# THE ECONOMIC JOURNAL

*SEPTEMBER* 1987

*The Economic Journal*, 97 (*September* 1987), 581–595
*Printed in Great Britain*


# THE SIMPLE ECONOMICS OF RESEARCH PORTFOLIOS*


*Partha Dasgupta and Eric Maskin*


## I. INTRODUCTION

### I.1. *The Issues*

One of the most notable discoveries in the sociology of science has been the observation that similar scientific advances are often made by several individuals and research teams working independently of one another (see Ogburn and Thomas (1922), and Merton (1961, 1963, 1973)). Robert Merton, who has examined the phenomenon at length has christened such occurrences, 'multiples.'[1] It is not difficult to provide an explanation for them. Much scientific and technological knowledge – the input for further advances – is in the public domain. Not only are many scientific problems on the public research agenda, many of the possible strategies (or avenues) for attacking them are publicly known. This is patently true in what one often calls the pure sciences, where the social ethos is geared to making (final) discoveries public. But it is also true – albeit to a lesser extent – in the applied sciences (or what is often called technology) even though much knowledge in this domain remains proprietary to individual laboratories and firms (see, e.g., Nelson (1982)). To put it another way, if certain scientific and technological problems are 'in the air,' so are their solutions in the air. Their time has come.[2]

These observations suggest that there are at least five interrelated questions in the economics of science and technology: (1) What problems ought to be on the agenda? (2) How many and what kinds of research projects (or research


* Research towards this article was conducted while the authors were visitors to, respectively, the Center for Economic Policy Research and the Institute for Mathematical Studies in the Social Sciences, Stanford University, during the summer of 1984. We are most grateful to these institutions for making our visits possible. Financial support was provided by the Center for Economic Policy Research at Stanford University and the US National Science Foundation, respectively. This revised version was made possible by a grant from the UK Economic and Social Research Council. We have benefited from discussions with Sudipto Bhattacharya, Zvi Griliches, Dilip Mookherjee, Ron Smith, Joseph Stiglitz, and, especially, from the comments of Steven Lippman. Our debt to the writings of Paul David and Nathan Rosenberg will be evident.


[1] The term is ambiguous. It is used by Merton to denote the occurrence of more than one research unit making what is essentially the same discovery or invention, and not that of a given research unit making more than one discovery.

[2] Reasons for this in technology are varied. The diffusion of knowledge occurs because some of the knowledge *is* pure science and is often generated in public entities such as universities. Moreover, the

strategies) ought to be pursued in tackling them? (3) How ought resources to be allocated among the chosen research projects? (4) Who ought to be conducting the research? and (5) How ought research personnel to be compensated?

We have mentioned the normative questions here. One may similarly pose their positive versions; that is, questions about how the 'market' resolves these issues (see below).

In the recent theoretical literature on the microeconomics of technological change (1) and (3) – and their positive counterparts – have been much discussed (e.g., Loury (1979) and Dasgupta and Stiglitz (1980*a*, *b*)). Question (2) has been addressed partially (e.g., Loury (1979) and Dasgupta and Stiglitz (1980*a*, *b*)), but under the unsatisfactory assumption that the risks associated with the available set of research strategies are independent of one another.[3] Question (5) has been carefully considered in recent years, for example by Lazear and Rosen (1981) and Nalebuff and Stiglitz (1983). Question (4) has for the most part been ignored.

Because the outcome of any research project is uncertain, it is generally in society's interest to hold a portfolio of active projects on any scientific or technological problem worth pursuing. That is, parallelism need not imply waste.[4] (Unless, of course, increasing returns to scale outweigh the benefits from diversification, in which case it would be desirable to pursue only one project.) Given this case for parallel research, society should be prepared to tolerate multiples. If scientists of similar calibre pursue similar research projects – that is, projects whose risks are highly correlated – duplication of success is precisely what we would expect, even when the number of competing research units is small.

Because of limitations in data, historians of science do not usually report on failures in the race for scientific and technological advances. The multiples Merton talks of are duplication in *discoveries* and *inventions*. But for every recorded success, or achievement, there presumably have been failures, where scientists independently attempted to solve the same problem but did not succeed, or at best solved a lesser problem.[5] One can thus argue that the

movement of scientific personnel across firms is conducive to their diffusion. Then again, the patenting mechanism itself divulges knowledge. Finally, it should be noted that to know that someone has solved a problem is to know a great deal: specifically, that the problem is solvable. For example, in their history of the transistor, Braun and MacDonald (1978, p. 52) argue that research on semiconductors was sufficiently far advanced in many places by the end of 1947 so that '...from the mere knowledge that such a thing as a transistor was possible, there were perhaps twenty-five organizations which could have made one.' This of course creates serious problems and tensions within research units in technology. For an analysis of such problems and the manner in which they are resolved, and for an analytical basis for distinguishing science from technology, as social institutions, see Dasgupta and David (1987).

[3] Dasgupta and Stiglitz (1980*a*, *b*) have also analysed the implications of one alternative hypothesis: that the available set of research strategies are perfectly (and positively) correlated with one another. But this is the other extreme and is equally unsatisfactory.

[4] The degree of parallelism can be influenced by direct government participation (e.g., through public funds) or by manipulating the reward system (for example, by choice of the length of patents; see Nordhaus (1969)).

[5] In basic research there is on occasion only one 'right' way of going about things, the problem being that one does not know in advance what the right way is. In such a case correlation among research projects may well be negative.

occurrence of multiples understates – although not necessarily severely – the duplication of research *effort* (see e.g., Kuznets (1962)).

