When is Good Good Enough? A Sructural Estimation of a Sequential R&D Model

Benjamin Bental   Dan Peled[†]

March 5, 2003

Abstract

Firm level data on productivity and R&D expenditures is used to estimate a model in which R&D investment solves an optimal stopping problem of repeated "bites of the R&D apple". The solution to this problem yields threshold levels for terminating the R&D process, which we use together with data on firms' output, inputs, and R&D expenditures to estimate the model. The estimation procedure overcomes some irregularity conditions that otherwise produce biased maximum likelihood estimates in similar models. Despite the preliminary nature of the results, the estimated parameters display some interesting similarities and differences across sectors. If reaffirmed by further elaboration of the basic model presened here, these sectoral differences can suggest benefitial refinements of R&D support policies.

A prerliminary version of the paper, prepared for the "Empirical Economics of Innovation and Patenting" conference, ZEW Mannheim, March 14-15, 2003.

[†] Both: Department of Economics, University of Haifa.
E-mail: <bbental@econ.haifa.ac.il>, <dpeled@econ.haifa.ac.il>

1. INTRODUCTION

1.1. The Importance of Explicit $R\&D$ Models

Ongoing R&D efforts to improve product and process quality consume resources which could instead be used for production of sellable goods and services. Nevertheless, the quest for market share and future profitability can make it worth while to divert some productive resources to R&D. Indeed, many firms spend as much as 25% of their sales revenue on R&D.[1] However, these R&D expenditures differ widely across sectors, across firms within the same sector, and over time for any given firm, (for a comprehensive study see Griliches (1995)). Moreover, due to uncertainty about the results of many R&D ventures, R&D results are only stochatically related to the levels of R&D expenditures.

Most existing studies of industrial R&D evaluate its contribution by adding corporate R&D expenditures as another explanatory "input" variable to econometric estimates of the production function.[2] Within the context of our framework, this approach may be problematic because the firm's input choices are made after it has realized the results of its R&D effort. As a result, the firm's chosen

[1] The R&D-sales revenue ratio, (R&D intensity), displays huge variety across firms in the US. For instance, the total R&D investment of the top 100 R&D-investing US companies in 1998, (estimated as 2/3 of the total R&D investment by US companies), amounted to 5.6% of their combined sales revenue. The average R&D intensity of the top 10 of these 100 R&D-leading companies, (covering among them 25% of the total Business Expenditure on R&D in the US), was 9.7%, while many among the 100 leading firms invested well over 25% of their revenue in R&D. Smaller firms with less than 500 workers had an average R&D intensity of about 13%. Moreover, variations within and across industries in this measure are huge. The median R&D intensity in the Drug & Pharmaceutical industry was 11%, with a 5%-95% range of 6% to 20%, while in Electronic & Electric Equipment the median R&D intensity was 3%, and the 5%-95% range was 1% to 7%. See: Research Technology Management, Vol. 43, No. 1, (January-February 2000), pp. 16-27, Industrial Research Institute.

[2] See, for instance, Griliches (1995).

inputs are endogenous to their productivity levels. Consequently, the estimated productivity attributed to the various inputs used in production, including R&D, are likely to be biased.

In this paper we propose an explicit model of optimizing R&D behavior, and estimate it using firm-level data. We believe that eventually our results may be useful in understanding the vastly different R&D behavior and performance of different kinds of firms, as well in as in assessing the likely impact of various forms of government support programs ranging from R&D subsidies to commissioned research.

Understanding the process by which firms decide how much to invest in R&D, and when, as a function of available resources, technological position, and the likelihood of success - can improve forecasts of technological progress, estimates of productivity gains, and identification of the sources of these beneficial outcomes. As a result, a more efficient system of allocating R&D supports may be developed. A structural estimation like the one proposed here may serve as a tool to answer some basic questions facing policy makers who are interested in the impact of R&D support. There are several levels of this issue, which we may be able to address. First, using the structural model, we may assess the effectiveness of supporting R&D, for instance in the context of estimating how much *additional* R&D is generated by R&D subsidies. Second, we may identify the industries which are more likely to benefit from such a support. Third, such a model can suggest why R&D conduct is different across firms of different sizes.

The underlying theoretical model used here borrows from the search-theoretic mechanism of endogenous growth, which we explored in Bental and Peled (1996), (originally proposed by Evenson and Kislev (1975), and sharing some common elements with the evolutionary industry growth models of Nelson and Winter (1982)). In our framework, the search for better technologies consumes resources which can be used instead in production. Profit maximizing, competitive firms must continuously decide how to allocate their resources between these two alternative uses. This decision process is at the focus of our model. [3]

The model assumes that firms can combine capital and labor at each point to produce marketable output using their current available technology. However, firms can also attempt to improve that technology by allocating some resources to R&D. The R&D process precedes production, and is conducted sequentially by each firm. R&D consists of sampling from a "pool of untried ideas", represented as a distribution function of productivity factors. Each draw from that distribution requires a fixed amount of resources, and reveals the profitability and usefulness of the sampled "idea". Once observed, the firm can decide whether to exploit the idea just drawn with its remaining resources, or keep on searching. The optimal solution to this sequential search problem generates a random outcome for the firm, consisting of actual R&D expenditure for the period, the technological improvement achieved by the R&D efforts, and the resulting output and profits. Interpreting observable firm-level measures of output, inputs, and R&D expenditures as stemming from this process, allows us to estimate the parameters of the R&D-search environment in which firms operate. In particular, the procedure

