0	217	63.56	64.57
Long lag (≥ 15 years)	648	62.95	99.16	Short lag (< 15 years)	480	67.00	64.37
Spillovers not considered	999	62.91	90.10	Spillins only	120	77.76	42.09
				Spillouts only	1	84.00	0
				Both spillins and spillouts	8	84.00	39.44
Distortions not considered	899	66.90	92.23	Farm program distortions	110	57.13	47.85
				Exchange rate distortions	53	46.17	50.19
				Deadweight losses from taxation	26	79.04	49.53
				Environmental impacts	10	129.48	142.04
				Other distortions considered	51	30.46	25.10
Overall average rate of return	1128	64.66	86.08				

Note: [a] Number of observations in the regression sample having particular attributes.

Table 7 The regression results

Default category	Explanatory variable included	Estimated coefficient	t-statistic
	Intercept Term	86.57	3.68***
	Characteristics of the rate-of-return measure		
Real	Nominal	−1.54	−0.10
All other observations	Nominal × LDC interaction	5.98	0.39
All other observations	Nominal × 1970s interaction	26.12	2.12**
Ex ante	Ex post	17.65	2.04**
Average	Marginal	7.20	0.78
Private	Social	14.32	0.91
Research only	Extension only	−57.70	−2.04**
	Both research and extension	−33.63	−5.76***
Reported	Derived from a benefit-cost ratio	162.67	12.13***
	Characteristics of the analyst		
First author affiliation, government	University	−15.05	−2.06**
	International research centre	5.09	0.42
	International funding body	2.54	0.13
	Private sector	−60.94	−3.78**
	Unknown affiliation	−48.65	−4.07***
Independent assessment	Self-evaluation	−22.00	−2.65**
	Unclear if self-evaluation or not	2.96	0.43
	Characteristics of the research		
Government research performer	University research performer	2.46	0.35
	International research organization	−2.84	−0.22
	Private sector	18.13	1.16
	Other (international funder or unknown)	8.07	0.95
All agriculture	Tree crops	18.88	1.22
	Field crops	25.10	2.50**
	Livestock	12.09	1.07
	Natural resources (forestry and fisheries)	−94.46	−6.40***
	Unspecified research focus	7.73	0.65
Not specified as basic research	Specified as basic research	−34.52	−1.33
Public research	Private research	18.97	0.69
	Both private and public research	−4.10	−0.30
Developing-country performer	Developed-country performer	13.20	1.71*
Median year of benefits [a,b]		3.24×10^{-3}	0.51
Median year of benefits squared [a,b]		2.03×10^{-7}	0.11
	Characteristics of the research evaluation		
Publication date [a]		−0.84	−1.92*
Single project evaluation	Research program evaluated	−41.33	−4.53***
	Research institution evaluated	−68.91	−4.83***
	Multiple research institutions evaluated	−53.13	−4.91***

Table 7 *Continued*

Default category	Explanatory variable included	Estimated coefficient	t-statistic
Non-journal publication	Evaluation published in a refereed journal	−15.58	−2.55**
Non-econometric study	Econometrically estimated supply shift	−18.53	−1.61
Benefits calculated directly from an econometric model	Benefits imputed using an explicit surplus measure with a pivotal supply shift	10.09	1.33
	Using an explicit surplus measure with neither a pivotal nor parallel supply shift	−54.23	−2.38**
	Using an implicit surplus measure	17.66	2.20**
Industry data for supply shift	Experimental data for supply shift	10.46	1.37
Gestation lag length (years)[a]		−4.59	−7.47***
Long lag (≥ 15 years)	Short lag (< 15 years)	−11.62	−1.49
Long lag and econometrically estimated supply shift	Short lag and econometrically estimated supply shift	38.30	3.37***
Spillovers not considered	Spillins only	2.67	0.26
	Spillouts only	21.90	0.30
	Both spillins and spillouts	−34.50	−1.22
Distortions not considered	Farm program distortions	−5.00	−0.62
	Exchange rate distortions	−15.56	−1.24
	Deadweight losses from taxation	8.92	0.55
	Environmental impacts	39.98	1.30
	Other distortions considered	−9.31	−0.78
	MODEL R^2	0.35	
	NUMBER OF OBSERVATIONS	1128	

Notes:
[a] These variables are entered in continuous not dichotomous form.
[b] Variable is median year of benefit stream minus 2000.
* Significant at the 90 per cent confidence level; ** significant at the 95 per cent confidence level; ***significant at the 99 per cent confidence level.