### I.2. *The Problem*

In this article we will compare efficient portfolios of research projects with portfolios that would emerge under competition among rival research units – what we will call 'market' portfolios. Specifically, we will examine three issues: (*a*) whether competition encourages too rapid a rate of technological advance (Section IV, see Proposition 4); (*b*) whether it encourages the rivals to undertake excessively risky research projects (Section V, see Proposition 5); and (*c*) whether it encourages them to choose overly similar (i.e., correlated) projects (Section III, see Propositions 2 and 3). To state the motivations behind these questions in words: affirmative answers (and we shall argue below that the answers *are* affirmative) would imply, respectively, that the market induces, on average, the development of too many inventions, too large a spread in the distribution of the quality of these inventions, and an excessive occurrence of Merton's multiples. We are thus concerned here with parallel research, an aspect of question (2) above.

It is as well to ask first why one might expect market portfolios to be inefficient. There are in fact several reasons (see footnote 9 on p. 585 below). In this article we will devote our attention exclusively to one reason, distinctive of the economics of science and technology. It is that of the discoveries (or inventions) made by rival research units involved in parallel research only the best is worthwhile to society (see equation (3) below). For instance, among available ways of manufacturing a commodity society wants to use only the best-practice technique. To take another example, there is no value added when the same discovery is made a second, third, or fourth time.[6] To put it sharply (and thus somewhat inaccurately), the winning research unit is the sole contributor to social surplus.

All this is to state the matter in an extreme form to highlight a distinctive feature of knowledge viewed as a commodity. The point here is that the same piece of knowledge can be put to use over and over again without any 'wear and tear' (Arrow, 1962). To be sure, discoverers involved in 'multiples' rarely discover the *same* thing. There are always qualitative differences, even if ever so small; for example, in their exact timing of discovery, methods of proof in the case of mathematical theorems, characteristics of the product when a new commodity is invented, technology of production if the discovery is a process invention, and so on. Despite such differences the fact is that parallel research involves goal-oriented research where researchers search for the same type of thing. There are thus only quality variations in the outcomes of such research, and society gains a disproportionate amount from one of the realised qualities.

The above observation concerns the manner in which *society* values the outcomes of parallel research. Turning next to the way in which research units are compensated in the 'market,' we find that the compensation schemes are

[6] By this we do not of course mean independent *confirmations* of a scientific discovery, which is a different matter altogether.

approximately of the form: 'the winner takes all' (see equations ($6a$, $b$) below). The institution of patents obviously mimics this rather ruthless mode of compensation. But even in the absence of patents the first firm to develop a product often makes great inroads into the market and thus reaps a large share of the rents from the invention. Here too then the winner of the race earns most of the profits that are to be earned. The point thus is not that inferior inventions earn nothing, merely that they earn a disproportionately low amount; that is why they are called inferior. The winner-takes-all hypothesis is merely an idealisation.[7]

These remarks about market compensation for research and development effort apply to technology. However, it would appear that in the pure sciences as well such a ruthless compensation scheme operates. Priority matters greatly to scientists, and this concern is continuously fostered by the scientific community (see Merton (1973) and Boorstin (1983), Chapter 53). In science priority *is* the prize. It awards 'moral possession' to the scientist (Medawar (1982), p. 260) even though legal possession may not be possible, nor indeed desired by any party. The reason why in scientific races the winner(s) is awarded the entire 'purse' unlike tennis tournaments, is not hard to find: it is often impossible to discover subsequently how far behind the winner(s) the losers of the race were at the time the race was completed.[8] But not always. Writing about the race for the discovery of the structure of DNA, Medawar (1968) asserts,

> If Watson and Crick had not made it, someone else would certainly have done so – almost certainly Linus Pauling, and almost certainly very soon. It would have been that same discovery too; nothing else could take its place. (Medawar, 1982, p. 272).

This leads us to the reason for market failure of concern here: the non-congruence of social goals and private aspirations occasioned by the 'winner–takes–all' compensation scheme. The point is that society does not care *who* is successful in solving a given scientific or technological problem, it cares that the problem is solved. But for the individual scientist (or the research team)

[7] There are occasions when being a late arrival (or developer) is an advantage, as was argued forcefully by Thorstein Veblen. The critical question to ask here is whether the participants in the 'race' knew in advance that being first would be disadvantageous. If they did, then the phenomenon under study would not have been a race but a war of attrition, which requires a somewhat different analysis. If they did not, then presumably unforeseen things happened; that is, the participants did not know the 'game' they were engaged in. The problem we are analyzing in this paper is neither of these.

In fact there is a deeper point to be made about those innovations for which late arrival was subsequently seen as advantageous. The key question to ask concerns the *vantage* point of the evaluation of relative advantages. Thus, broad gauges (Veblen's example), which were installed much later, were economically superior to narrow gauges. Veblen's argument was that the country which had earlier installed the narrow gauge suffered in competition with those which subsequently innovated and installed the broad gauge. But during the period *between* the two innovations the early innovator enjoyed great benefits. Before either innovation an early innovation with the narrow gauge might well have looked good even if it had been known then that the broad gauge would subsequently come to dominate. Indeed, with a high enough discount rate it *would* have looked good.