[3] In our previous work we showed that the model is capable of generating some results that resemble actual performance of economies at the aggregate level. In particular, the model predicts that as the technology level rises, the search for further technological improvements requires an ever increasing investment in R&D. Concurrent with evidence from aggregate data, this feature generates a feedback mechanism, whereby successful R&D leads to technological improvements and higher output which, in turn, can provide resources for increased investments needed for further technological improvements generated by continued R&D.

allows us to differentially estimate the behavior of firms in different sectors and compare the R&D prospects in these sectors. [4]

One of the potential benefits of estimating the R&D decision rules in this explicit form, is that it delivers a sharp prediction about the probability distribution of the returns to the firm from having one more unit of resource at the beginning of the period, (the marginal gain from what is later denoted by q). The rates at which firms can borrow additional funds in the capital markets, (which we currently do not have in the model), in comparison with the estimated marginal gain from q, may provide us with an important independent measure for gauging the plausibility of the model. In addition, this estimate can be contrasted with the exceedingly high *indirect* estimates of "rate of return to R&D" obtained by others researchers (e.g., Coe and Helpman (1995)). Another potential benefit of an explicit model of a firm's R&D decision is the ability to predict its responses to R&D support programs, and the ability to refine such programs accordingly.[5]

In the next section we present the underlying model of R&D as a sequential search. Section 3 describes the estimation procedure, and section 4 the standard firm-level data needed for this procedure. In section 5 we present some initial results obtained by applying the procedure to a particular US firm-level data set. Section 6 concludes with some extensions and implications of the derived estimates.

[4]The idea that R&D is an inherently a sequential activity that can be examined using insight from optimal search theory was also recently exploited by Lach and Sauer (2001). Their theoretical analysis, which applied search theory to examine firms' R&D responses to subsidies, concludes that the results depend on the model's parameters and distributional assumptions, which must be determined empirically.

[5]See Lach (2002) for a recent empirical examination of the change in firms' own R&D expenditures under a cost-sharing subsidy program in Israel.

2. THEORETICAL FRAMEWORK

2.1. An Overview

In order to understand the emphasis on the sequential nature of R&D it is convenient to start with a simplified model of the variety to be used below for estimation. Particular extensions and elaboration, which are needed to match data characteristics and modify over-simplification of this basic version, are discussed in the next section.

Consider a firm that produces a single output by using installed capital and technical know-how. For the sake of this concise presentation we ignore other variable inputs, but these are added in the actual work. These inputs are transformed into output via an appropriately defined objective function, (which can be a standard production or a profit function), multiplied by a function of a firm-specific productivity factor, θ, that represents the know-how of the firm. The productivity factor is amenable to costly improvement by the firm, and captures all impacts that R&D can have on the firm's profitability. The firm enters the "period", (the length of which in empirical application is discussed below), with three initial conditions: (i) a given stock of installed capital, K_0, which captures the firm's current size; (ii) a default productivity factor, θ_0; and (iii) an amount of "flexible" resources, q_0, which can be freely allocated between R&D and "production", where the latter captures all other possible productive uses of these flexible resources.

In order to introduce the notion of R&D as a sequential search, we first present a skeleton version of the model, and defer further elaboration required for empirical application. Let $\phi(K, \theta)$ be the firm's payoff when it produces with K units of capital, and has productivity θ. Before embarking on production, the firm goes through a "search phase", in which it can attempt to improve upon its initial productivity θ_0. This search takes the form of successive draws from a distribution of "untried" ideas, $H(\cdot \mid \Lambda)$, where Λ is a vector of parameters of a distribution known to the firm.[6] Each draw costs α units of the resource. Observing the outcome of each draw, the firm decides whether to stop the search, add all its remaining resources to the installed capital and start production using its attained productivity, or take another draw. Let $V(q, K_0, \theta)$ be the (optimal) value of the firm during the search phase, when it has q remaining units of the resource, and its current best productivity is θ. Then, this value function must satisfy the following Bellman equation:[7]

$$V(q, K_0, \theta \mid \theta_0) = Max\left\{\phi(K_0 + q, \theta), \phi(K_0 + q, \theta_0), E\left[V\left(q - \alpha, K_0, \tilde{\theta} \mid \theta_0\right)\right]\right\}. \tag{2.1}$$

The first two terms in the brackets on the right represent the option to stop the R&D process and apply the better technology at hand, $(\max\{\theta, \theta_0\})$, to all remaining resources. The last term in the maximand represents the decision to continue for at least one more draw, and the expectations are taken with respect to the realization of the next draw from the known distribution $H(\cdot \mid \Lambda)$.

For given ϕ and H the functional equation in (2.1) can be explicitly solved for the optimal R&D decision rule, which takes the following form:

[6] We, of course, do not know the distribution's parameters, but we pretend to "know" the family of distributions from which the firm draws. The parameters are the subject of the estimation procedure described below.

[7] The following formulation assumes that the search is conducted "without recall" of sampled productivities which were initially rejected. This is an important, albeit unrealistic, simplification. See Wolpin (1987, footnote 5) for a brief comment on this issue.

