The Australian Journal of
Agricultural and
Resource Economics

AARES
AUSTRALASIAN AGRICULTURAL & RESOURCE ECONOMICS SOCIETY

Australian Journal of Agricultural and Resource Economics, 64, pp. 977–1001

Recalibrating the reported returns to agricultural R&D: what if we all heeded Griliches?*

Xudong Rao, Terrance M. Hurley and Philip G. Pardey[†]

Zvi Griliches' seminal analysis of hybrid corn spawned a large literature seeking to quantify and demonstrate the value of agricultural research and development (R&D) investments. The most important metric for quantifying the rate of return to R&D emerging from this literature is the internal rate of return (IRR), even though Griliches was sceptical of its usefulness as a metric in this context. An alternative metric, also reported by Griliches but not as commonly used in the subsequent returns-to-research literature, is the benefit–cost ratio (BCR). We assess how the implications of the returns to agricultural R&D literature may have differed if the BCR had become the standard rather than the IRR. We reveal that the IRR and BCR produce substantially different rankings of agricultural R&D projects, differences that persist even under various commodity and geographical aggregations of the BCR and IRR estimates. The median across 2,627 reported IRRs is 37.5 per cent per year. Using data gleaned from 492 research evaluation studies, we developed and deployed a methodology to impute 2,126 BCRs (median of 5.4) and modified internal rates of returns (MIRRs, median of 16.4 per cent per year) assuming a uniform 10 per cent per year discount rate and a 30 year research timeline.

Key words: agriculture, benefit–cost ratio, internal rate of return, modified internal rate of return, research and development.

1. Introduction

An extensive literature on the returns to agricultural research and development (R&D) has emerged since Zvi Griliches' seminal analysis of hybrid corn research (Griliches 1958). While Griliches' initial estimate of approximately $7.00 was constructed to reflect the annual return in perpetuity (ARP) to every dollar invested in hybrid corn research, he also reported a corresponding benefit–cost

*All the authors are affiliated with the University of Minnesota's International Science and Technology Practice and Policy (InSTePP) centre. Rao is also a Faculty Fellow at the Sheila and Robert Challey Institute for Global Innovation and Growth at North Dakota State University. They thank Connie Chan-Kang, Michelle Hallaway, Louise Letnes and Robert Andrade for their excellent research assistance. This paper was prepared with support from the CGIAR's Standing Panel on Impact Assessment (SPIA), with additional support from the University of Minnesota and the Minnesota Agricultural Experiment Station (MIN-14-061 and MIN-14-134).

[†]Xudong Rao (email: xudong.rao@ndsu.edu) is Assistant Professor at the Department of Agribusiness and Applied Economics, North Dakota State University, Fargo, North Dakota, USA. Terrance M. Hurley a Professor and Philip G. Pardey a Professor are with the Department of Applied Economics, University of Minnesota, St. Paul, Minnesota, USA.

doi: 10.1111/1467-8489.12388

ratio (BCR) of approximately 150:1, and, on a suggestion from Martin J. Bailey (Griliches 1958, footnote 16), an internal rate of return (IRR) between 35 and 40 per cent.[1] Of these three metrics, it is the IRR that became the standard – of the 2,242 project/program evaluations from 371 different studies summarised by Hurley *et al.* (2014), an overwhelming 91.4 per cent reported an IRR estimate. BCR estimates were also commonly reported (28.2 per cent of the evaluations), sometimes in conjunction with an IRR estimate (19.6 per cent of the evaluations), while Griliches' ARP did not stand the test of time.

The IRR's predominance emerged even though Griliches expressed consternation with it:

> My objection to this procedure is that it values a dollar spent in 1910 at \$2,300 in 1933. This does not seem very sensible to me. I prefer to value a 1910 dollar at a reasonable rate of return on some alternative social investment. (Griliches 1958, p. 425)

But, it is not just the agricultural R&D literature that has opted to rely heavily on the IRR as a summary statistic for quantifying investment returns. It has been used by researchers to summarise the economic consequences of manufacturing R&D (see the tabulation in Hall *et al.* 2010, Table 2) and health-related R&D (e.g. HERG *et al.* 2008; Deloitte 2014; Glover *et al.* 2014), as well as investments in education (e.g. European Commission 2005; Heckman *et al.* 2006). Private businesses also commonly use it, in conjunction with other metrics, to evaluate prospective investments (e.g. Graham and Harvey 2001; Ryan and Ryan 2002; Truong *et al.* 2008; Bennouna *et al.* 2010; Daunfeldt and Hartwig 2014).

Critiques of using the IRR to measure the value of an investment date back to Griliches' seminal paper (e.g. Hirshleifer 1958) and continue to the present (e.g. Alston *et al.* 2011; Hurley *et al.* 2014, 2017). Defences of the IRR also date back to Griliches' seminal paper (e.g. Bailey 1959) and continue to the present (e.g. Oehmke 2017). While proponents of the IRR often cite its computation without an explicit discount rate assumption as a virtue, Hurley *et al.* (2014) identified a potentially unpalatable discount rate assumption that is implicit in its computation – the IRR is the value that equates the MIRR with the discount rate used to calculate it. Burns and Walker (1997) attributed the IRR's attractiveness as a metric to its interpretability as a percentage, much like the percentage commonly reported for a range of financial products (e.g. mortgages, certificates of deposit, mutual funds and credit cards).[2] However, Hurley *et al.*

[1] The \$7.00 rate of return in perpetuity (or 'external rate of return') splits the difference between estimates assuming 5 per cent and 10 per cent discount rates: \$7.43 and \$6.89 (see Griliches 1958, p. 425 and Table 2 where we corrected the reported \$0.03 to \$7.43 based on the formulas and data provided in the table). The BCR of 150 assumed a 5 per cent discount rate. Griliches also reported a BCR of 70 assuming a discount rate of 10 per cent.

[2] Anecdotally, we have attended presentations by distinguished economists and government officials where the IRR for a project/program is erroneously compared with the annualised rate of return being earned on their retirement accounts to illustrate just how attractive a project/program was as an investment.

(2014) showed that such an annualised percentage rate of return interpretation is generally incorrect and often leads to incredible implications (see also Alston *et al*. 2011, pp. 1,271–1,272). They also pointed out how the modified internal rate of return, MIRR (Lin 1976), offers an alternative to the IRR that can be reasonably interpreted as an annualised percentage rate of return.

Is such quibbling over assumptions and interpretations more than just an academic exercise? The objective of agricultural R&D evaluation research is to provide information to policymakers about the profitability of investments as standalone programs as well as relative to other investments. While the IRR provides a useful metric for determining whether or not an investment is profitable as a standalone program, it is not generally recommended by textbooks for comparing the relative profitability of alternative investments (Kierulff 2008; Daunfeldt and Hartwig 2014; Robison *et al*. 2015). This is particularly true when the investment costs span an extended time period, which is the usual case for agricultural R&D. It is also well established that the ranking of investments by the IRR need not match the ranking implied by other popular metrics like the BCR, which is often recommended by textbooks as a useful metric for comparing among different research projects/programs. The potential for differences in rankings between the BCR and IRR suggests that the implications of the extensive agricultural R&D literature may have looked different had researchers preferred the BCR to the IRR, which makes the IRR debate of practical concern as well as an academic exercise.

Our primary objective in this study is to determine the practical importance of using the BCR instead of the IRR to evaluate and rank investments in agricultural R&D. This objective is accomplished by comparing the rankings of 412 program evaluations from the International Science and Technology Practice and Policy (InSTePP) centre's returns-to-research (RtR) database (version 3) (Pardey *et al*. 2016) that estimated both the BCR and IRR. Since rankings based on the BCR are sensitive to the interest rate used to discount an investment's benefits and costs, and the discount rate used in the literature has varied, the robustness of this comparison is further explored by recalibrating the BCRs following Hurley *et al*. (2014), first with a common 10 per cent discount rate, and then with a common 20 per cent discount rate. Given the substantial and robust differences in rankings revealed by this primary objective, we then use a formulaic relationship between IRR and BCR to econometrically project and summarise additional 1,714 BCRs based on IRR estimates from the RtR database that did not already report a corresponding BCR.

Our contributions in this article are twofold. First, we demonstrate that the debate surrounding the choice of metrics for evaluating and comparing R&D investments in agriculture is more than just an academic exercise. With the two metrics evaluated here producing substantially different rankings, the practical importance of understanding the key assumptions employed by each is heightened, as is the importance of choosing the best metric for measuring and communicating the value of alternative agricultural R&D investments. Second, we provide a detailed summary of the returns to agricultural R&D in terms of the BCR (and MIRR), so researchers and policymakers who share Griliches' (and our) concerns regarding the IRR have alternative benchmarks, including one that is reasonably interpretable in percentage terms, for comparing the profitability of past and future programs.