[8] Bell (1937) has an illuminating, indirect discussion of the role played by the institution of mathematical contests in the growth of science in the seventeenth and eighteenth centuries. These contests were for the most part initiated by scientific bodies, such as the French Academy.

the identity of the problem-solver matters greatly: *he* wants to be the successful one! Thus suppose there are two research units and two possible outcomes of each unit's research: success ($S$) or failure ($F$). Writing as ($S$, $F$) the event where the first research unit succeeds and the second fails, and so on for the remaining combinations, and ignoring research and development (R&D) costs, society clearly is indifferent between ($S$, $S$), ($S$, $F$) and ($F$, $S$). But when the market awards 'all' to the winner, neither competitor will be indifferent between them. (For example, the first competitor will prefer ($S$, $F$) to ($F$, $S$) and indeed, ($S$, $F$) to ($S$, $S$)). This non-congruence between private and social rankings of final outcomes is a source of distortion and a reason why market portfolios are likely to be inefficient.[9]

## I.3. *The Approach*

We will view a research project as a real-valued random variable whose realisations are (scalar) indices of the output (or outcome) of research. Except for the two-point example in Section III.2 we will consider continuous distributions. In Section II the basic model will be presented and conditions describing, respectively, efficient and market portfolios will be derived. In Sections III, IV and V the three sorts of biases we mentioned earlier will be derived. Section VI summarises the main conclusions.

Given the structure of both social and private rewards from R & D competition, our analysis necessarily concerns the statistics of extreme values (see expressions (3) and (4*a*)–(4*b*)). Thus, in particular, the portfolio problems we will be analysing would be nonconcave were care not taken to restrict the available class of research projects. Except for the example in Section III. 2, we will, for expositional ease, tailor the R & D *cost schedules* in such a way as to concavify the portfolio problems. If the reader finds our cost assumptions unduly restrictive he can ignore them, but should then interpret our analysis as *local*. That is, without concavity, the distortions we identify in market portfolios are local biases: useful information for policy reform, but not necessarily enlightening about the nature of the global optimum.

[9] There are at least five other reasons why market portfolios would generally be inefficient: (i) the phenomena of moral hazard and adverse selection are acutely present in research and development; (ii) the winner of the scientific competition does not collect the social surplus; (iii) free-entry into the race for the prize might lead to the wrong number of parallel research teams because rents are dissipated (Dasgupta and Stiglitz, 1980*b*); (iv) private attitudes to risk differ from 'society's attitude to risk; and (v) at least one of the participants faces a nonconcave maximisation problem. Moral hazard has been much discussed in the principal-agent literature and is particularly relevant in answering question (5) in the text. Adverse selection is the phenomenon under study in question (4) in the text. In what follows we avoid (i) by assuming that effort can be monitored and that the characteristics of the research units are common knowledge. We will avoid (ii) by assuming that the winner in the market is awarded the entire social surplus. We will block the avenue opened by (iii) by keeping the number of research units fixed when comparing market portfolios of R & D projects with socially efficient ones. We will (except for Section III.2), avoid (iv) by postulating away differences in risk attitudes. And finally, we will bypass (v) by so choosing cost functions that the maximisation problems are concave (see below). What remains as a possible source of distortion is the feature discussed in the text above; for, recall from the work of Marschak and Radner (1972 p. 153) that for market portfolios to be efficient one in general needs social and private goals to be identical and the common valuation function to be strictly concave (and differentiable everywhere).

## II. THE MODEL

### II.1. *The Setting*

There are two research units (or firms), labelled $i = 1, 2$. Each chooses a research project from a continuum of projects labelled by numbers in the unit interval $[0, 1]$. Let $a_i \in [0, 1]$ denote the project selected by $i$ and let $C_i(a_i)$ be the cost of the project. Given $a_1$ and $a_2$ there is a joint probability distribution of the outputs of the two research projects. Denoting by $x_1$ and $x_2$ the research outputs of the two firms, we assume that they are non-negative real numbers with common support, $[0, \bar{x}]$, and that the joint probability distribution is given by the density function $g(x_1, x_2; a_1, a_2)$. (Thus $a_1$ and $a_2$ are parameters of the distribution.) The variable $x_i$ can be interpreted as the quality of the product developed by the research unit. Alternatively, we can think of it as the inverse of the time required for development.

From $g$ we define the marginal distributions as:

$$f_1(x_1; a_1) \equiv \int_0^{\bar{x}} g(x_1, x_2; a_1, a_2)\, dx_2 \tag{1a}$$

and

$$f_2(x_2; a_2) \equiv \int_0^{\bar{x}} g(x_1, x_2; a_1, a_2)\, dx_1. \tag{1b}$$

Thus we have assumed in (1 $a$) and (1 b) that there is no technical interference in the R & D processes of the research units: the marginal distribution of the random variable $\tilde{x}_1$ (resp. $\tilde{x}_2$) depends solely on $a_1$ (resp. $a_2$).

To present our arguments in a sharp form it will be useful to suppose that the firms are symmetric. In fact nothing essential will be lost in our making this assumption. It will allow us as well to economise in notation. Formally, we have

$$g(x_1, x_2; a_1, a_2) = g(x_2, x_1; 1-a_2, 1-a_1) \quad \text{and} \quad C_1(a) = C_2(a) = C(a)$$
$$\text{for all} \quad x_1, x_2 \geqslant 0 \quad \text{and} \quad \text{for all} \quad a, a_1, a_2 \in [0, 1]. \tag{2}$$

We are interested, moreover, in symmetric solutions to the efficient portfolio problem and in symmetric market equilibria. This means in particular that the two research teams locate on different sides of the mid-point, $\frac{1}{2}$, of the space of research projects, unless they both happen to choose $a_1 = a_2 = \frac{1}{2}$. Thus we may as well restrict firm 2 to choosing $a_2$ from the range $[0, \frac{1}{2}]$ and firm 1 to choosing $a_1$ from the range $[\frac{1}{2}, 1]$. In other words, the assumption that the two firms have access to the same set of research projects – i.e. all scientific knowledge is publicly known – is not necessary.