Characteristics of the analyst

Several aspects of the affiliation of the research evaluator had statistically significant effects on the rate of return measure. Most evaluations are done by government employees (table 6). When the evaluation was done, instead, by an analyst employed in a university or the private sector, or the employer was unknown, the rate of return was statistically significantly lower (by 15, 61, or 49 per cent per annum, respectively).

Self-evaluations — a more direct measure of any tendency to bias estimates — provide significantly lower rate of return estimates (by 22 per cent per annum). At first blush, it may seem surprising to find that self-evaluations yield rates of return that are lower than more independent studies. Perhaps self-evaluators are simply better informed, have access to better data, and are less biased as a result. As can be seen in the opening quote from Wheeler McMillen, another explanation is that self-evaluators want to be plausible and are inclined to bias their estimates down (noting that many find the typical estimates too high to be really plausible) for that reason.[7]

Characteristics of the research

The returns to research do not seem to depend on who does the research. The default category of *research performer* is government; there are no statistically significant effects and the point estimates are all small for other categories of research performer. Also, there is no measurable difference in estimated rates of return between privately and publicly performed research.

Research focus does matter. The estimated coefficients on the variables representing the *research focus* suggest that compared with all agriculture, the rates of return were 25 per cent per annum higher for research on field crops and 95 per cent per annum lower for research on natural resources. It should be noted that only 61 studies fell into this category, mostly concerning forestry and some fisheries research, and these might not be representative of the broad subject matter of natural resources research, much of which has not been the subject of evaluation studies.

There is no significant difference in rates of return related to whether studies reported basic or other categories of research, nor between research that was identified by authors as private in nature versus public in nature. Where the research was conducted may matter. The point estimates indicate that if the research took place in a developed country, the rate of return was higher by 13 per cent per annum, perhaps because of better research infrastructure or better research training, but this effect was only statistically significant at the 10 per cent level.

Some suggest that the rate of return to agricultural R&D ought to be expected to decline over time, owing to some loose notion of diminishing returns, or the view that the easy problems have already been solved — nature is increasingly niggardly. On the other hand, others have said that the new biotechnology offers potential for an unprecedented technological revolution. Both the linear and quadratic time trend terms were statistically insignificant. Hence, there is no evidence that the rate of return to agricultural R&D has declined over time (in fact, the point estimates of both coefficients were positive).

[7] More generally we might expect to find a bias towards the conventional wisdom, with 'low' estimates being biased up and 'high' estimates being biased down.

Characteristics of the research evaluation

The impact of progress in research evaluation methodology on the measured rate of return can be proxied by the publication date of the evaluation. The coefficient on publication date indicates a significant downward trend of about 1 per cent per annum per year over the post-war period, but this was only significant at the 10 per cent level.

The remaining results confirm some of our predictions concerning the implications of certain modelling assumptions. First, as anticipated, more aggregative studies generally mean lower rates of return. The coefficients are significant and negative for evaluations of entire programs of research, institution-wide research, and research by multi-institutional agencies. These results suggest that rates of return are 40 to 70 per cent per annum lower for evaluations of more aggregated research investments, relative to single-project evaluations — probably a reflection of selection bias in the less aggregative studies (i.e., evaluating only impressive projects or programs or parts thereof).

A published result may be expected to have been more heavily scrutinised and this might lead to lower rates of return. This hypothesis is supported in our regression. The rate of return measure is estimated to be 16 per cent per annum lower when the results were reported in a refereed journal than in the default category of 'grey' literature.

The next block of variables refers to the approach used to compute benefits. First, there was no statistically significant difference in the estimated rate of return between econometric and non-econometric studies, but the point estimate suggests that when the supply shift was estimated econometrically, the rate of return was lower. Assumptions about the form of the research-induced supply shift had some effect in studies using explicit or implicit surplus measures. The default category is a parallel shift. Everything else being equal, a pivotal supply shift is known to result in smaller estimates of research benefits than a parallel one, so it is surprising that this was not reflected in a lower rate of return in studies using a pivotal supply shift. However, rates of return were significantly lower, by 54 per cent per annum, in the very small number of estimates (13) that used neither parallel nor pivotal shifts. The use of an implicit surplus model (i.e., $\text{GARB} = kPQ$) rather than an explicit model to compute benefits, implied an 18 per cent per annum higher rate of return, a statistically significant difference. The use of experimental yields to measure the supply shift versus the default, industry yields, did not affect the rate of return.