2. Data

The data for our analysis come from the InSTePP returns-to-research (RtR) database (version 3) (Pardey *et al*. 2016). The goal of the RtR database is to curate and periodically update a comprehensive listing of agricultural R&D evaluation studies, and to extract a common set of information from these studies for use by researchers and policymakers to better understand the consequences of agricultural R&D investment decisions. Version 1 of InSTePP's RtR database was developed in 2000 building on the reviews by Evenson *et al*. (1979), Echeverría (1990) and Alston and Pardey (1996). It included 1,886 program/project evaluations reported in 292 different studies and provided the data used in the analysis reported in the influential and highly cited agricultural R&D surveys by Alston *et al*. (2000a,b). Version 2 of the RtR database was an update to 2011 and formed the basis for the analysis in Hurley *et al*. (2014). It included the evaluation of 2,241 program/project evaluations from 371 different studies. Version 3 mainly consists of new literature published worldwide since 1975, with a targeted effort to include relevant literature for sub-Saharan Africa, Latin America and the Caribbean. It includes 2,829 program/project evaluations from 492 different studies. Each evaluation includes an estimate of the IRR, the BCR or both, for a total of 3,426 returns-to-research estimates.

The RtR database also includes information concerning the study authors, type of publication, who performed the research, where the research was performed, the nature of the research (e.g. basic or applied, or public or private), the commodity or environmental coverage of the research, details of the methodological features of the research and the types of rate of return estimates (i.e. the IRR, BCR or both).

3. The anatomy of agricultural R&D rate of return calculations

A review of the definitions for the IRR, BCR and MIRR, as well as how they are related, facilitates a better understanding of the comparisons and analysis that follow.

Researchers typically characterise the stream of costs and benefits associated with agricultural R&D projects (or programs) using four dates as Figure 1 illustrates. The first date (0 in Figure 1) is the date that resources start to be invested into the project reflecting expenditures. These

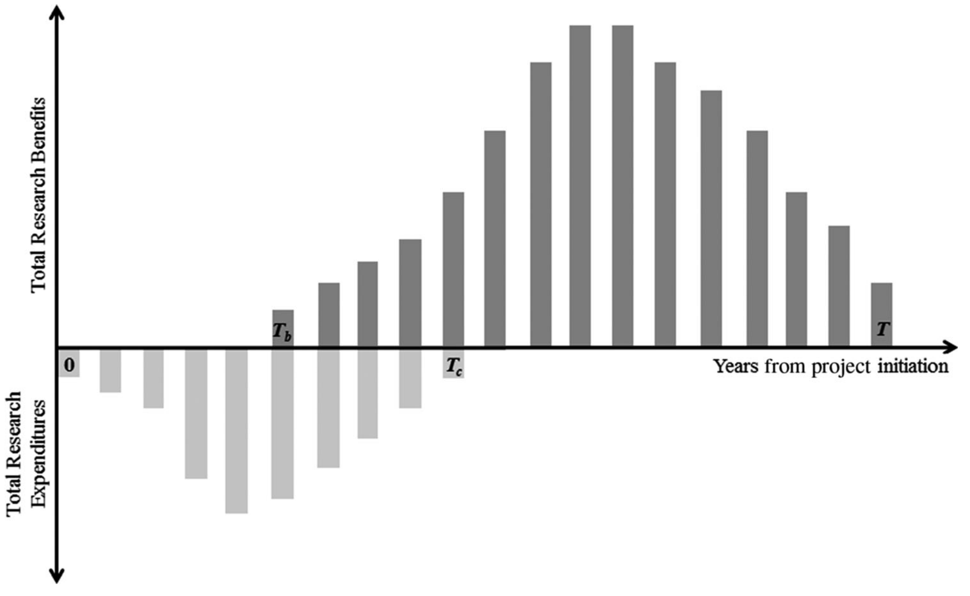

Figure 1 Illustrative research cost and benefit streams.
Source: Authors' construction.

expenditures often span a period, at first growing, then declining, and eventually leading to a second date that signifies the point at which these expenditures cease – T_c in Figure 1. The third date is the date when innovations resulting from these investments begin to reap benefits in terms of increased productivity or reduced costs for example – T_b in Figure 1. As with the investment's expenditures, annual returns tend to initially grow over time, but eventually start to shrink and even cease as new innovations replace the old – T in Figure 1. On average, the evidence we gleaned from the 2,418 evaluations summarised in Table 1 has benefits beginning 6.1 years after the initiation of research costs, cost streams that run over 22.7 years and benefits that accrue for 30.8 years after the initiation of the research. Table 1 also shows there is considerable variation around the sample averages in these various components of the overall research lag.

Within this framework, the present value of an investment's expenditures or research costs is defined by the sum $\mathrm{PVC}(\delta_c) = \sum_{t=0}^{T_c} c_t(1+\delta_c)^{-t}$ where $c_t \geq 0$ is the research cost t years from the project's initiation and $\delta_c > 0$ is a discount rate that reflects the time value of money or the opportunity cost of the resources used to finance the project. The present value at year zero of the return on investment or research benefit is defined by the sum $\mathrm{PVB}(\delta_b) = \sum_{t=T_b}^{T} b_t(1+\delta_b)^{-t}$ where $b_t \geq 0$ is the research benefit t years from the project's initiation and $\delta_b > 0$ is a discount rate that reflects the time value of money or opportunity cost of investing these returns elsewhere. With these present value formulas, the popular IRR is defined as:

Table 1 Research lags, Ts, and discount rates, δs, from the published evaluations

	Sample	Sample size	Mean	SD	Min	25th percentile	Median	75th percentile	Max
		(years)							
T_b	Full	2,418	6.12	6.32	0	1	5	9	49
T_c	Full	2,418	22.72	17.43	0	9	21	32	102
T	Full	2,418	30.77	16.70	0	20	27	42	142
		(% per year)							
δ	Full	1,161	7.88	3.33	2	5	9	10	20
	Reporting IRR	974	8.06	3.36	2	5	9.5	10	20
	Reporting BCR	747	6.9	3.27	2	5	5	10	15
	Reporting IRR and BCR	560	6.89	3.35	2	4	5	10	15
	Reporting nominal ROR	201	9.08	2.79	3	6	10	10	20
	Reporting real ROR	800	7.57	3.54	2	5	8	10	20

Source: Authors' construction from the InSTePP returns-to-research database, version 3.0.

$$\mathrm{IRR} \in \{\delta(\mathrm{PVC}(\delta) = \mathrm{PVB}(\delta))\}. \tag{1}$$

The features of this calculation that helped to make it so popular include the ability to interpret it as a percentage rate and the lack of an explicit assumption regarding the discount rate. Of course, as Equation (1) makes explicit, the IRR need not be unique, which is a well-known drawback that appears to have been of little practical significance in the agricultural R&D literature. The less popular, but still often used, BCR is defined as:

$$\mathrm{BCR}(\delta) = \frac{\mathrm{PVB}(\delta)}{\mathrm{PVC}(\delta)}. \tag{2}$$

Unlike the IRR, to evaluate a project's rate of return using Equation (2), a researcher must make an explicit assumption regarding the discount rate δ. The MIRR, which has gained increasing attention recently, is defined as:

$$\mathrm{MIRR}(\delta_b, \delta_c, T_e) = \sqrt[T_e]{\frac{(1+\delta_b)^{T_e}\mathrm{PVB}(\delta_b)}{\mathrm{PVC}(\delta_c)}} - 1, \tag{3}$$

where T_e reflects the point in time chosen to assess program benefits. This calculation requires assumptions on δ_b, δ_c and T_e, though it was originally proposed assuming $\delta_b = \delta_c = \delta$ and $T_e = T$ (Lin 1976) such that it required no more restrictive assumptions than the BCR. By differentiating between when the final benefits of a program accrue (T), which often differs across programs, and when the benefits of the program are assessed (T_e), it is possible to standardise the MIRR so it has a common interpretation across programs.

These definitions point immediately to several relationships, some more insightful than others:

$$\mathrm{BCR}(\mathrm{IRR}) = 1, \tag{4}$$

$$\mathrm{MIRR}(\mathrm{IRR}, \mathrm{IRR}, T_e) = \mathrm{IRR}\,\text{and}, \tag{5}$$

$$\mathrm{MIRR}(\delta, \delta, T_e) = \sqrt[T_e]{(1+\delta)^{T_e}\mathrm{BCR}(\delta)} - 1. \tag{6}$$

Equation (4) follows immediately from rearranging the equality in Equation (1) and adds little additional insight into the relationship between the BCR and IRR beyond their definitions. Equation (5) was demonstrated by Hurley *et al*. (2014) and is insightful because it provides a clear interpretation of the implicit assumption made in the calculation of the IRR: specifically, the IRR is the discount rate that equals the annualised return on an initial (lump sum) investment of $\mathrm{PVC}(\delta)$ that yields $(1+\delta)^{T_e}\mathrm{PVB}(\delta)$ at year T_e. Both Athanasopoulos (1978) and Negrete (1978) established Equation (6), which is also insightful because it shows that there is a one-to-one correspondence between the BCR and MIRR if a common discount rate and evaluation period (T_e) are used to calculate the rates of returns for different projects. This one-to-one correspondence means that the BCR and MIRR will rank projects identically.