In Section III we impose the following restriction on the class of available R & D projects: we will suppose that the closer are $a_1$ and $a_2$ the more highly 'positively correlated' are the research strategies (conditions (10$a$) and (10$b$) below). For vividness the reader may wish to suppose (although this interpretation is not necessary) that the projects are perfectly (and positively) correlated if $a_1 = a_2 = \frac{1}{2}$ and are independent (or even negatively correlated) if $a_1 = 1$ and $a_2 = 0$.

In Section IV we are concerned with the rate of technological progress. Here it is natural to interpret $x_i$ as the inverse of development time. We suppose that varying $a_i$ changes the distribution of $x_i$ in a first-order stochastically dominating way.

In Section V, by contrast, the emphasis is the degree of risk-taking in R & D portfolios. Hence, we suppose here that $a_i$ does not affect the mean of $x_i$ but rather determines how variable the distribution is, i.e. varying $a_i$ affects the distribution according to second-order stochastic dominance.

To focus attention on those features of R & D competition due to its particular reward structure, it is simplest to ignore risk-aversion. Thus, we will assume, except in Section III. 2, that all decision-makers are risk neutral. But risk neutrality implies we are apt to obtain corner solutions to both the market and socially planned portfolio problems. Specifically, in Section III, increasing $a_1$ (decrease $a_2$) reduces the correlation between firm 1's and 2's outcomes. If the only effect of increasing $a_1$ is to reduce the correlation – i.e. if the *marginal* distribution of $x_1$ is independent of $a_1$ – then, in the absence of risk aversion, such a change is an unequivocal improvement both privately and socially. Thus, without a counterbalancing force, both market equilibrium and the social optimum will place firms 1 and 2 at $a_1 = 1$ and $a_2 = 0$ (c.f., Bhattacharya and Mookherjee (1986)). Similarly, in Section IV, increasing $a_1$ (decreasing $a_2$) induces a first-order stochastic improvement in firm 1's (2's) distribution, and again we are in danger of finding the firms on opposite extremes of the spectrum.

To induce interior solutions – and, therefore, to permit a genuine trade off between costs and benefits – we shall assume that a firm's costs rise as its research strategy deviates from the midpoint, $\frac{1}{2}$.[10] Specifically, we shall suppose that the cost function $C(a)$ is symmetric about the point $\frac{1}{2}$ and strictly convex.

## II.2. *Efficient R & D Portfolios*

If $x_1$ and $x_2$ are the research outputs of the two teams, the social payoff is $\max\{x_1, x_2\}$. Because society is assumed to be risk-neutral, choice of the pair of projects $a_1$ and $a_2$ yields the expected net social surplus (or benefit):

$$\int_0^{\bar{x}}\int_0^{\bar{x}} \max\{x_1, x_2\}\, g(x_1, x_2; a_1, a_2)\, dx_1\, dx_2 - C(a_1) - C(a_2). \tag{3}$$

An efficient portfolio is a pair of research projects, $(a_1^s, a_2^s)$, which maximises (3) subject to the constraints $0 \leqslant a_2 \leqslant \frac{1}{2}$ and $\frac{1}{2} \leqslant a_1 \leqslant 1$.

In what follows we will assume that $g(.)$ and $C(.)$ are such that the solution is interior and defined by the first-order conditions of the maximisation. (Indeed, this will be so if $C(.)$ is sufficiently convex.) These first-order conditions are:

$$\int_0^{\bar{x}}\int_0^{\bar{x}} \max\{x_1, x_2\}\,[\partial g(x_1, x_2; a_1, a_2)/\partial a_1]\, dx_1\, dx_2 = dC(a_1)/da_1. \tag{4a}$$

and

$$\int_0^{\bar{x}}\int_0^{\bar{x}} \max\{x_1, x_2\}\,[\partial g(x_1, x_2; a_1, a_2)/\partial a_2]\, dx_1\, dx_2 = dC(a_2)/da_2. \tag{4b}$$

[10] The intuitive idea behind the assumption is that, the farther away is the project from $\frac{1}{2}$ the more 'unusual' is the research strategy and thus the more costly in terms of materials and so forth.

Thus $(a_1^s, a_2^s)$ is a solution of equations $(4a)$ and $(4b)$. Since we are analysing a symmetric solution, we have $0 < a_2^s = 1 - a_1^s \leq \frac{1}{2}$.

### II.3. *Market Portfolios*

We will assume that $g(.)$ and $C(.)$ have forms that result in a symmetric equilibrium. Since correlation is then never perfect (except possibly at $a_1 = a_2 = \frac{1}{2}$, which we will avoid as a possible outcome) the probability of a tie is nil. We suppose, as discussed in Section I.2, that the winner takes all and that the compensation to the winner is the social payoff. Each firm is risk neutral. Thus, if $a_1$ and $a_2$ are the two chosen research projects, firm 1's expected net benefit is

$$\int_0^{\bar{x}} \int_{x_2}^{\bar{x}} x_1 g(x_1, x_2; a_1, a_2)\, dx_1\, dx_2 - C(a_1), \tag{5a}$$

and firm 2's expected net benefit is

$$\int_0^{\bar{x}} \int_{x_1}^{\bar{x}} x_2 g(x_1, x_2; a_1, a_2)\, dx_2\, dx_1 - C(a_2). \tag{5b}$$

Firms choose their projects simultaneously and with complete information about their rival's benefit and cost functions. A Nash equilibrium, $(a_1^*, a_2^*)$, therefore describes a market outcome. Assuming the equilibrium to be interior, we note that it must be a solution of the two individual first-order conditions:

$$\int_0^{\bar{x}} \int_{x_2}^{\bar{x}} x_1[\partial g(x_1, x_2; a_1, a_2)/\partial a_1]\, dx_1\, dx_2 = dC(a_1)/da_1. \tag{6a}$$

$$\int_0^{\bar{x}} \int_{x_1}^{\bar{x}} x_2[\partial g(x_1, x_2; a_1, a_2)/\partial a_2]\, dx_2\, dx_1 = dC(a_2)/da_2. \tag{6b}$$

We are interested in a symmetric equilibrium. Thus $0 < a_2^* = 1 - a_1^* \leq \frac{1}{2}$.