Keep drawing from H as long as the level of flexible resources is sufficiently high to allow another draw, until the first draw of a technology surpasses a threshold, $\theta^*(q, K_0, \theta_0)$, that depends on the remaining level of "flexible" resources,q, the installed capital,K_0, and the default productivity at hand, θ_0.

Optimal R&D decision rules of the kind implied by the functional equation in (2.1) can be derived for alternative specifications. These can include alternative specifications of the distribution of untried ideas, and different objective functions attributed to the firm.[8]

2.2. The General Estimation Approach

The theoretical framework described above maps structural parameters, including the characteristics of the distribution of the technological innovations, into observables. In particular, the model generates stopping rules for firms engaged in R&D, a distribution of R&D expenditures, as well as a distribution of productivity measures and of profits (or rates of return on investments).

All these objects depend on the underlying parameters. In particular, the parameters governing the distribution of the potential technological innovations and the sampling cost are central to the calculation of the optimal stopping rules and through those, to all observables.

We propose to utilize these relationships in order to estimate a structural model of R&D investment. We interpret the data on individual firms engaged in R&D as stemming from our model. In particular, we think of the observed R&D expenditure as if it results from a sequence of "stop or continue" decisions. The

[8] Detailed description of the procedure for recursively solving similar functional equations is provided in Appendix A.

data is also used to provide us with both the default and the resulting productivity improvement, as measured by total factor productivity. A sufficiently large sample of observations on ex-ante similar firms engaged in R&D which contains these bits of information, combined with the theoretical structure of the stopping rules, allows us to identify and estimate the structural parameters.

We proceed by describing in some detail how firm level data can be used to estimate the parameters of such a model of R&D.

The model produces threshold acceptance levels for productivity, $\{\theta^*(q, K_0, \theta_0)\}$, for any given level of remaining flexible resources, q. The $\theta^{*\prime}s$ depend in a known way on:

The amount of available "flexible" resources for the period, q,

"Installed production capital", K,

The cost of a "search step", α,

The number of search steps taken up to that stage, n,

The parameters of the distribution of new productivity levels, Λ,

The objective function, $\phi(\cdot, \cdot)$.

Observables at the firm level over a given period need to include:

Amount of installed capital at the beginning of the period, K_0,

Initial and final levels of "productivity": θ_0 and θ_e,

Investment in capital during the period, I,

$R\&D$ expenditure during the period, R.

We reconstruct the initial level of "flexible" resource which was available to the firm at the beginning of the period as:

$$q_0 = R + I. \tag{2.2}$$

Accordingly, the *maximum* number of R&D-search steps the firm could have taken during the period is:

$$N = [q_0/\alpha]$$

whereas the *actual* number of R&D-search steps taken by the firm during the period is:

$$n = [R/\alpha]. \tag{2.3}$$

Using the above structure, we can construct a likelihood function for a sample of searching firms. In particular, given observations on K_0, R and θ_0 and $\theta_e$and q_0, and the inference on n and N, conditional on the unobserved search cost parameter α, and the parameters of the distribution $H(\cdot \mid \Lambda)$, the contribution of an individual firm to the likelihood of a given sample is:

$$\begin{aligned} & \Pr\left\{\text{first } n-1 \text{ draws were unacceptable, } n^{th} \text{ draw } \geq \theta^*\right\} \cdot \Pr\left\{\theta_e \mid \theta_e \geq \theta^*\right\} \\ = \;& \left\{\prod_{i=1}^{n-1} H\left[\theta^*(q_0 - i\alpha, K_0, \theta_0)\right]\right\} h(\theta_e; \Lambda) \end{aligned}$$

where h is the (unconditional) density of H.

Accordingly, a maximum likelihood procedure allows us to estimate the parameters of the model using a sample of firms with the observables described above. In particular, we can estimate the parameters of the distribution of "untried ideas". Moreover, samples of firms from different sectors, will enable us to estimate sector-specific parameters. Such estimates potentially allow us to compare the marginal value of flexible resources, (capital infusion), across sectors, and for differently sized firms within a sector, and predict how firms would react to a particular R&D-support program, should one be offered to them.

In practice, there are many additional details that need to be added to the model to prevent the implied likelihood from being set to zero by the data. In the next section we explain in details how we accomplish that within a more elaborate version of the static R&D-search model presented above.

3. AN ESTIMABLE MODEL OF SEQUENTIAL R&D

We begin by specifying functional forms for some of the components of the full model which will be used in the estimation. Let $\phi(k,\theta) = \theta^{1/\gamma}k$ be the outcome of applying technology θ to k units of capital, (after labor and other variable inputs have been optimized).[9]

The second specification involves the distribution of the untried technologies, H. Here we assume that potential technologies are drawn from an Exponential distribution[10]

$$H(\theta) = 1 - \exp^{-\lambda(\theta - b)},\ \theta \geq b \geq 0, \quad \lambda \geq 0. \tag{3.1}$$

Although this distribution has two parameters, we focus at this stage on the standard case where $b = 0$, and hence θ is the only parameter to be estimated.[11]

[9] Such an objective results from treating capital as a fixed factor, and substituting the profit maximizing labor input for the production function $y = \theta A K^{\gamma} L^{1-\gamma}$, given a real wage rate. See Bental and Peled (1996) or Appendix A for details.