4. Rate of return rankings

Figure 2 compares rankings based on the reported IRRs and BCRs for the 412 project evaluations in the InSTePP database that report both an IRR and BCR. Panel (a) shows the scatter plot of IRR and BCR rankings along with the line of best fit. This line of best fit explains only 26 per cent of the variation in the rankings. The line's positive intercept and slope of less than one imply that the BCR systematically ranks projects with a high IRR lower and projects with a low IRR higher, on average. If they ranked projects identically, the intercept would be zero and the slope would be one. This trend is further highlighted in Panel (b), which shows the differences between the rank based on the BCR and the rank based on the IRR arrayed in ascending order (from left to right) by the IRR's rank. What is more evident in Panel (b) is that there are very large differences in rankings, over 300 places in several instances and commonly over 100 places when only 412 projects are being ranked.

The discount rates used to calculate the reported BCRs for the analysis in Figure 2 differ widely among each of the evaluations, which means we could also expect to see differences in ranking between the BCR and MIRR. These differences can be attributed to differences in the timing and magnitude of both the expenditures and the economic returns over time, or they could be attributable to differences in the discount rate employed in the calculations. To focus on how differences in the timing and magnitude of each project's

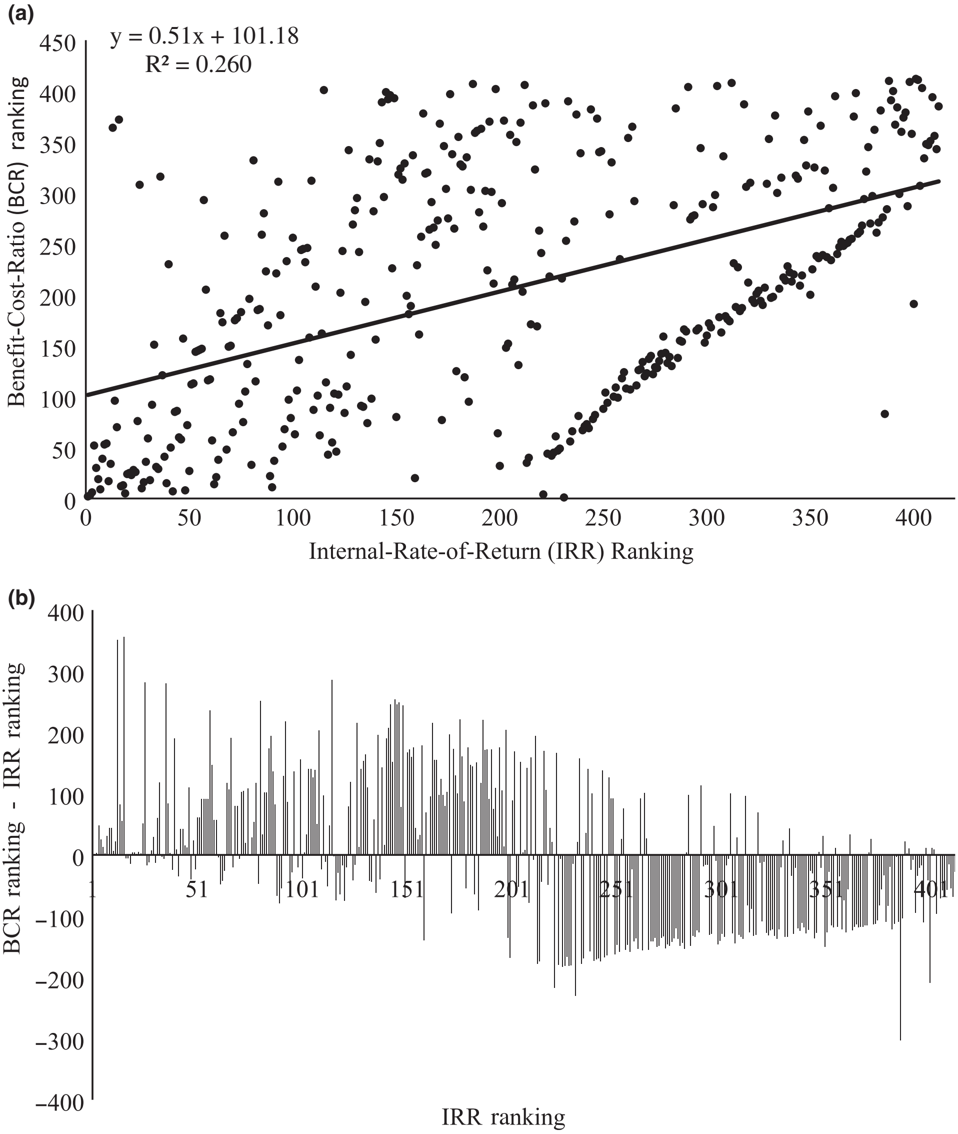

Figure 2 Reported internal rate of returns (IRRs) versus benefit–cost ratios (BCRs) using the study's delta. Panel (a): Scatter plot and ordinary least squares regression results for these rankings. Panel (b) Difference in rankings arranged according to the IRR rankings*Notes*: Comparison of IRR and BCR rankings for the 412 evaluations that reported an IRR and BCR.
Source: Authors' construction.

expenditures and economic returns affect the relative ranking of projects, we used the methodology reported in Hurley *et al.* (2014) to approximate the BCRs for a common discount rate. Figures 3 and 4 show the same patterns as Figure 2, but now with the BCRs imputed using a common annual discount rate of 10 and 20 per cent, respectively.

With a 10 per cent discount rate, the line of best fit between the IRR and BCR explains 68 per cent of the variation in ranks (Figure 3a). Alternatively, 91 per cent of the variation in ranks is explained with a discount rate of 20 per cent (Figure 4a). In both cases, the intercept of the line of best fit is positive, while the slope is less than one. Again, this implies that the ranking based on the BCR is systematically lower for higher IRRs and higher for lower IRRs, on average. The maximum difference in rankings occasionally exceeds 200 and commonly exceeds 100 when using an annual discount rate of 10 per cent (Figure 3b). With an annual discount rate of 20 per cent, the difference in rankings is less pronounced with infrequent incursions above 100 (Figure 4b) and most differences below 50.

The relationships reported in Equations (5) and (6) help to explain why the correspondence between the BCR and IRR improves as the discount rate increases. From Equation (6), we know that the BCR and MIRR will produce identical rankings with a common discount rate and evaluation period. Equation (5) indicates that the IRR and MIRR are equal when the MIRR is calculated using a discount rate equal to the IRR. With an average and median IRR of 63.2 and 38 per cent per year, respectively, for the 412 evaluations with both IRRs and BCRs, increasing the annual discount rate from 10 to 20 per cent tends to reduce the difference between the discount rate and IRR, making the IRR and MIRR ranking results more consistent. Thus, the IRR and BCR ranking results also become more consistent. With that said, 20 per cent represents the maximum annual discount rate used by studies in the InSTePP database to calculate BCRs, while 10 per cent per year is just above the median discount rate used to calculate BCRs in the InSTePP database. Thus, the revealed preferences of returns-to-research analysts who explicitly choose a discount rate favour the results in Figure 3 rather than Figure 4.

5. Recalibrating returns-to-research distributions

To derive BCR estimates for evaluations in the InSTePP database that only included an IRR, we deploy a three-step procedure. The first step is to use the methodology in Hurley *et al.* (2014) to approximate BCRs for the subset of 412 evaluations in the InSTePP database that reported a BCR, an IRR and the time-related information (i.e. T_c, T_b and T), as was done in the previous section. The second was to use regression methods to identify the best-fitting relationship between the reported IRRs and approximated BCRs while accounting for differences in T_c, T_b and T. Finally, these regression results were used to project BCRs for all reported IRR estimates that did not report