### II.4. *The Bias*

We wish to analyse the nature of the bias in the market portfolio. Because we are restricting attention to symmetric solutions we may as well compare equations $(4a)$ and $(6a)$ only. (Comparison of equations $(4b)$ and $(6b)$ follows analogously.) We turn to equation $(4a)$. Note first that

$$\max\{x_1, x_2\} = x_2 + \max\{(x_1 - x_2), 0\}. \tag{7}$$

Using condition $(1b)$ and equation (7) in equation $(4a)$ yields

$$\int_0^{\bar{x}} \int_{x_2}^{\bar{x}} (x_1 - x_2)\, [\partial g(x_1, x_2; a_1, a_2)/\partial a_1]\, dx_1\, dx_2 = dC(a_1)/da_1. \tag{8}$$

One may now compare equations $(6a)$ and (8) to conclude that what determines the deviation of the market equilibrium from social efficiency is the term

$$-\int_0^{\bar{x}} \int_{x_2}^{\bar{x}} x_2[\partial g(x_1, x_2; a_1, a_2)/\partial a_1]\, dx_1\, dx_2. \tag{9a}$$

Expression (9 $a$) is marginal net social surplus (social minus private benefit) with respect to $a_1$ when evaluated at $a_1 = a_1^*$. If it is negative, then (at least in a local sense) the market solution, $a_1^*$, is too large; that is, $a_1^* > a_1^s$. By the same token, if (9 $a$) is positive, then $a_1^*$ is too small, that is, $a_1^* < a_1^s$. We thus obtain

PROPOSITION 1. *If expression (9 a) is negative (positive) at* $(a_1^*, 1-a_1^*)$, *then the market portfolio consists of projects that are unduly far apart (close together). The opposite conclusion follows if*

$$-\int_0^{\bar{x}}\int_{x_1}^{\bar{x}} x_1 \left[\partial g(x_1, x_2; a_1, a_2)/\partial a_2\right] dx_2\, dx_1 \tag{9b}$$

*is negative (positive).*

In the next three sections we will investigate the sign of expressions (9 $a$) and (9 $b$) under different hypotheses about the set of feasible R & D projects.

## III. CORRELATION

### III.1. *Continuous Distributions*

We assume in this section that, the closer are $a_1$ and $a_2$, the more highly 'positively correlated' are the two corresponding research projects. One way of formalising this assumption is to suppose that as $a_2$ moves toward $a_1$, the increase in the probability of a pair $(x_1, x_2)$ is biggest if $x_1 = x_2$ and declines monotonically as the difference between $x_1$ and $x_2$ grows. That is, we can write $(\partial/\partial a_2)\, g(x_1, x_2; a_1, a_2)$ as

$$\frac{\partial}{\partial a_2} g(x_1, x_2; a_1, a_2) = h(|x_1 - x_2|, x_1; a_1, a_2), \tag{10a}$$

where $h$ is decreasing in its first argument. From symmetry, we have

$$\frac{\partial}{\partial a_1} g(x_1, x_2; a_1, a_2) = -h(|x_1 - x_2|, x_2; 1-a_2, 1-a_1). \tag{10b}$$

We shall also assume that for any pair $(a_1, a_2)$, the distribution of $(x_1, x_2)$ is symmetric in the sense that

$$g(x_1, x_2; a_1, a_2) = g(\bar{x}-x_1, \bar{x}-x_2; a_1, a_2). \tag{11}$$

PROPOSITION 2. *If g satisfies conditions* (10 $a$), (10 $b$), *and* (11), *then the market portfolio is characterised by excessive correlation among research projects.*

*Proof.* Consider $x_1 \geqslant \bar{x}/2$. From (1 $a$) we have

$$\begin{aligned} 0 = \frac{\partial f_1}{\partial a_2}(x_1; a_1) = & \int_0^{2x_1-\bar{x}} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2 \\ & + \int_{2x_1-\bar{x}}^{x_1} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2 \\ & + \int_{x_1}^{\bar{x}} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2. \end{aligned} \tag{12}$$

From (10 $a$),

$$\frac{\partial g}{\partial a_2}(x_1, x_1 - b; a_1, a_2) = \frac{\partial g}{\partial a_2}(x_1, x_1 + b; a_1, a_2) \text{ for all } b.$$

Hence, the second and third terms of the right side of (12) must be equal. Now, for $x_2 < 2x_1 - \bar{x}$, the absolute value, $|x_1 - x_2|$, is greater than that for $x_2$ between $x_1$ and $\bar{x}$. Hence, from (10a), the integrand in the first term is less than that of the third term. We conclude, from (12) and the fact that the second and third terms are equal, that

$$\int_{x_1}^{\bar{x}} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2 > 0, \quad \text{for} \quad x_1 \geqslant \frac{\bar{x}}{2}. \tag{13}$$

Now,

$$\int_{x_1}^{\bar{x}} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2 = \int_0^{\bar{x}-x_1} \frac{\partial g}{\partial a_2}(\bar{x}-x_1, x_2; a_1, a_2)\, dx_2$$

$$= -\int_{\bar{x}-x_1}^{\bar{x}} \frac{\partial g}{\partial a_2}(\bar{x}-x_1, x_2; a_1, a_2)\, dx_2, \quad \text{for} \quad x_1 \geqslant \frac{\bar{x}}{2}, \tag{14}$$

where the first equality follows from (11) and the second from (12). Thus,

$$\int_0^{\bar{x}} x_1 \int_{x_1}^{\bar{x}} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2\, dx_1$$

$$= \int_{\bar{x}/2}^{\bar{x}} [x_1 - (\bar{x} - x_1)] \int_{x_1}^{\bar{x}} \frac{\partial g}{\partial a_2}(x_1, x_2; a_1, a_2)\, dx_2\, dx_1,$$

which, from (13), is positive. But the left-hand side of this last expression is just the negative of (9b). Hence, from Proposition 1 we conclude that, in equilibrium, $a_1$ and $a_2$ are too close together, i.e. there is excessive correlation.