Several key assumptions about the lag structure were found to have significant implications for the reported rate of return. First, a longer gestation lag meant a lower rate of return (lower by 4.6 per cent per annum for each additional year of gestation). Second, overall lag length matters. Studies that assumed short lags (≤ 15 years) for research benefits found rates of return similar to those that used longer lags, although the point estimate suggests that truncation of the lag reduces the rate of return. This effect would be expected in a non-econometric study where truncation of the lags means the omission of some benefits. However, Alston, Craig and Pardey (1999) showed that in econometric studies of returns to research, the arbitrary truncation of the lag distribution for the stream of net benefits could lead to serious upward biases in the estimated rate of return. As they predicted, *econometric* studies that used short lags found rates of return that were 38 per cent per annum higher than those that used longer lags. This statistically significant coefficient reflects the result that, because of the omitted variables problem discussed earlier, truncation of lags in the stream of net benefits from research biases the rate of return up. It is noteworthy that the regression analysis picked up both the positive and negative biases from truncation of lags.

The remaining sets of coefficients that relate to the effects of allowing for *research spillovers*, and allowing for *distortions* are all statistically insignificant, and mostly small. In many of these instances, theory does not give any clear-cut indication of the likely sign of the effect, but in three instances, the signs of coefficients were unexpected: referring to studies that took account of exchange rate distortions, the deadweight losses from taxation, or environmental impacts. In each of these instances, the anomalous sign could easily have been a result of a small-sample problem, or selection bias (in table 6, these categories included only 53, 26, and 10 observations, respectively).

6. Conclusion

This study has compiled a comprehensive meta-dataset of studies representing the entire post-war history of quantitative assessment of rates of return to agricultural research. Compared with previous, narrative reviews, this data base is much more comprehensive. The consequences for drawing conclusions from this literature are both good and bad. The range of rates of return is large, which makes it harder to discern meaningful patterns in the rates of return, and to identify those factors that account for the systematic variation in the evidence. But these are the data, and it is better to use objective and systematic methods to filter the results rather than an *ad hoc* sample selection, which may entail corresponding bias.

To make our assessment of the evidence more meaningful, we excluded 39 observations that were statistically determined to be outliers exerting significant influence on the regression results, and a further 652 observations were lost because they did not include full information on all the explanatory variables in the model. This left 1128 observations to analyse. Even so, it was difficult to confidently draw meaningful inferences from the tabulations and simple pairwise comparisons. It may be important to control for some of the systematic sources of variation in order to isolate a particular effect, especially given the importance of within-group variability.

For the most part there is a close connection between our key results from the multivariate analysis and our prior beliefs based on theory. Some issues, however, are strictly empirical, and these were a significant motivation for the study. Five questions were stated in the introduction, and we have been able to answer some of them clearly; others remain the subject of further analysis.

1 There is no evidence to support the view that the rate of return has declined over time.
2 The rate of return to research may be higher when the research is conducted in more-developed countries.
3 The rate of return to research varies according to problematic focus, in ways that make intuitive sense. In general, we would expect to see longer production cycles associated with lower rates of return, and the regression results indicate a significantly lower rate of return for natural resource management research (primarily forestry) compared with the other categories, and a higher rate of return to research into (typically annual) crops.
4 A lower rate of return is found in studies that combine research and extension, and especially studies of extension only, compared with studies evaluating research only.
5 Characteristics of the research evaluation itself, particularly the scope of the research being evaluated and choices about lags, were found to have important, systematic effects on the estimated rates of return, and most of these effects are reasonable.

In addition to these primary questions, we considered other possible systematic aspects of rate-of-return estimates that might reflect characteristics of the true rates of return, or sources of bias in the estimates. For instance, characteristics of the measures themselves, and of the analyst conducting the evaluation affected the rate of return measure in ways that were expected, and self-evaluations yielded significantly lower rates of return. On the other hand, we were unable to detect any effect of accounting for spillovers or market distortions on measured rates of return to research.

References

Alston, J.M., Chan-Kang, C., Marra, M.C., Pardey, P.G. and Wyatt, TJ. 2000 (forthcoming), *A Meta-Analysis of the Rates of Return to Agricultural R&D: Ex Pede Herculem?,* IFPRI Research Report Number 113, International Food Policy Research Institute, Washington, DC.

Alston, J.M., Craig, B.J. and Pardey, P.G. 1999, *Dynamics in the Creation and Depreciation of Knowledge, and the Returns to Research*, EPTD Discussion Paper No. 35, International Food Policy Research Institute, Washington, DC, August 1998 (revised August 1999).

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

Alston, J.M. and Pardey, P.G. 1996, *Making Science Pay: The Economics of Agricultural R&D Policy,* American Enterprise Institute Press, Washington, DC.

Ballard, C.L. and Fullerton, D. 1992, 'Distortionary taxes and the provision of public goods', *Journal of Economic Perspectives*, vol. 6, pp. 117–31.