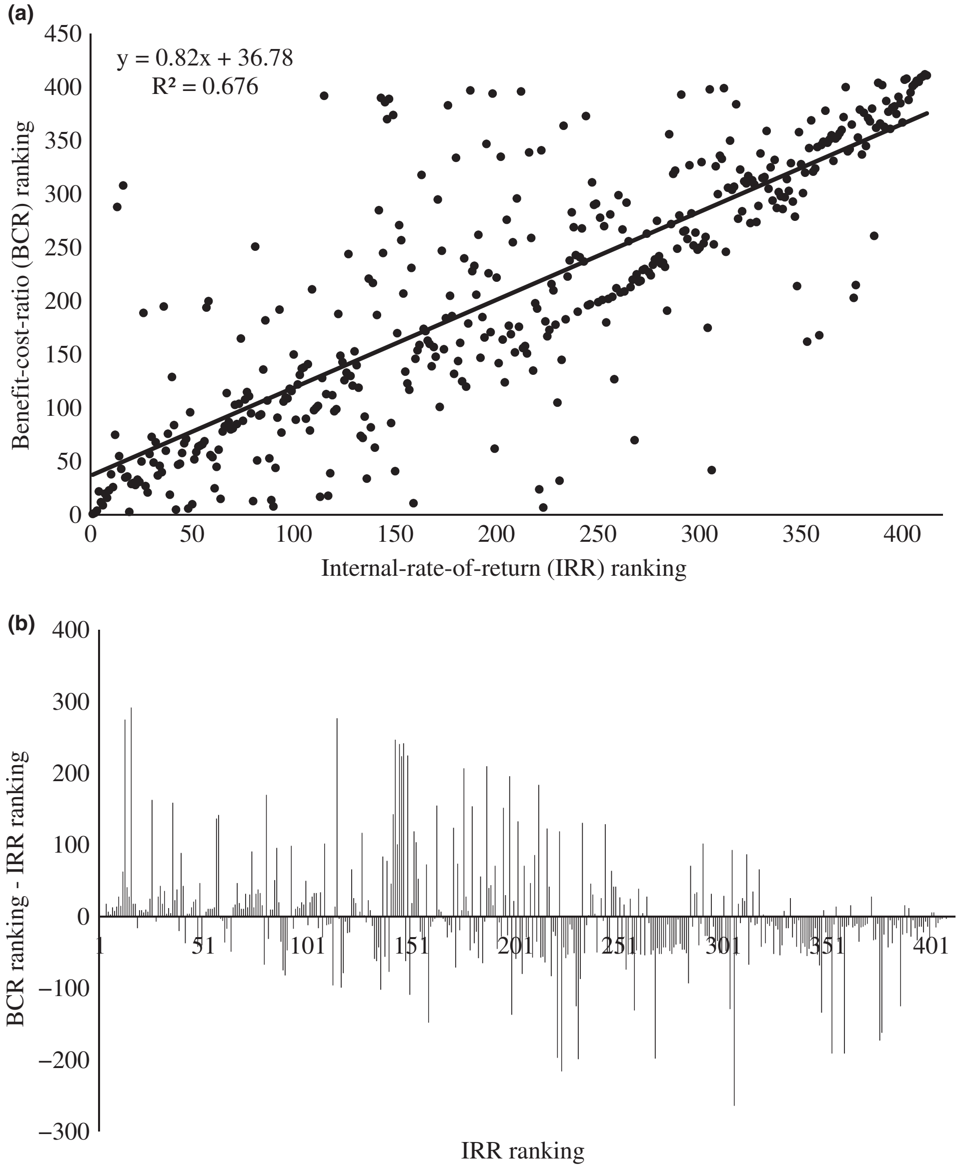

Figure 3 Rankings of reported internal rate of return (IRRs) versus calibrated benefit–cost ratios (BCRs) assuming a 10 per cent discount rate (δ = 0.10). Panel (a): Scatter plot and ordinary least squares regression results for these rankings. Panel (b) Difference in rankings arranged according to the IRR rankings.
Source: Authors' construction.

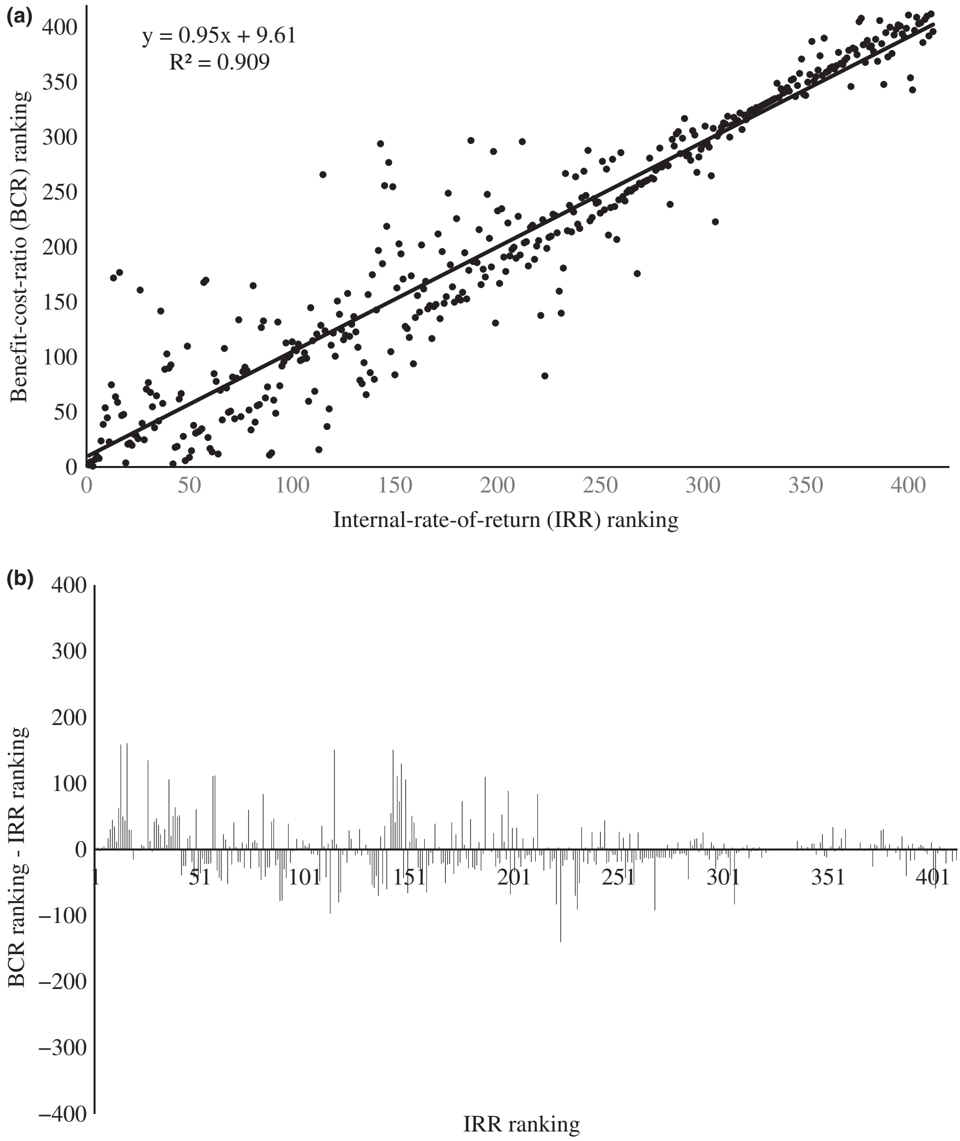

Figure 4 Rankings of reported internal rate of return (IRRs) versus calibrated benefit–cost ratios (BCRs) assuming a 20 per cent discount rate (δ = 0.20). Panel (a): Scatter plot and ordinary least squares regression results for these rankings. Panel (b) Difference in rankings arranged according to the IRR rankings.
Source: Authors' construction.

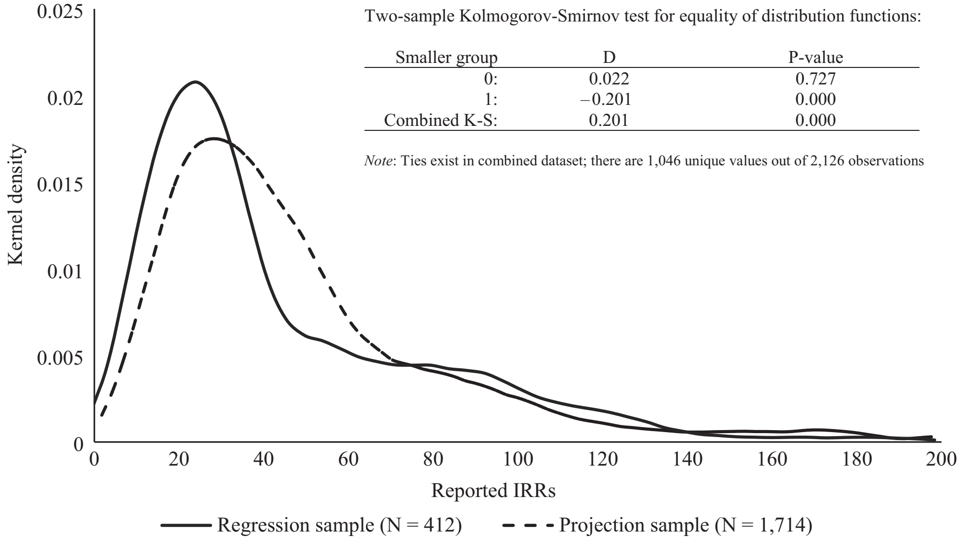

Figure 5 Kolmogorov–Smirnov test of the regression sample versus the prediction sample.

sufficient information to approximate a BCR using the approach developed in Hurley *et al*. (2014), and were within the support of the regression analysis (Figure 5).[3]

5.1 Regression details

Hurley *et al*. (2014) showed the relationship between the BCR and IRR can be written as:

$$\mathrm{BCR} = \frac{\sum_{t=0}^{T_c} w_{c_t}(1+\mathrm{IRR})^{-t}}{\sum_{t=T_b}^{T} w_{b_t}(1+\mathrm{IRR})^{-t}} \frac{\sum_{t=T_b}^{T} w_{b_t}(1+\delta)^{-t}}{\sum_{t=0}^{T_c} w_{c_t}(1+\delta)^{-t}}, \tag{7}$$

where w_{c_t} and w_{b_t} are the proportions of the total costs and total benefits occurring at time t that are usually not reported in the original evaluation studies. This relationship is used to guide the specification of a regression equation. Taking the natural log of Equation (7) yields:

[3] The combined Kolmogorov–Smirnov test for equality of the regression sample (N = 412) and projection sample (N = 1,714) is 0.201 with a P-value < 0.000, indicating that the two distributions are not equal (see Figure 5). Furthermore, the one-sided test and summary statistics (e.g. the median and percentiles) suggest that the projection sample lies to the right of the regression sample. To avoid out-of-sample forecasts, projections were restricted to IRRs, T_cs, T_bs and Ts within the minimum and maximum values observed in the regression sample.

$$\begin{aligned}\ln(BCR)= & \ln\left(\sum_{t=0}^{T_c} w_{c_t}(1+\mathrm{IRR})^{-t}\right)-\ln\left(\sum_{t=T_b}^{T} w_{b_t}(1+\mathrm{IRR})^{-t}\right)\\ & +\ln\left(\sum_{t=T_b}^{T} w_{b_t}(1+\delta)^{-t}\right)-\ln\left(\sum_{t=0}^{T_c} w_{c_t}(1+\delta)^{-t}\right).\end{aligned} \tag{8}$$

Define $f((x)\mathbf{w},T_l,T_u)=\ln\left(\sum_{t=T_l}^{T_u} w_t x^{-t}\right)$ where $w_t>0$ for $t = T_l,\cdots,T_u$ and $\sum_{t=T_l}^{T_u} w_t = 1$. The first and second derivatives of this function with respect to x are:

$$f'((x)\mathbf{w},T_l,T_u)=\frac{-\sum_{t=T_l}^{T_u} t w_t x^{-t-1}}{\sum_{t=T_l}^{T_u} w_t x^{-t}}, \text{ and} \tag{9}$$

$$f''((x)\mathbf{w},T_l,T_u)=\frac{\sum_{t=T_l}^{T_u} t(t+1) w_t x^{-t-2}\sum_{t=T_l}^{T_u} w_t x^{-t}-\left(\sum_{t=T_l}^{T_u} t w_t x^{-t-1}\right)^2}{\left(\sum_{t=T_l}^{T_u} w_t x^{-t}\right)^2}. \tag{10}$$

The second-order Taylor series approximation can then be written as:

$$\begin{aligned}f((1+\delta)\mathbf{w},T_l,T_u) &=-\sum_{t=T_l}^{T_u} t w_t\delta+\left(\sum_{t=T_l}^{T_u} t w_t+\sum_{t=T_l}^{T_u} t^2 w_t-\left(\sum_{t=T_l}^{T_u} t w_t\right)^2\right)\frac{\delta^2}{2!}\\ &+\varepsilon((\delta)\mathbf{w},T_l,T_u)=\left(\sum_{t=T_l}^{T_u} t w_t\right)\delta\left(\frac{\delta}{2!}-1\right)\\ &+\left(\sum_{t=T_l}^{T_u} t^2 w_t-\left(\sum_{t=T_l}^{T_u} t w_t\right)^2\right)\frac{\delta^2}{2!}+\varepsilon((\delta)\mathbf{w},T_l,T_u)\\ &=\mu(\mathbf{w},T_l,T_u)\delta\left(\frac{\delta}{2!}-1\right)+\sigma^2(\mathbf{w},T_l,T_u)\frac{\delta^2}{2!}\\ &+\varepsilon((\delta)\mathbf{w},T_l,T_u),\end{aligned} \tag{11}$$

where $\mu(\mathbf{w},T_l,T_u)=\sum_{t=T_l}^{T_u} t w_t$ and $\sigma^2(\mathbf{w},T_l,T_u)=\sum_{t=T_l}^{T_u} t^2 w_t-\mu(\mathbf{w},T_l,T_u)^2$. Substitution back into Equation (8) then yields:

$$\begin{aligned}\ln(\mathrm{BCR}) &= (\sigma^2(\mathbf{w}_c,0,T_c)-\sigma^2(\mathbf{w}_b,T_b,T))\left(\frac{IRR^2}{2!}-\frac{\delta^2}{2!}\right)\\ &+(\mu(\mathbf{w}_c,0,T_c)-\mu(\mathbf{w}_b,T_b,T))\left(\mathrm{IRR}\left(\frac{\mathrm{IRR}}{2!}-1\right)-\delta\left(\frac{\delta}{2!}-1\right)\right)\\ &+\varepsilon((\mathrm{IRR})\mathbf{w}_c,0,T_c)-\varepsilon((\delta)\mathbf{w}_c,0,T_c)+\varepsilon((\delta)\mathbf{w}_b,T_b,T)\\ &-\varepsilon((IRR)\mathbf{w}_b,T_b,T).\end{aligned} \tag{8'}$$

In Equation (8'), terms with $\mathbf{w}_c$ and T_c are linearly separable from terms with $\mathbf{w}_b$ and T_b and T. Defining $x_1(\delta)=\left(\mathrm{IRR}\left(\frac{\mathrm{IRR}}{2!}-1\right)-\delta\left(\frac{\delta}{2!}-1\right)\right)$ and $x_2(\delta)=\left(\frac{\mathrm{IRR}^2}{2!}-\frac{\delta^2}{2!}\right)$, the regression equation we propose respects this separability, such that

$$\begin{aligned}\ln(BCR_i(\delta)) &= \alpha_0+\alpha_1 T_{c_i}+\alpha_2 T_{c_i}^2+\alpha_3 T_{b_i}+\alpha_4 T_{b_i}^2+\alpha_5 T_i+\alpha_6 T_i^2+\alpha_7 T_{b_i} T_i\\ &+(\beta_0+\beta_1 T_{c_i}+\beta_2 T_{b_i}+\beta_3 T_i+\beta_4 T_{b_i} T_i)x_{1_i}(\delta)\\ &+(\gamma_0+\gamma_1 T_{c_i}+\gamma_2 T_{b_i}+\gamma_3 T_i+\gamma_4 T_{b_i} T_i)x_{2_i}(\delta)\\ &+\varepsilon_i(\delta).\end{aligned} \tag{12}$$

Equation (12) indicates how the reported BCRs relate to other information (i.e. time-related variables, the discount rate δ and the IRR) that is often reported in R&D evaluation studies. Line 1 of this equation includes time-related variables, their squared terms and the interaction term only between T_b and T since Equation (8') suggests T_b and T are linearly separable from T_c but not from each other. Line 2 includes the interaction terms between time-related variables and x_1, which boils down to the reported IRR, assumed discount rate (δ) and their squared terms. Line 3 differs from line 2 in that line 3 includes only the squared terms of the reported IRR and δ. Finally, $\mathbf{w}_{c_t}$ and $\mathbf{w}_{b_t}$, the proportions of the total costs and total benefits occurring at time t, are usually not reported in the original evaluation studies, and their effects on the reported BCR are captured in the parameter vectors of $\boldsymbol{\alpha}$,$\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ and the residual term $\varepsilon_i(\delta)$. To estimate these coefficients so we can project the BCR into more observations, we recalibrate the 412 BCRs with discount rates equal to 5, 10, 15 and 20 per cent to generate 1,648 total observations.

Being aware of evident outliers in the 1,648 observations, we use and compare various robust regression estimators to obtain the best-fitting empirical relationship between the reported IRRs and approximated BCRs. The OLS estimates are first derived to benchmark the comparisons (Table 2, Column 2). We then estimated several variants of the M-estimator (Huber

1973) – a commonly used robust regression estimator that is robust to vertical outliers, but not bad leverage points.[4] Specifically, the least absolute value (LAV) estimator (Table 2, Column 4) focuses on the median rather than the mean errors to mute the influence of outliers. However, the LAV estimator suffers from low efficiency, so the Huber M-estimator (Table 2, Column 5) and the bisquare M-estimator (Table 2, Column 3) are employed as two alternatives with improved efficiency. The S-estimator from the second generation of robust regression estimators is also employed to detect potential bad leverage points. For the purposes of this study, we also deploy an MM-estimator (Yohai 1987) and report the results in Table 2, Column 6. By combining an M-estimator with an S-estimator, MM-estimators have high efficiency while preserving a high breakdown point.[5]

Comparing the coefficient estimates from the different models, we find that the signs, magnitude and statistical significance vary appreciably. Table 3 reports the results from projecting the 2,126 BCRs (in the log form) using the relationship identified by each model. While the coefficients for the various models differ appreciably, the projected medians as well as the 25th and 75th percentiles are similar. What differs more appreciably among these projections is the extreme values. With more muted extremes, the OLS estimates are less inclined to produce nonsensically high or low projections, which is why our subsequent results rely on these estimates.

5.2 Imputed BCRs and MIRRs

Table 4 reports distribution statistics for the IRRs and imputed BCRs using four different annual discount rates – specifically 5, 10, 15 and 20 per cent, which span the range of discount rates in the InSTePP database (Table 1). The BCR tells us the present value of the benefits of a project per dollar invested, also in present value, which is not as ubiquitous a metric as the annualised rates of return commonly reported for financial products (e.g. mortgages, certificates of deposit, mutual funds and credit cards). However, the MIRR does have an interpretation that is consistent with the annualised rates of return commonly reported for these types of financial products. Therefore, to make our result more comparable to the common annual rates of returns reported (and that better concord with the rates of return more

[4] Rousseeuw & Leroy (1987) categorised outlying observations into three types: vertical outliers, bad leverage points and good leverage points. Vertical outliers refer to observations that have outlying values for the corresponding error term but not for the explanatory variables. Good leverage points refer to those with outlying values for the explanatory variables but are located close to the regression line. Bad leverage points refer to observations that have outlying values for the explanatory variables but are located well away from the regression line.