Q.E.D.

The intuition behind this proposition can be conveyed in terms of market externalities. As a firm moves away from its rival in the space of research projects, it bestows a positive externality. Specifically, the likelihood that the rival is successful when the firm in question is not, increases. This is socially desirable, but it is not picked up in the firm's private calculation. There is thus undue similarity of project characteristics in the market portfolio.

### III.2. *Two-Point Distributions*

Let us drop the hypothesis in Proposition 2 that the winner captures the entire social benefit from a discovery. Let us also eliminate risk-neutrality. Without these two assumptions, the general problem does not offer any simple analysis. But two-point distributions do yield answers. For this reason we consider here the important special case when research outcome is either a success ($S$) or a failure ($F$).

Let $p(a_i)$ be the unconditional probability of success for firm $i$ if it chooses project $a_i$ (the unconditional probability of failure is thus $1 - p(a_i)$). Assume that $0 \leqslant p(a_i) \leqslant \frac{1}{2}$ for all $a_i$, and that $p(.)$ is strictly concave with its maximum at $\frac{1}{2}$. We shall continue to impose symmetry. Thus, $p(a) = p(1-a)$ for $0 \leqslant a \leqslant 1$. For simplicity assume in this subsection that all projects cost the same (and therefore that costs can be ignored).

As before, let $0 \leqslant a_2 \leqslant \frac{1}{2}$ and $\frac{1}{2} \leqslant a_1 \leqslant 1$. Then let $[1-(a_1-a_2)]$ be the correlation coefficient between projects $a_1$ and $a_2$. The table below gives the probabilities of the four outcomes: $(S, S)$, $(S, F)$, $(F, S)$ and $(F, F)$.

Table 1

| | $S$ | $F$ |
|---|---|---|
| $S$ | $(a_1-a_2)p(a_1)p(a_2)$ $+(1-a_1+a_2)[p(a_1)+p(a_2)]/2$ | $(1+a_1-a_2)p(a_1)/2$ $-(1-a_1+a_2)p(a_2)/2$ $-(a_1-a_2)p(a_1)p(a_2)$ |
| $F$ | $(1+a_1-a_2)p(a_2)/2$ $-(1-a_1+a_2)p(a_1)/2$ $-(a_1-a_2)p(a_1)p(a_2)$ | $1-(1+a_1-a_2)p(a_1)/2$ $-(1+a_1-a_2)p(a_2)/2$ $+(a_1-a_2)p(a_1)p(a_2)$ |

We first consider symmetric efficient portfolios. Let $B(>0)$ be the social benefit from successful completion of at least one project. (The social benefit from failure of both projects is nil.) It follows that we are to find the pair $(a^s, 1-a^s)$ which maximises

$$B[(1+a_1-a_2)p(a_1)/2-(a_1-a_2)p(a_1)p(a_2)+(1+a_1-a_2)p(a_2)/2].$$

Now, if the solution is interior, the first-order condition (with respect to $a_1$) is

$$\tfrac{1}{2}\{1+(a_1-a_2)[1-2p(a_2)]\}\frac{dp(a_1)}{da_1}-p(a_1)p(a_2)+[p(a_1)+p(a_2)]/2=0. \quad (15)$$

A symmetric, efficient portfolio is a pair of projects $a_1$ and $a_2$ satisfying equation (15) with $a_1 = 1-a_2$.

Turn now to symmetric market portfolios. Each research unit is concerned with its expected net benefits. Let $U_F$ be the unit's net benefit if it fails in its research, $U_{SS}$ if both succeed, and $U_{SF}$ if it succeeds and the rival fails.

We may as well normalise and set $U_F = 0$ and $U_{SS} = 1$. It clearly makes sense to suppose that $U_{SF} > U_{SS}$. It follows that firm 1 maximises

$$(a_1-a_2)p(a_1)p(a_2)+(1-a_1+a_2)[p(a_1)+p(a_2)]/2$$
$$+U_{SF}[p(a_1)(1+a_1-a_2)/2-p(a_2)(1-a_1+a_2)/2-(a_1-a_2)p(a_1)p(a_2)].$$[11]

Firm 1's first-order condition for a Nash equilibrium is then

$$\frac{dp}{da_1}(a_1)\{1+U_{SF}+(a_1-a_2)(U_{SF}-1)[1-2p(a_2)]\}/2$$
$$+(U_{SF}-1)\{[p(a_1)+p(a_2)]/2-p(a_1)p(a_2)\}=0. \quad (16)$$

(The second-order condition for a maximum is satisfied if $U_{SF} \geqslant 2U_{SS}$, i.e. a 'monopoly' is at least as profitable as a 'joint duopoly.') A symmetric market portfolio is a pair of projects $a_1$ and $a_2$ that satisfies equation (16) with $a_1 = 1-a_2$.

[11] Recall that we are assuming that all projects cost the same and so can ignore costs.