Belsey, D.A., Kuh, E. and Welsch, R.E. 1980, *Regression Diagnostics*, John Wiley & Sons Inc., New York.

Boyle, K.J., Poe, G.L. and Bergstrom, J.C. 1994, 'What do we know about ground water values? Preliminary implications from a meta analysis of contingent-valuation studies', *American Journal of Agricultural Economics*, vol. 76, pp. 1055–61.

Cook, T.D., Cooper, H.M., Cordray, D.S., Hartmann, H., Hedges, L.V., Light, R.J., Louis, T.A. and Mosteller, F. 1992, *Meta-Analysis for Explanation: A Casebook*, Russell Sage Foundation, New York.

Cooper, H.M. and Hedges, L.V. (eds) 1994, *Handbook of Research Synthesis*, Russell Sage Foundation, New York.

Echeverría, R.G. 1990, 'Assessing the impact of agricultural research', in Echeverría, R.G. (ed.), *Methods for Diagnosing Research System Constraints and Assessing the Impact of Agricultural Research, Vol. II, Assessing the Impact of Agricultural Research*, ISNAR, The Hague.

Evenson, R.E., Waggoner, P.E. and Ruttan, V.W. 1979, 'Economic benefits from research: an example from agriculture', *Science*, vol. 205, pp. 1101–7.

Farley, J.U. and Lehmann, D.R. 1986, *Meta-Analysis in Marketing*, Lexington Books, Lexington, MA.

Feldman, K.A. 1971, 'Using the work of others: some observations on reviewing and integrating', *Sociology of Education*, vol. 44, pp. 86–102.

Fox, G.C. 1985, 'Is the United States really underinvesting in agricultural research?', *American Journal of Agricultural Economics*, vol. 67, pp. 806–12.

Fuglie, K., Ballenger, N., Day, K., Klotz, C., Ollinger, M., Reilly, J., Vasavada, U. and Yee, J. 1996, *Agricultural Research and Development: Public Investments under Alternative Markets and Institutions*, Agricultural Economic Report No. 735, USDA-ERS, Washington, DC, May.

Fullerton, D. 1991, 'Reconciling recent estimates of the marginal welfare cost of taxation', *American Economic Review*, vol. 81, pp. 302–8.

Griliches, Z. 1957, 'Hybrid corn: an exploration in the economics of technological change', *Econometrica,* vol. 25, pp. 501–22.

Hall, P. 1984, 'Resampling a coverage pattern', *Stochastic Processes and their Applications*, vol. 20, pp. 231–46.

Hall, P., Horowitz, J.L. and Jing, B.-Y. 1995, 'On blocking rules for the bootstrap with dependent data', *Biometrika*, vol. 82, pp. 561–74.

Hedges, L.V. and Olkin, I. 1985, *Statistical Methods for Meta-Analysis,* Academic Press, Inc., New York.

Kealey, T. 1996, *The Economic Laws of Scientific Research*, Macmillan Press Ltd., Houndmills.

Lakshminarayan, P.G., Gassman, P.W., Bouzaher, A. and Izaurralde, R.C. 1996, 'A metamodeling approach to evaluate agricultural policy impact on soil degradation in Western Canada', *Canadian Journal of Agricultural Economics*, vol. 44, pp. 277–94.

Marra, M.C. and Schurle, B. 1994, 'Wheat yield risk and aggregation: a meta-analysis approach', *Journal of Agricultural and Resource Economics,* vol. 19, pp. 69–77.

McMillen, W. 1929, *Too Many Farmers: The Story of What is Here and Ahead in Agriculture,* William Morrow and Company, New York.

Milchunas, D.G. and Lauenroth, W.K. 1993, 'Quantitative effects of grazing on vegetation and soils over a global range of environments', *Ecological Environments,* vol. 63, pp. 327–66.

Pasour, E.C. Jr. and Johnson, M.A. 1982, 'Bureaucratic productivity: the case of agricultural research revisited', *Public Choice*, vol. 39, pp. 301–17.

Schultz, T.W. 1953, *The Economic Organization of Agriculture*, McGraw-Hill, New York.

Smith, V.K. and Huang, J.-C. 1995, 'Can markets value air quality? A meta-analysis of hedonic property value models', *Journal of Political Economy,* vol. 103, pp. 209–27.

Smith, V.K. and Kaoru, Y. 1990, 'Signals or noise: explaining the variation in recreation benefit estimates', *American Journal of Agricultural Economics*, vol. 72, pp. 419–33.

Smith, V.K. and Osborne, L. 1993, 'Do contingent valuation estimates pass a scope test?: A preliminary meta analysis', selected paper, joint meetings of the American Economic Association and the Association of Environmental and Resource Economists, Boston, MA.