[5] A breakdown point is the number of outliers that can be included in the analysis before the analysis is adversely affected. The higher the value of the breakdown point is, the more robust the model estimates are. The OLS estimator admits a breakdown point of zero, thus being nonrobust to outliers.

Table 2 Regression results relating benefit–cost ratio to IRR – OLS and various robust regression models

	OLS	RREG	LAV	*M*	MM85
x_1	1.010*** (0.183)	−1.680*** (0.147)	0.396 (0.232)	0.283 (0.317)	2.302*** (0.689)
$T_c\ x_1$	−0.0561*** (0.010)	0.00735 (0.008)	−0.0418*** (0.012)	−0.0419** (0.015)	0.134*** (0.039)
$T_b\ x_1$	−0.445*** (0.041)	−0.113*** (0.033)	−0.432*** (0.052)	−0.404*** (0.057)	−0.445** (0.159)
$T\ x_1$	−0.101*** (0.008)	−0.117*** (0.006)	−0.115*** (0.010)	−0.0987*** (0.010)	−0.377*** (0.018)
$T\ T_b\ x_1$	0.00158 (0.001)	−0.00840*** (0.001)	0.000884 (0.001)	0.000663 (0.001)	−0.00226 (0.003)
x_2	−0.809*** (0.212)	−0.309 (0.169)	0.334 (0.268)	−0.159 (0.258)	5.585*** (1.087)
$T_c\ x_2$	−0.0148* (0.006)	−0.118*** (0.005)	−0.0135 (0.008)	−0.0290* (0.014)	0.270*** (0.037)
$T_b\ x_2$	0.604*** (0.072)	1.037*** (0.058)	0.381*** (0.091)	0.549*** (0.099)	0.51 (0.348)
$T\ x_2$	0.0918*** (0.010)	0.209*** (0.008)	0.0863*** (0.013)	0.0922*** (0.018)	−0.271*** (0.063)
$T\ T_b\ x_2$	−0.0169*** (0.003)	−0.0531*** (0.003)	−0.0178*** (0.004)	−0.0179** (0.006)	−0.0427*** (0.013)
T_c	−0.0306*** (0.007)	−0.0275*** (0.006)	−0.0290** (0.009)	−0.0267*** (0.006)	−0.0216 (0.013)
T_c^2	0.000346** (0.000)	0.000411*** (0.000)	0.000396** (0.000)	0.000327*** (0.000)	0.000396 (0.000)
T_b	−0.00822 (0.014)	0.0445*** (0.011)	0.00733 (0.017)	0.00654 (0.012)	0.00829 (0.012)
T_b^2	0.000696 (0.000)	−0.00057 (0.000)	0.000419 (0.000)	0.000407 (0.000)	0.000561 (0.001)
T	0.00896 (0.008)	0.0255*** (0.007)	0.0233* (0.010)	0.0160* (0.007)	0.0248 (0.016)
T^2	−0.000257* (0.000)	−0.000391*** (0.000)	−0.000427** (0.000)	−0.000308** (0.000)	−0.00049 (0.000)
Intercept	0.933*** (0.114)	−0.0506 (0.092)	0.307* (0.145)	0.514*** (0.120)	−0.0138 (0.091)
Goodness of fit	0.72	0.90	0.54	0.66	0.63
N	1,648	1,648	1,648	1,648	1,648

Note: $*P < 0.1$; $**P < 0.05$; $***P < 0.01$. OLS refers to the ordinary least square estimator; RREG refers to one version of the M-estimator that is based upon Cook's distance and estimated by the Stata command *rreg*; LAV refers to the least absolute value estimator, also known as the quantile regression estimator, and is estimated by the Stata command *qreg*; M and MM85 refer to the M-estimator and MM-estimator, respectively, and both are estimated by the Stata command *robreg* with the corresponding options. Robust standard errors in parentheses.
Source: Authors' construction.

Table 3 Descriptive statistics of projected BCRs (in the log form) under various robust regression models

	OLS	RREG	LAV	M	MM85
Sample size	2,126	2,126	2,126	2,126	2,126
Mean	1.43	−0.32	1.12	1.27	−0.38
SD	6.87	26.63	9.79	8.16	34.48
Minimum	−189.74	−748.22	−260.25	−224.06	−791.28
25th percentile	1.09	0.97	0.93	0.99	1.00
Median	1.64	1.64	1.62	1.62	1.85
75th percentile	2.54	2.37	2.62	2.54	2.84
Maximum	8.13	39.37	11.76	10.82	315.52

Note: Table entries report a trimmed sample of 2,126 evaluations, which includes 412 approximated BCRs and 1,714 regression predicted estimates. OLS refers to the ordinary least square estimator; RREG refers to one version of the M-estimator that is based upon Cook's distance and estimated by the Stata command *rreg*; LAV refers to the least absolute value estimator, also known as the quantile regression estimator, and is estimated by the Stata command *qreg*; M and MM85 refer to the M-estimator and MM-estimator, respectively, and both are estimated by the Stata command *robreg* with the corresponding options. The 'goodness of fit' statistics are included, but represent a mixture of R^2 and pseudo-R^2 statistics that are not directly comparable across the models.
Source: Authors' construction.

likely to be understood by policymakers), we also report the MIRR distribution statistics for the four discount rates assuming a time to maturity of 30 years (i.e. $T_e = 30$), which is the average life of projects in the InSTePP database.[6]

The mean rate of return for the IRRs is 59.2 per cent per year with a standard deviation of 84.1, median of 39.0 and interquartile range of 40.5. The mean rate of return based on the BCR (in the level form) with a 10 per cent annual discount rate is 18.0 with a standard deviation of 67.0, median of 5.4 and interquartile range of 10.5. Converting these BCRs to MIRRs using Equation (6) produces a mean rate of return of 16.7 per cent per year with a standard deviation of 9.4, median of 16.4, and interquartile range of 6.0 per cent per year. Comparing the BCR and MIRR results as the discount rate increases reveals apparently contradictory results. Increasing the discount rate results in lower BCRs, suggesting a declining value, but higher MIRRs suggesting an increasing value. Mathematically, the relationship between the BCR and the discount rate, and the MIRR and the discount rate is ambiguous, reiterating that while the two measures rank projects identically when evaluated with a common discount rate, this need not be the case when evaluations are made with heterogeneous discount rates.

Tables 5 and 6 report distribution statistics for the reported IRR and the BCRs imputed at a discount rate of 10 per cent per year for a range of different categories. The estimates reported in these tables have several notable features. In all cases, there are large ranges in the returns-to-research estimates around the central tendencies of the respective distributions.

[6] In the context of the returns-to-research evaluation literature, T_e represents the period from the initiation of research costs to the cessation of evaluated research benefits.

Table 4 Descriptive statistics, reported IRRs and imputed BCRs ($N = 2{,}126$)

	IRR	BCR – Imputed				MIRR – Imputed			
						$T_e = 30$ and			
		$\delta = 0.05$ % per year	$\delta = 0.10$	$\delta = 0.15$	$\delta = 0.20$	$\delta_b = \delta_c = 0.05$ % per year	$\delta_b = \delta_c = 0.10$	$\delta_b = \delta_c = 0.15$	$\delta_b = \delta_c = 0.20$
Mean	59.2	28.9	18.0	12.1	8.6	12.8	16.7	20.7	24.7
SD	84.1	111.7	67.0	41.9	27.1	9.2	9.4	9.7	10.0
Minimum	7.4	0.0	0.0	0.0	0.0	−100.0	−100.0	−100.0	−100.0
25th percentile	24.5	3.7	2.9	2.3	1.8	9.7	14.0	18.2	22.3
Median	39.0	8.2	5.4	3.9	3.1	12.6	16.4	20.3	24.6
75th percentile	65.0	20.0	13.5	9.8	7.3	16.0	20.0	24.1	28.2
Maximum	1,736.0	3,549.3	2,179.9	1,357.4	856.9	37.9	42.1	46.3	50.3

Note: The sample includes 412 directly estimated IRRs (for which there were corresponding BCRs) plus 1,714 econometrically derived estimates formed from the regression parameters summarised in Table 2, with the projected IRRs trimmed to avoid out-of-sample forecasts.
Source: Authors' construction.