Dividing the left-hand side of equation (16) by $(U_{SF}-1)$ and subtracting it from the left-hand side of equation (15) we obtain

$$\frac{dp}{da_1}(a_1)\,[1-(1+U_{SF})/(U_{SF}-1)]/2. \tag{17}$$

Because we are looking at an interior solution, $\frac{1}{2} < a_1 < 1$. Therefore $(dp(a_1)/da_1) < 0$. Thus expression (17) is positive when evaluated at a symmetric Nash equilibrium. It follows that the left-hand side of equation (15) is positive when evaluated at the market portfolio. We therefore deduce

PROPOSITION 3. *If the outcome of research is either success or failure then market research portfolios consist of projects that are too highly correlated.*

## IV. FIRST-ORDER STOCHASTIC DOMINANCE

In this and the following section we are interested in the characteristics of market portfolios when the projects can be ranked in terms of stochastic dominance of some order. So, suppose that, for all $a_1$ and $a_2$, $\tilde{x}_1$ and $\tilde{x}_2$ are *independent* random variables. In this section we will also assume that an increase in $a_1$ (resp. decrease in $a_2$) implies a first-order stochastic improvement in $x_1$ (resp. $x_2$). Writing

$$F_i(x_i;\, a_i) \equiv \int_0^{x_i} f_i(t;\, a_i)\, dt \quad \text{for} \quad i = 1, 2,$$

we assume therefore that

$$\partial F_1(x_1;\, a_1)/\partial a_1 < 0 \quad \text{for all} \quad x_1 \text{ in } (0, \bar{x}) \tag{18a}$$

$$\partial F_2(x_2;\, a_2)/\partial a_2 > 0 \quad \text{for all} \quad x_2 \text{ in } (0, \bar{x}). \tag{18b}$$

Because $\tilde{x}_1$ and $\tilde{x}_2$ are independent, we can rewrite expression (9a) as

$$\int_0^{\bar{x}} x_2[\partial F_1(x_2;\, a_1)/\partial a_1] f_2(x_2;\, a_2)\, dx_2, \tag{19}$$

which from (18a) is negative. But if expression (19) is negative then $a_1^* > a_1^s$ (see Proposition 1); or in other words, the market portfolio dominates the efficient portfolio in the first-order stochastic sense. Recall however that $C(a_1)$ is increasing in $a_1$. Thus at a market equilibrium there is excessive R & D expenditure (if $C(a)$ is independent of $a$ it is easy to see that $a_1^* = a_1^s = 1$ and $a_2^* = a_2^s = 0$). To summarise, we have:

PROPOSITION 4. *In a market equilibrium firms select projects having expected yields that are greater than the expected yields of the efficient portfolio.*

The point here is that the private cost of losing in the R & D race is high. Under market competition, each firm is prompted to spend more than is socially efficient so as to give itself a better chance of being the winner.[12] To put it another way, if there is a good chance that the research performance of a given

[12] For related results, see Loury (1979) and Lee and Wilde (1980).

firm will be exceptionally good – (it has invested heavily in R & D) – there is no *social* advantage in the rival's investing much in R & D. But the rival may well have a private advantage to do so.

## V. SECOND-ORDER STOCHASTIC DOMINANCE

Continue to assume that $\tilde{x}_1$ and $\tilde{x}_2$ are independent random variables for all $a_1$ and $a_2$. Suppose now that an increase in $a_1$ (decrease in $a_2$) implies a mean-preserving spread of the distribution of $x_1(x_2)$. In other words, for all $x_1$ and $x_2$,

$$\partial\left[\int_0^{x_1} F_1(t;\, a_1)\, dt\right]\Big/\partial a_1 \geqslant 0 \tag{20a}$$

and

$$\partial\left[\int_0^{x_2} F_2(t;\, a_2)\, dt\right]\Big/\partial a_2 \leqslant 0, \tag{20b}$$

where the inequalities hold with equality at the endpoints and otherwise hold strictly.

Integrating expression (19) by parts we have

$$-\int_0^{\bar{x}} \{x_2[\partial f_2(x_2;\, a_2)/\partial x_2] + f_2(x_2;\, a_2)\} \int_0^{x_2} [\partial F_1(x_1;\, a_1)/\partial a_1]\, dx_1\, dx_2. \tag{21}$$

Now using (20a) and (21) we conclude that expression (21) is negative if $1 + (x_2/f_2)\,\partial f_2/\partial x_2 > 0$. But if (21) is negative, then $a_1^* > a_1^s$, (Proposition 1). Thus we have

PROPOSITION 5. *If* $1 + (x_i/f_i)\,\partial f_i/\partial x_i > 0$ $(i = 1, 2)$ *firms in market equilibrium undertake excessively risky research projects.*

One way of seeing why Proposition 5 is intuitively plausible is to note that if $f_i(x_i;\, a_i)$ does not fall too rapidly with increasing $x_i$, a firm inflicts an external diseconomy on its rival when it chooses a more risky research project. Suppose, for example, that firm 2 has chosen a project that gives rise to the uniform distribution on the interval [4, 5]. Assume, moreover, that firm 1 can choose between the uniform distribution on [4, 5] and the uniform distribution on [3, 6]. The latter is a mean-preserving spread of the former. Were firm 2 by chance to realise a good result from its project, say the value $x_2 = 5$, it would be a sure winner against firm 1, if firm 1 chose the project [4, 5]. But it would by no means be a necessary winner were firm 1 to choose the project defined on [3, 6]. The argument, of course, cuts the other way as well since if by bad luck firm 2 were to draw the lowest possible value, $x_2 = 4$, it would no longer necessarily be the loser against firm 1 choosing [3, 6]. The crucial point, though, is that the loss incurred by firm 2 at the upper end of its chosen distribution is greater than the gain it enjoys at the lower end when firm 1 chooses the uniform distribution on [3, 6] rather than on [4, 5]. In short, by choosing a more risky research project a firm inflicts an external diseconomy on its rival, and we should expect market equilibrium to be characterised by an excess of this diseconomy. The force of the sufficiency condition, $1+(x_i/f_i)\,\partial f_i/\partial x_i > 0$ is clear. There must be sufficient weight at the upper tail (as there is with uniform distributions) for the marginal losses imposed by a mean preserving spread to outweigh the marginal gains.[13]