[10] Our previous work that has investigated the implications of the sequential search model for economic growth, showed that to sustain growth the distribution of untried technologies has to be Pareto with a sufficiently "thick tail". Empiriacllay, the Pareto distribution is rejected by the data.

[11] As a matter of fact, $b = 0$ turned out to yield the highest likelihood.

3.1. The Firm's Search Problem

As noted above, the firm begins the period with three initial conditions. First, it has what we call "installed capital", K_0, which represents the firm's market size or basic production capacity. This is the minimal level of operation that the firm will undertake with the random outcome of the R&D-search process. It is a minimal measure of scale, since the firm might augment that with some of its flexible resources, should there be any left from the sequential R&D.

Second, the firm begins the period with a given amount of flexible resources, q_0, which can be used for R&D or production. This amount can be retained earnings from operations during the previous period, or an external investment or loan. Since the model is based on just a single period snapshot of the firm, there is no way it can explain the source of these funds, or examine the interaction between the source of such funds and the firm's conduct during the observed period.

The third initial condition is a *default productivity level*, θ_0, which we take to be the previous productivity level achieved by the firm in the previous period. At any point during the observed period, the firm can abort the R&D process, and attempt to utilize θ_0 with all its remaining resources. However, we assume that the previous knowledge depreciates at the rate δ, and accordingly, when such a default option is exercised, it yields an effective productivity of $(1-\delta)\,\theta_0$.[12] Specifically, if a firm ends the R&D process after it consumed $R \leq q_0$ units of the flexible resources and attempts to exploit the default technology θ_0, then the outcome is given by $\phi\left(K_0 + (q_0 - R)\,,(1-\delta)\,\theta_0\right)$.

The firm has to decide whether to engage in R&D at the beginning of the

[12] The "knowledge depreciation" parameter is required by the data. One may think of it as representing the cost of remaining with an old technology, while other firms adopt a new one.

period, and if so when to stop it. It faces a cost of α units of the flexible resource for each R&D-search step, and faces a sequence of "go" "no-go" R&D investment decisions.

The optimal solution to the firm's sequential R&D-search problem is obtained, as in the previous section, by dynamic programming. Let $q \geq \alpha$ be the remaining amount of the flexible resource during the search phase, when the firm has just drawn a technology θ. Then the value function satisfies:

$$v(q, \theta \mid K_0, \theta_0) = \max \left\{ \begin{array}{c} E[\phi(q + K_0, \theta)] \ , \\ E[\phi(q + K_0, (1-\delta)\theta_0)] \, , \\ E\left[v\left(q - \alpha, \tilde{\theta} \mid K_0, \theta_0\right)\right] \end{array} \right\} \tag{3.2}$$

The first term on the RHS corresponds to adopting θ and applying to all remaining resources. The second term corresponds to resorting to the default technology with all remaining resources. The third term corresponds to taking another technology draw, $\tilde{\theta}$, and then decide what to do next.

For $0 < q < \alpha$ we have:

$$v(q, \theta \mid K_0, \theta_0) = \max \{\phi(q + K_0, \theta) \ , \ \phi(q + K_0, (1-\delta)\theta_0)\} \tag{3.3}$$

The functional equations (3.2) and (3.3) imply a set of threshold values, $\{\theta_i^*\}_{i=1}^{\overline{N}}$, with the maximal number of R&D-search steps given by $\overline{N} = \left[\frac{q_0}{\alpha}\right]$. These threshold values depend on $q, K_0, \theta_0, \delta, \alpha$, and the parameters of the distribution of untried technologies, $H(\cdot)$. The firm proceeds in the R&D process as long as the technology it has sampled falls below the threshold value corresponding to its remaining flexible resources. The threshold values can be computed recursively, (see appendix A).

3.2. The likelihood function

It is convenient to abstract, at this stage, from the parameters that pertain to the firm's technology and regard them as if they are known. In particular, the values of α , δ, and θ_0, that are not directly observable and will eventually need to be estimated. However, at this stage they are regarded as known. This leaves only the parameter λ of the exponential distribution (see (3.1)) to be estimated.

The probability that a firm with default technology θ_0 and K_0 units of capital will take N R&D steps before stopping is:

$$\left((1 - H(\theta_N^*)) \prod_{i=1}^{N-1} H(\theta_i^*) \right).$$

The pdf of the actual productivity finally applied by a firm conditioned on finding a suitable one at the N^{th} step, denoted by θ_e, is:

$$\frac{h(\theta_e)}{1 - H(\theta_N^*)}$$

Accordingly, the contribution to the likelihood function of a firm which took N R&D steps, and displayed productivity level θ_e is:

$$\left(\prod_{i=1}^{N-1} H(\theta_i^*) \right) h(\theta_e) \tag{3.4}$$

3.3. Consistency Constraints

(a) End-of-search constraint:

Search is terminated when the firm draws a θ which surpasses the appropriate threshold. Thus, if firm j in the sample, started the R&D process with $(q_{0j}, K_{0j}, \theta_{0j})$, preformed N_j R&D steps, and ended with actual observed productivity θ_{ej}, then necessarily the result of the N_j^{th} R&D step must have been acceptable. That is:

$$\theta_{ej} \geq \theta^*_{N_j} \left(q_{0j} - N_j \alpha, K_{0j}, (1-\delta)\theta_{0j}\right). \tag{3.5}$$

This constraint is not explicitly incorporated into the likelihood function. However, that function is valid only if the constraint is satisfied.