Table 5 Categorised distribution statistics for the reported internal rate of returns (% per year)

	N	Mean	SD	Min	25th percentile	Median	75th percentile	Max
Crops	1,068	58.1	73.7	7.7	28.0	42.9	69.0	1,736.0
Wheat	181	48.9	34.0	8.8	26.9	41.9	62.7	290.0
Maize	156	55.8	40.7	9.0	29.2	43.0	66.3	291.4
Rice	84	78.1	73.0	15.9	32.0	60.6	95.0	466.0
Millet and sorghum	100	47.0	40.6	7.7	21.5	31.5	57.0	179.0
Other cereals	40	40.9	22.3	14.6	22.0	32.5	53.0	90.9
Roots and tubers	76	51.8	32.9	10.0	30.3	41.9	63.9	202.0
Horticulture (fruit, vegetables and nuts)	42	130.4	293.5	8.0	28.3	41.3	87.3	1,736.0
Other crops	389	57.0	45.2	10.0	29.0	46.8	66.0	350.0
Livestock	195	75.4	88.6	9.0	33.0	56.0	91.4	910.0
All agriculture	695	50.9	85.0	7.4	20.1	28.5	45.6	1219.0
Natural resources	28	46.6	30.9	9.3	15.9	39.5	74.5	111.2
Developed countries	1,143	61.6	100.1	7.4	21.9	35.5	64.3	1,736.0
The United States	796	54.6	76.5	7.4	21.6	32.7	54.0	910.0
Australia and New Zealand	112	95.0	201.4	9.0	19.0	29.3	76.0	1,736.0
Other developed countries	235	69.3	95.3	10.5	28.0	56.0	84.6	1,219.0
Developing countries	848	57.9	58.7	7.7	28.0	43.0	68.5	1,000.0
Asia and Pacific	247	83.9	91.7	14.0	36.7	52.0	95.0	1,000.0
Latin America and the Caribbean	358	46.8	28.1	8.0	26.9	40.1	59.0	191.0
Sub-Saharan Africa	243	47.6	37.1	7.7	23.2	37.3	59.0	350.0
Multinational	101	50.6	78.4	10.0	26.8	35.0	51.5	677.0
Global	13	44.0	23.2	10.0	26.0	48.0	52.0	84.2
High income	1,171	61.5	99.2	7.4	22.0	36.0	64.3	1,736.0
Middle income	683	61.4	67.6	7.7	28.8	43.6	73.0	1,000.0
Low income	109	40.5	26.3	9.0	23.5	33.7	49.0	188.0
All estimates	2,126	59.2	84.1	7.4	24.5	39	65	1,736

Note: The income groups in this table come from the World Bank Analytical Classification for the fiscal year of 2017 (accessed from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519). In this study, 'middle-income' countries include both upper-middle (UM)- and lower-middle (LM)-income countries from the World Bank classification.
Source: Authors' construction.

Table 6 Categorised distribution statistics for the imputed benefit–cost ratios with $\delta = 0.1$ (10%)

	N	Mean	SD	Min	25th percentile	Median	75th percentile	Max
Crops	1,068	16.0	34.9	0.9	3.3	6.4	14.4	607.4
Wheat	181	13.2	18.9	1.3	3.4	6.6	14.7	132.2
Maize	156	13.7	15.7	1.3	4.7	8.3	17.2	117.7
Rice	84	34.3	90.0	1.3	3.4	7.6	25.2	607.4
Millet and sorghum	100	13.1	16.8	1.4	2.6	4.5	18.5	64.5
Other cereals	40	10.2	10.8	1.9	2.9	4.6	13.6	39.6
Roots and tubers	76	10.2	19.0	1.3	2.3	4.4	8.4	129.7
Horticulture (fruit, vegetables and nuts)	42	33.6	70.6	0.9	3.5	7.6	25.2	378.0
Other crops	389	14.9	23.3	1.1	3.3	6.8	12.9	141.6
Livestock	195	16.8	27.0	0.0	3.2	7.0	22.5	274.2
All agriculture	695	22.0	105.2	0.0	2.5	4.2	8.8	2,179.9
Natural resources	28	7.0	6.0	0.9	2.3	5.0	10.8	21.2
Developed countries	1,143	20.0	85.8	0.0	2.8	4.7	12.0	2,179.9
The United States	796	20.2	100.0	0.0	2.6	4.4	9.0	2,179.9
Australia and New Zealand	112	16.2	42.8	0.8	3.0	4.8	11.2	378.0
Other developed countries	235	21.0	32.5	0.0	3.2	8.9	24.6	182.2
Developing countries	848	16.1	35.2	0.0	3.2	6.4	14.3	607.4
Asia and Pacific	247	19.7	45.5	0.0	5.1	9.4	16.2	607.4
Latin America and the Caribbean	358	14.5	35.2	0.9	2.7	5.0	10.7	503.7
Sub-Saharan Africa	243	14.9	19.6	1.1	2.6	6.8	20.1	141.6
Multinational	101	15.4	25.7	1.4	3.3	6.9	19.0	176.0
Global	13	5.9	5.0	1.3	2.4	3.1	7.8	14.1
High income	1,171	19.6	84.8	0.0	2.8	4.8	11.8	2,179.9
Middle income	683	16.6	38.0	0.0	3.4	6.6	14.3	607.4
Low income	109	8.9	10.2	1.1	2.8	4.4	10.3	47.9
All estimates	2,126	18.0	67.0	0.0	2.9	5.4	13.5	2,179.9

Note: See Table 5.
Source: Authors' construction.

Table 7 Categorised distribution statistics for the imputed MIRRs (% per year) with $\delta_b = \delta_c = 0.1$ (10%) and $T_e = 30$

	N	Mean	SD	Min	25th percentile	Median	75th percentile	Max
Crops	1,068	17.8	4.3	9.7	14.4	17.0	20.2	36.2
Wheat	181	17.7	4.0	10.8	14.6	17.2	20.3	29.4
Maize	156	18.3	3.7	11.0	15.8	18.1	20.9	28.9
Rice	84	18.9	5.6	11.0	14.6	17.7	22.5	36.2
Millet and sorghum	100	17.2	4.5	11.3	13.5	15.6	21.2	26.4
Other cereals	40	17.1	3.7	12.4	14.0	15.7	20.0	24.4
Roots and tubers	76	16.4	3.8	11.0	13.2	15.6	18.1	29.4
Horticulture (fruit, vegetables and nuts)	42	18.9	5.9	9.7	14.6	17.7	22.5	34.1
Other crops	389	17.8	4.2	10.2	14.5	17.3	19.8	29.7
Livestock	195	17.7	7.6	−61.9	14.3	17.3	22.0	32.6
All agriculture	695	15.6	10.5	−100.0	13.3	15.4	18.3	42.1
Natural resources	28	15.8	3.8	9.5	13.1	16.0	19.1	21.8
Developed countries	1,143	15.8	12.1	−100.0	13.8	15.8	19.5	42.1
The United States	796	14.9	13.5	−100.0	13.6	15.5	18.4	42.1
Australia and New Zealand	112	17.1	4.5	9.1	14.0	15.9	19.2	34.1
Other developed countries	235	18.2	8.8	−85.9	14.3	18.3	22.4	30.8
Developing countries	848	17.8	4.5	−8.6	14.3	17.0	20.2	36.2
Asia and Pacific	247	18.8	4.5	−8.6	16.2	18.5	20.7	36.2
Latin America and the Caribbean	358	17.0	4.3	9.7	13.7	16.1	19.1	35.4
Sub-Saharan Africa	243	17.8	4.5	10.2	13.6	17.3	21.6	29.7
Multinational	101	18.0	4.2	11.3	14.5	17.3	21.4	30.7
Global	13	15.5	3.1	11.0	13.3	14.2	17.8	20.2
High income	1,171	15.8	11.9	−100.0	13.8	15.9	19.4	42.1
Middle income	683	17.8	4.5	−8.6	14.5	17.2	20.2	36.2
Low income	109	16.5	3.6	10.2	13.9	15.6	18.9	25.1
All estimates	2,126	16.7	9.4	−100.0	14.0	16.4	20.0	42.1

Note: See Table 5.
Source: Authors' construction.

Moreover, the differences in implication between the IRR and BCR persist even with significant aggregation. For example, based on the IRR, the median return to agricultural R&D in the United States outpaced the median return in Australia and New Zealand. Based on the BCR, the median return to agricultural R&D in Australia and New Zealand outpaced the median return in the United States. By commodity, investments in rice yielded the highest median return based on the IRR, while maize investments yielded the highest median return based on the BCR. The median BCR ranked horticultural crops third out of eight commodity categories, while the median IRR ranked horticultural crops only sixth out of eight. Investments in agricultural R&D yielded a higher median IRR in Latin America and the Caribbean than in sub-Saharan Africa, but sub-Saharan Africa had a higher median BCR. An important comparison where the two measures agree is that middle-income countries as a group had the highest rates of returns to agricultural R&D followed by the high- and then low-income country groups.

Table 7 reports the same results in terms of the MIRRs that have been imputed with a standardised discount rate (10 per cent) and time to maturity of 30 years (i.e. $T_e = 30$). Again, the median categorical rankings for the MIRRs are identical to the median categorical rankings for the BCRs, but not for the IRRs. The median rate of return for all studies based on the MIRR was 17.0 per cent per year. For crop investments, the median MIRR ranges from 18.1 per cent per year for maize to 15.6 per cent per year for roots and tubers and millet and sorghum. The median MIRR for livestock investments (17.3 per cent per year) was higher than that for crops overall, but not by much. The median MIRR for developed country investments was 15.8 per cent per year, which was smaller than for developing countries. Among developing countries, the median MIRR was highest at 18.5 per cent per year for Asian and Pacific countries. The median MIRR for middle-income countries was 17.2 per cent per year, which was about 1.5 percentage points higher than for high- and low-income countries.