## VI. CONCLUSIONS

In this article we have undertaken some simple exercises in the economics of research portfolios. Three questions were raised: (*a*) whether competition among rival research units leads to too high a rate of technological advance; (*b*) whether competition encourages rivals to undertake excessive risks; and (*c*) whether competition induces them to choose overly similar research projects.

R & D portfolios differ markedly from their financial counterparts in that often only one element of the portfolio pays off, i.e., contributes to social surplus. Although society does not care who this winner is, individual firms care very much when payment is of the typical winner-take-all variety. Private aversion to losing the R & D race leads them to select too high an expected rate of technological change (Proposition 4). In a large class of cases, it also induces excessively risky research projects (Proposition 5).

We had no settled prior beliefs about the manner in which the market coordinates rival research units in their choice over the degree of similarity between projects. The fact that 'multiples' are observed is no ground for thinking that the market encourages too much correlation. The correct question to ask is if there are too many multiples on average. Our tentative conclusion is that there are (Propositions 2 and 3).

Taking all things into account, whether the market encourages undue investment in R & D is therefore a question that is unresolved by our analysis. Propositions 4 and 5 identify circumstances where it does and Proposition 2 circumstances where it does not. But the most important moral is probably the one that emerges from Proposition 2: it is that academics are not lone wolves trying as much to differentiate themselves from their rivals; they work on projects which are far too similar!

*University of Cambridge*

*Harvard University*

*Date of receipt of final typescript: January 1987*

## References

Arrow, K. J. (1962). 'Economic welfare and the allocation of resources for invention.' In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, (ed. R. R. Nelson). Princeton University Press.

Bell, E. T. (1937). *Men of Mathematics*. New York: Simon and Schuster.

Bhattacharya, S. and Mookherjee, D. (1986). 'Portfolio choice in research and development.' *Rand Journal of Economics*, vol. 17, pp. 594–605.

Boorstin, D. J. (1983). *The Discoverers: A History of Man's Search to Know His World and Himself*. New York: Random House.

Braun, E., and MacDonald, S. (1978). *Revolution in Miniature: The History and Impact of Semiconductor Electronics*. Cambridge: Cambridge University Press.

[13] For a related result, see Klette and de Meza (1985).

Dasgupta, P. and David, P. (1987). 'Information disclosure and the economics of science and technology.' In *Arrow and the Foundations of the Theory of Economic Policy* (ed. G. Feiwel). London: Macmillan.

—— and Stiglitz, J. E. (1980*a*). 'Industrial structure and the nature of innovative activity.' ECONOMIC JOURNAL, vol. 90, pp. 266–93.

—— and —— (1980*b*). 'Uncertainty, industrial structure and the speed of R & D.' *Bell Journal of Economics*, (spring), pp. 1–28.

Kleete, T. and de Meza, D. (1986). 'Is the market biased against R & D.' *Rand Journal of Economics*, vol. 17 (spring), pp. 113–39.

Kuznets, S. (1962). 'Inventive activity: problems of definition and measurement.' In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, (ed. R. R. Nelson). Princeton University Press.

Lazear, E. and Rosen, S. (1981), 'Rank order tournaments as optimum labor contracts.' *Journal of Political Economy*, vol. 89, pp. 841–64.

Lee, T. and Wilde, L. (1980). 'Market structure and innovation: a reformulation.' *Quarterly Journal of Economics*, 94, pp. 429–36.

Loury, G. (1979). 'Market structure and innovation.' *Quarterly Journal of Economics*, vol. 93, pp. 395–410.

Marschak, J. and Radner, R. (1972). *Economic Theory of Teams*. Yale University Press.

Medawar, P. (1968). '"Lucky Jim": review of J. D. Watson, *The Double Helix*.' *New York Review of Books*, 28 March. Reprinted in Medawar (1982).

—— (1982). *Pluto's Republic*. Oxford: Oxford University Press.

Merton, R. K. (1961). 'Singletons and multiples in scientific discovery.' *Proceedings of the American Philosophical Society*, vol. 105, no. 5, pp. 470–86. Reprinted in Merton (1973), chapter 16.

—— (1963). 'Resistance to the systematic study of multiple discoveries in science.' *European Journal of Sociology*, vol. 4, pp. 239–49. Reprinted partially in Merton (1973), chapter 17.

—— (1973). *The Sociology of Science: Theoretical and Empirical Investigations* (ed. N. W. Storer) Chicago: University of Chicago Press.

Nalebuff, B. J. and Stiglitz, J. E. (1983). 'Prizes and incentives: towards a general theory of compensation and competition.' *Bell Journal of Economics*, vol. 14 (spring), pp. 21–43.

Nelson, R. (1982). 'The role of knowledge in R & D efficiency.' *Quarterly Journal of Economics*, vol. 97, pp. 453–70.

Nordhaus, W. (1969). *Invention, Growth and Welfare*. Cambridge, Massachusetts: MIT Press.

Ogburn, W. P. and Thomas, D. S. (1922). 'Are inventions inevitable.' *Political Science Quarterly*, vol. 37, pp. 83–100.