(b) Continued search constraint:

For the search to continue for N_j steps, it must be true that: $(1-\delta)\,\theta_{0j} < \theta^*\left(q_{0j} - n\alpha, K_{0j}, z_{0j}\right)$ for $n = 0, 1, 2, ..N_j - 1$. Since the threshold levels are decreasing, this amounts to:

$$(1-\delta)\,\theta_{0j} < \theta^*_{N_j-1}\left(q_{0j} - (N_j - 1)\,\alpha, K_{0j}, (1-\delta)\theta_{0j}\right). \tag{3.6}$$

This constraint is not explicitly incorporated in the estimation procedure of the likelihood function below. Notice, however, that in general it is more readily satisfied the higher is δ, and the lower is λ, (lower λ increases θ^*).

c) Non-participation constraint:

For firm j to abstain from even starting the R&D process, it must have:

$$(1-\delta)\,\theta_{0j} \geq \theta^*\left(q_{0j}, K_{0j}, (1-\delta)\theta_{0j}\right). \tag{3.7}$$

Note that this constraint has the exact opposite implications on δ and λ than

(3.6).

At this stage, we have not incorporated information on firms which do not report any R&D expenditures.[13]

4. DATA AND PRELIMINARY CALCULATIONS

4.1. Data

We use two data sets. General industry characteristics are derived from the NBER-CES Manufacturing Data base described by Bratelsman and Gray (1996). This data set contains a panel of 4-digit ISIC industry indicators, from which we compute the capital shares (denoted by γ) for the sectors under consideration as explained below.

The second data set we use is the Compustat based firm-level data compiled from a US Productivity Panel by Bronwyn Hall for the years 1977-1989.[14] We selected five sectors from this data set which roughly correspond to high-, medium, and low-tech sectors. The sectors chosen were: (i) electronics and other electric equipment, instruments, computers and office equipment; (ii) industrial machinery and equipment; (iii) tobacco products, lumber and wood products, furniture and fixures, stone clay and glass products, and miscellaneous manufacturing industries; (iv) paper, paper products, printing and publishing; and (v) chemical and allied

[13] The specific data set we use contains information only on firms that report some positive amount of R&D. We were forced to actually curtail this set further, and take into account only firms that report a sufficiently high level of R&D (1.1 million dollars). This restriction was dictated by the maximal number of search steps that could be handled by the PCs we use. Clearly, this constraint may be relaxed by using more powerful computers.

[14] The data set is available from ftp://elsa.berkeley.edu/users/bhall/index.html.

products and drugs.[15]

We constructed a sample consisting of all firms in the data for each of our five sectors in the years 1985-6 and 1987-8.[16] The observations we have on these firms include measures of output, ($y =$ value of sales net of material cost), capital stock at the end of the period, (K_0), investment, (I), and expenditure on R&D, ($R\&D$) as well as expenditures on labor and materials. Some descriptive statistics on the distributions of firm characteristics across sectors in the two time periods considered are presented in Tables 1 below. For each industry, in each of the two periods, we report statistics on the number of firms in the sample, (and the number of firms excluded from estimation), sales, capital stock, R&D expenditure, and R&D intensity, (R&D expenditures to sales). We also report some measures which we compute from that data, according to our model. Those include the fraction of flexible resources allocated to R&D, which we proxy by: $R\&D/(R\&D + I)$, denoted in the table as $R\&D/q_0$, and the productivity levels of the firms before and after conducting R&D. The latter are reported in the last two rows of Table 1 as $A\theta_0$ and $A\theta_e$, computed as $\frac{y_t}{K_t^{\gamma} L_t^{1-\gamma}}$ for the $t = 1985$-6, and for $t = 1987$-8, respectively, where the capital share γ is computed from the NBER Manufacturing database, as described below.

[15] Precise sectoral classicifications are described in references in the data documentation, (see previous footnote).

[16] Here too the sample is dictated by the capacity constraints of our PCs. See below.

Table 1: Descriptive Satistics

1986 & 1988 US R&D Manufacturing Firms (used in estimation)