6. Conclusion

Zvi Griliches' seminal analysis of hybrid corn spawned a large literature seeking to quantify and demonstrate the value of agricultural research and development (R&D) investments. While his first inclination was to represent the value of these investments as an annual rate of return in perpetuity (ARP), he also framed the results of his analysis in terms of benefit–cost ratios and, at the behest of Martin J. Bailey, an internal rate of return (IRR). Of the three alternatives, it was the IRR that became predominant, while the ARP essentially disappeared from the literature. Unlike the ARP, the BCR did not disappear from the literature with its use occurring about a quarter as often as the IRR.

Even though the IRR became the predominant measure of the returns to investments in agricultural R&D, it has had its critics, including Griliches.

This is potentially important because metrics like the IRR and BCR need not provide consistent conclusions when comparing the value of different investments. These potential inconsistencies are well referenced in most textbooks dealing with capital budgeting, as is the warning against using the IRR to compare among alternative investments. But are these warnings of practical concern making the debate among critics and proponents merely an academic exercise? For example, while textbooks warn that the IRR need not be unique, this concern has been of little practical consequence in the returns-to-research literature because there are no examples of nonunique IRRs in the 2,418 IRR estimates found in the InSTePP returns-to-research (RtR) database (version 3).

The primary objective of this article was to assess the practical consequences of relying on the IRR as a metric for measuring the rates of return to agricultural R&D. We accomplished this objective by comparing the rankings of agricultural R&D projects in the InSTePP RtR database based on the IRR and BCR for those projects that reported both. This comparison revealed the two measures produce vastly different results. One explanation of the differences is that the BCRs reported in the RtR database did not always use consistent discount rate assumptions. Recalculating the BCRs with consistent discount rate assumptions made the BCR and IRR rankings more consistent, particularly when the BCRs were calculated with a high 20 per cent discount rate. With a more commonly used 10 per cent discount rate, substantial differences between the IRR and MIRR rankings remained.

Given these substantial differences, a secondary objective of these articles was to impute BCR estimates from the IRR estimates for as many agricultural R&D evaluations as possible, so this literature could be reconsidered through the lens of the BCR. This exercise resulted in different relative assessments of the returns to different categories of crop research.[7] It also led to different results when comparing the value of investments across different developing and developed country groupings. These results demonstrate that the debate over the most appropriate metrics for weighing the value of investments in agricultural R&D is more than just an academic exercise. Analysts need to carefully consider the assumptions used to construct different metrics of the returns to research as well as the appropriate interpretation of those metrics. The choice of metric is also critical for clearly communicating to policymakers and others who draw on returns-to-research estimates when allocating resources to agricultural R&D. With the ability to evaluate project profitability with both the IRR and BCR, textbook warnings against comparing investments using the IRR, and large practical differences

[7] Moreover, using the imputed BCRs and relevant research costs makes it possible to assess the magnitude of the PVBs associated with the respective projects being evaluated, which has additional informational value when assessing a portfolio of research projects (see, e.g., Alston *et al.* 2020).

between IRR and BCR rankings of past agricultural R&D investments, reasonable care dictates analysts and policymakers should avoid using the IRR.

Data availability statement

Author elects to not share data.

References

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, NY.

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

Alston, J.M., Marra, M.C., Pardey, P.G. and Wyatt, T.J. (2000b). Research returns redux: a meta-analysis of the returns to agricultural R&D, *Australian Journal of Agricultural and Resource Economics* 44, 185–215.

Alston, J.M., Andersen, M.A., James, J.S. and Pardey, P.G. (2011). The economic returns to U.S. Public Agricultural Research, *American Journal of Agricultural Economics* 93, 1,257–1,277.

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

Alston, J.M., Pardey, P.G. and Rao, X. (2020). *The Payoff to Investing in CGIAR Research*. SoAR Foundation, Washington, DC.

Athanasopoulos, P.J. (1978). A note on the modified internal rate of return and investment criterion, *Engineering Economist* 23, 131–133.

Bailey, M.J. (1959). Formal criteria for investment decisions, *Journal of Political Economy* 67, 476–488.

Bennouna, K., Meredith, G.G. and Marchant, T. (2010). Improved capital budgeting decision making: evidence from Canada, *Management Decision* 48, 225–247.

Burns, R.M. and Walker, J. (1997). Investment techniques among the Fortune 500: a rationale approach, *Managerial Finance* 23, 3–15.

Daunfeldt, S. and Hartwig, F. (2014). What determines the use of capital budgeting methods? Evidence from Swedish listed companies, *Journal of Finance and Economics* 2, 101–112.

Deloitte Centre for Health Solutions (2014). *Measuring the Return on Pharmaceutical Innovation 2014: Turning a Corner?*. London, UK: " Deloitte LLP (United Kingdom).

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. 2. Assessing the Impact of Agricultural Research*. International Service for National Agricultural Research, The Hague. 1–31.

European Commission (2005). *The Returns to various Types of Investment in Education and Training*. Final Report to Directorate General Education and Culture. European Commission, Brussels. Available from URL: https://londoneconomics.co.uk/wp-content/uploads/2011/09/82-Study-on-the-returns-to-various-types-of-investment-in-education-and-training.pdf [accessed 9 July 2020].

Evenson, R.E., Waggoner, P.E. and Ruttan, V.W. (1979). Economic benefits from research: an example from agriculture, *Science* 205, 1,101–1,107.

Glover, M., Buxton, M., Guthrie, S., Hanney, S., Pollitt, A. and Grant, J. (2014). Estimating the returns to UK publicly funded cancer-related research in terms of the net value of improved health outcomes, *BMC Medicine* 12, 99.

Graham, J. and Harvey, C. (2001). The theory and practice of corporate finance: evidence from the field, *Journal of Financial Economics* 60, 187–243.

Griliches, Z. (1958). Research costs and social returns: hybrid corn and related innovations, *Journal of Political Economy* 66, 419–431.

Hall, B.H., Mairesse, J. and Mohnen, P. 2010. Measuring the returns to R&D, chapter 24 in Hall, B. and Rosenberg, N. (eds), *Handbook of the Economics of Innovation*, Vol. 2. Elsevier, Amsterdam. 1,033–1,082.

Heckman, J., Lochner, L.J. and Todd, P.E. (2006). Earnings functions, rates of return and treatment effects: the mincer equation and beyond, Chapter 7 in *Handbook on the Economics of Education*, Vol. 1. Elsevier, Amsterdam. 307–458.

HERG (Health Economics Research Group), OHE (Office of Health Economics) and RAND Europe (2008). *Medical Research: What's it Worth? Estimating the Economic Benefits from Medical Research in the UK*. UK Evaluation Forum, London. Available from URL: www.mrc.ac.uk/publications/browse/medical-research-whats-it-worth/ [accessed 24 September 2020].

Hirshleifer, J. (1958). On the theory of optimal investment decision, *Journal of Political Economy* 66, 329–352.

Huber, P.J. (1973). Robust regression: asymptotics, conjectures and Monte Carlo, *The Annals of Statistics* 1, 799–821.

Hurley, T.M., Rao, X. and Pardey, P.G. (2014). Re-examining the reported rates of return to food and agricultural research and development, *American Journal of Agricultural Economics* 96, 1,492–1,504.

Hurley, T.M., Rao, X. and Pardey, P.G. (2017). Re-examining the reported rates of return to food and agricultural research and development: reply, *American Journal of Agricultural Economics* 99, 827–836.

Kierulff, H. (2008). MIRR: a better measure, *Business Horizons* 51, 321–329.

Lin, S.A.Y. (1976). The modified internal rate of return and investment criterion, *The Engineering Economist* 21, 237–247.

Negrete, G.L. (1978). The modified internal rate of return and investment criterion, a reply, *Engineering Economist* 23, 133–134.

Oehmke, J. (2017). Re-examining the reported rates of return to food and agricultural research and development: comment, *American Journal of Agricultural Economics* 99, 818–826.

Pardey, P.G., Andrade, R., Rao, X. and Hurley, T.M. (2016). Documentation: InSTePP Returns to Research (RtR) database. Available from URL: https://www.instepp.umn.edu/sites/instepp.umn.edu/files/product/downloadable/Pardey%20et%20al%202016%20–%20InSTePP%20RTR%20v3.0%20documentation%2826OCT2016%29_2.pdf [accessed 13 October 2016].

Rao, X., Hurley, T.M. and Pardey, P.G. (2019). Are Agricultural R&D returns declining and development dependent?, *World Development* 122, 27–37.

Robison, L.J., Barry, P.J. and Myers, R.J. (2015). Consistent IRR and NPV rankings, *Agricultural Finance Review* 75, 499–513.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection*. John Wiley and Sons, New York, NY.

Ryan, P.A. and Ryan, G.P. (2002). Capital budgeting practices of the Fortune 1000: how have things changed?, *Journal of Business and Management* 8, 355–364.

Truong, G., Partington, G. and Peat, M. (2008). Cost-of-capital estimation and capital-budgeting practice in Australia, *Australian Journal of Management* 33, 95–121.

Yohai, V. (1987). High breakdown-point and high efficiency estimates for regression, *The Annals of Statistics* 15, 642–665.