		Electronics		Machinery		Tobacco		Paper		Chem	
		1986	1988	1986	1988	1986	1988	1986	1988	1986	1988
N_s		262	231	73	66	32	25	17	16	81	67
excl.$^{(a)}$		68	56	23	18	15	14	12	11	28	28
$y^{(b)}$	avg	333.43	376.70	806.72	833.04	889.26	693.51	1791.98	1968.91	938.24	928.20
	med	114.33	126.41	297.90	328.65	426.15	171.88	1743.31	1665.08	355.39	387.04
	min	4.79	11.16	18.38	6.96	42.90	18.30	416.94	541.68	2.34	2.25
	max	5111.35	6395.67	7555.21	5158.47	5429.49	6957.59	4440.77	5186.06	5720.33	5288.98
$K_0^{(b)}$	avg	97.06	100.60	310.14	264.98	341.12	235.80	1519.12	1438.52	433.17	389.87
	med	27.63	29.12	119.39	104.91	134.81	48.55	1458.51	1272.82	98.06	99.48
	min	1.33	1.04	4.10	2.06	12.47	7.94	98.45	111.10	2.74	1.46
	max	1113.32	1213.19	3971.18	1701.29	2516.67	1556.19	4023.94	3452.59	2785.37	2763.96
$R\&D^{(b)}$	avg	21.28	22.97	22.62	20.30	11.71	7.80	22.39	23.09	42.48	38.79
	med	7.14	7.78	9.96	9.11	4.40	3.19	11.78	12.23	11.73	13.96
	min	1.11	1.12	1.22	1.28	1.12	1.14	4.23	4.83	1.21	1.30
	max	315.48	263.30	225.87	208.61	63.36	60.33	111.56	107.15	337.19	287.83
RDI (%)	avg	8.08	7.48	4.03	3.82	1.82	1.59	1.31	1.25	16.45	17.52
	med	7.23	6.86	2.63	2.58	1.30	0.88	1.23	1.25	4.31	4.24
	min	0.45	0.70	0.36	0.18	0.37	0.07	0.21	0.23	0.77	0.86
	max	60.60	28.74	21.76	35.30	5.61	5.43	2.83	2.45	414.75	380.02
$R\&D/q_0$	avg	0.53	0.52	0.42	0.49	0.23	0.31	0.14	0.16	0.43	0.44
	med	0.54	0.54	0.38	0.47	0.19	0.27	0.10	0.13	0.41	0.43
	min	0.05	0.04	0.03	0.07	0.03	0.01	0.02	0.02	0.04	0.10
	max	0.94	0.92	0.88	0.95	0.59	0.79	0.32	0.45	0.99	0.94
$A \cdot \theta_0$	avg	9.55	11.02	8.88	9.88	8.01	9.14	5.28	6.46	4.69	5.78
	med	8.34	9.39	8.37	9.31	7.39	8.02	4.19	5.77	4.02	5.12
	min	2.91	3.62	3.50	4.27	4.34	3.28	2.42	3.10	0.53	0.72
	max	39.06	42.96	19.41	22.35	21.95	23.02	11.09	12.30	12.28	33.35
$A \cdot \theta_e$	avg	10.21	10.99	9.15	10.25	8.44	9.41	5.56	6.65	4.96	5.57
	med	8.93	9.64	8.56	9.90	7.83	7.69	4.41	5.45	4.41	5.27
	min	2.09	3.32	3.61	3.65	5.05	3.53	2.67	3.32	0.77	0.67
	max	44.53	36.25	20.88	23.17	18.23	34.04	11.28	12.83	11.72	14.00

$^{(a)}$ Number of firms excluded from estimation due to incomplete data, or R&D expenditures which were either less than \$1.1m or total investments (q) exceeding \$550.m.

$^{(b)}$ All dollar figures are in millions of current \$.

As can be seen, the five industries chosen span sectors with highly varied R&D efforts. In particular, the R&D intensity ranges from an median value of 6-7% in Electronics to around 1% in Tobacco. However, few intensive R&D firms in Chemicals make the RDI average in that industry the highest among the five considered. The share of flexible resources allocated to R&D is on average in the 40%-53% range for Electronics, Machinery and Chemicals, while it is only in the 15%-30% range for Tobacco and Paper. The average size of firms, as measured by the average installed capital, is the highest in the Paper industry, and the lowest in Electronics. Finally note that within each industry there is a huge variation in size as well as R&D conduct, which is reflected in the minimum and maximum levels, and the skewness captured by the mean/median ratios.

5. ESTIMATION PROCEDURE

For given values for α and δ, the likelihood function involves only λ, the parameter of the distributions $H(\cdot)$. Given a sample of size N_s of searching firms from one of the five sectors, the likelihood function implied by the model is:

$$\prod_{j=1}^{N_s}\left\{\left(\prod_{n=1}^{N_j-1} H\left(\theta_n^{j*}\right)\right) h\left(\theta_{ej}\right)\right\} \tag{5.1}$$

where θ_{ej} is the actual productivity displayed by firm j at the end of the period, and $\theta_n^{j*} \equiv \theta^*\left(q_{0j} - n\alpha, K_{0j}, (1-\delta)\theta_{0j}\right)$, for $n = 0, 1, 2, ..N_j$, are the threshold technology levels for firm j, given its default technology θ_{0j}, its installed capital, K_{0j}, and the residual amount left of the flexible resource after n R&D steps.

5.1. The regularity problem

Given the observed variables and all other parameters, one is tempted to pick λ simply by maximizing the likelihood of the entire sample of firms. This procedure turns out to fail the regularity conditions needed to make the maximum likelihood estimator consistent (Wald (1949)). In particular, the effect of the parameter λ on θ_i^* biases the maximizing value of (5.1) downwards (see Appendix B).[17] As a result, we have devised the following procedure: We pick an arbitrary value of λ, say λ_0. We use λ_0 to compute the values of θ_i^* for all i. We then find $\lambda_1(\lambda_0)$ that maximizes (5.1), using the computed values θ_i^*. If $\lambda_1(\lambda_0) = \lambda_0$, we stop the procedure. If not, we replace λ_0 by λ_1, compute the values of θ_i^*,and again maximize (5.1) to obtain $\lambda_2(\lambda_1)$. We iterate in a similar fashion until $\lambda_Q(\lambda_{Q-1}) = \lambda_{Q-1}$. In Appendix B we show (by a simple simulation) that (5.1) is indeed maximized at the true value of λ, when that value is used to compute θ_i^*as well.

5.2. Other Parameters

The discussion above assumed that the values of the remaining parameters: A, γ, α, and δ are known. We turn now to describe the procedures adopted for estimating these parameters.

[17] The problem seems to be related to that discussed by Donald and Paarsch (1993a,b). There, the problem is that the support of the distribution depends on the parameters that are estimated. The solution they propose circumvents the problem. While that solution is not directly applicable to our case, we do use it in spirit for the estimation of other parameters. See below.

5.2.1. Estimation of the Capital Share, γ

We cannot estimate the capital share directly from the production function, since our theory implies that the capital and labor inputs are correlated with the "error term" (the technology factor) in any such equation. Therefore, we estimate the labor share in each of the sectors as the average of the ratios between labor expenditure, (total pay), and output, (value of shipments net of material expenditure) in the industries comprising that sector for three consecutive years. The capital share is the complemet of this ratio to one. The results are reported in Table 2 below.

5.2.2. Estimation of α

In principle, a consistent estimator of α is the smallest observed amount of R&D in the sample.[18]

In practice, since very low values of α imply huge matrices of θ^*, we had to bind the value of α from below (see footnote 13 above).[19]

5.2.3. The choice of δ

Here we need to take into account the constraints (3.6). The value of δ is chosen so as to satisfy **all** these constraints.[20] The actual procedure is again iterative.

[18] See, for example, footnote 4 in Donald and Paarsch (1993b).

[19] The choice of α is not innocuous. In particular, lower values of α imply that λ has to decrease, to "justify" the increased number of search steps.

[20] This procedure follows, in spirit, the idea suggested by Donald and Paarsch (1993a). In their case, the lower bound of the distribution they are interested in is determined by the smallest observed value of a relevant variable. The bound is also a function of the parametrs they want to estimate.They suggest to solve out one of the parameters, and then estimate the rest of the parameters by maximizing the "concentrated" likelihood function. Setting δ so as to satisfy all "continued search" constraints in effect does the same.

Given A, we start with an initial value of δ and estimate λ. We then compute the constraints. If some are violated, δ must be changed and λ is re-estimated. The procedure stops when the "marginal" constraint is satisfied with equality.

Notice that the value of δ estimated in this way is likely to violate the non-participation constraints, (3.7). We ignore this problem at this stage of our work.

5.2.4. The choice of the technology parameter A

We have explained above how we compute the value of $A\theta_{0j}$.and $A\theta_{ej}$. It is crucial, of course, to identify A in order to be able to compute θ_{0j} and θ_{ej} that are the basic ingredients of our estimation procedure. Since we cannot identify A in any direct fashion, we search over a grid of admissible values of A and choose that value that, in the final analysis, yields the highest value of the likelihood function.

6. RESULTS

As explained above, we had to bind the value of the search-cost, α, from below, and the value of flexible resources, q_0, from above, in order to keep the size of the matrices involved within the memory capacities of the PC. The combination we use is $\alpha = 1.1$, and $q = 550$ (all in millions of \$). Accordingly, the largest number of potential search steps is 500. These choices dictate the sample sizes for the various sectors and years. Table 2 below summarizes the estimated parameters of our model for the five industries over two different time periods.

	Table 2: Estimation Results					
		Electronics	Machinery	Tobacco	Paper	Chemical
	Capital Share (γ)	0.56	0.52	0.54	0.57	0.70
1986	N	262	73	32	17	82
	A	3.4	3.3	5.1	1.7	0.6
	δ	0.94	0.88	0.85	0.90	0.98
	λ	2.4807	2.6254	2.5416	2.4354	2.3020
	standard error	0.0398	0.0790	0.1587	0.1490	0.0437
	5% confidence	2.4022-2.5591	2.4679-2.7830	2.2181-2.8652	2.1194-2.7514	2.2150-2.3890
	mean LogL	-9.8492	-9.3393	-4.8676	-10.7015	-30.2892
1988	N	231	66	23	15	67
	A	3.5	4.1	5.5	2.7	0.9
	δ	0.95	0.87	0.77	0.88	0.99
	λ	2.5428	2.6281	2.5157	2.5041	2.3283
	standard error	0.0420	0.0879	0.1842	0.1857	0.0536
	5% confidence	2.4600-2.6250	2.4525-2.8038	2.1337-2.8977	2.1057-2.9023	2.2213-2.4353
	mean LogL	-10.721	-8.2946	-4.9981	-8.1749	-20.9215

Table 2 reveals the following:

1. The λ coefficients seem to be quite stable over the two time periods we have used. In particular, the coefficients of of 1986 and of 1988 in all sectors are basically within one standard error of one another (except electronics).They are clearly within the 5% confidence intervals of one another.

2. The λ coefficients are surprisingly similar across sectors.

3. The λ coefficients of the machinery, electronics and chemical sectors seem to be ranked (with chemicals being lowest, machinery highest).

4. The λ coefficients of tobacco and paper are less sharply estimated. Their 5% confidence interval is quite similar, and basically includes the ranges of the three other sectors. In particular, the higher end of the interval for these two sectors tends to exceed that of the other sectors.

3. The A coefficients are all smaller in the 1986 sample than in the 1988

sample. This implies that, on average, the technology level in 1988 is higher than that of 1986.

4. The depreciation rates required to guarantee the "continued search" constraints are very high. They are particularly high for the chemical products sector.

7. Discussion

It is quite revealing to compare the results in Table 2 to the data characterization of Table 1. Looking at averages, Table 1 is characterized by the following rankings (where >>> indicates a difference by a factor of about 4, and >> indicates a difference by a factor of about 2):

y: Paper >> Chemicals >> Machinery $\simeq$ Tobacco >> Electronics

K_0: Paper >>> Chemicals > Machinery $\simeq$ Tobacco >> Electronics

$R\&D$: Chemicals >> Paper $\simeq$ Machinery $\simeq$ Electronics >> Tobacco

RDI: Chemicals >> Electronics >> Machinery >> Tobacco $\simeq$ Paper

$R\&D/q_0$: Electronics > Machinery $\simeq$ Chemicals >> Tobacco >> Paper

This ranking reflects that of the estimated "R&D environment" (indicated by the value of λ) for the R&D Intensity. In other words, the model implies that R&D intensity is as high as it is for the chemical industry because that industry faces a more favorable R&D environment than that of electronics and machinery. The low R&D intensity of the tobacco and paper industries is, of course, consistent with the upper end of the confidence interval of λ for these industries.

These findings highlight the question of industrial policies our model raises. While the R&D environment seems to be most favorable for the chemical industry, it is also true that that industry is by far the most capital intensive in our sample. Therefore, government support programs that are restricted to R&D may, in the final analysis, be less suitable to the chemical industry than to electronics, where capital intensity is much lower. Exact calculation are required to make this judgement.[21]

While we believe that the initial success of the structural model proposed above is quite encouraging, we realize that there is a lot that has to be added to the model.

One of the striking findings is actually the failure of the model to fit the data under a Pareto distribution of the untried technologies (see footnote XX above). The Exponential distribution, on the other hand, yields what seem to us to be quite high values of λ, that imply very low expected productivity gains. It seems reasonable to ask whether a specification that will encompass the features of both the Pareto and the Exponential distributions may not be more appropriate. Accordingly, we plan to re-estimate the model, using the Gamma distribution.

The use of the Gamma distribution may help also resolve the problem of the very high "knowledge depreciation" rates the model requires at this stage. These exceedingly high rates reflect the low threshold values generated under the Exponential distribution. Another way to resolve this problem is to abandon the requirement that all firms in the industry face the same "knowledge depreciation" rate. Doing this will relax the need to satisfy the "continued search" constraint of the marginal firm. However, at this stage it is not clear how this route is to be pursued.

[21] In Bental and Peled (2002) we have a macroeconomic model that compares government support programs restricted to R&D, to programs that allow firms to decide on the allocation of government funds between R&D and capital formation. The calibration of the model to Israeli data reveals that the aggregate result is quite insensitive to the particular policy.

We have mentioned above the question of calculating the return to another unit of malleable resource to different firms in different industries. This calculation, while not complicated, needs still to be carried out in order to see whether our model also yields exceedingly high returns to R&D as predicted by more traditional estimation methodologies.

Finally, at this stage the model is static, and looks at every time period as being independent of the next. Clearly, this is a gross simplification. What is needed is a dynamic version of the model, that may exploit the panel data, and allow for the possibility of multi-year R&D endeavours. Clearly, this is a very ambitious goal.

References

[1] Bartelsman, E. J., and W. B. Gray, (1996), "The NBER Manufacturing Productivity Database", NBER, Technical Working Papers 205, October 1996.

[2] Bental, B. and D. Peled, "The Accumulation of Wealth and the Cyclical Generation of New Technologies: A Search Theoretic Approach", *International Economic Review* 37(3), (1996), 687-718.

[3] ——————————, "Quantitative Growth Effects of Subsidies in a Search Theoretic R&D Model", *Journal of Evolutionary Economics*, 12 (2002), 397-423.

[4] Coe, D. T. and E. Helpman, "International *R&D* Spillovers," *European Economic Review* 39(5), (May 1995), 859-887.

[5] Donald, S., and H. Paarsch, (1993a), "Piecewise Pseudo-Maximum Lilelihood Estimation in Empirical Models of Auctions", *International Economic Review*, 34(1), (Feb. 1993), 121-148.

[6] ——————————, (1993b), "Maximum likelihood estimation when the support of the distribution depends upon some or all of the unknown parameters", Working Paper, 1993.

[7] Evenson, R. E. and Y. Kislev, "A Stochastic Model of Applied Research," *Journal of Political Economy* 84 (1975), 265-281.

[8] Griliches, Z., "R&D and Productivity: Econometric Results and Measurement Issues", in Handbook of the Economics of Innovations & Technical Change, Paul Stoneman, Ed., Oxford, Basil Blackwell Ltd., 1995, pp. 52-89.

[9] Lach, S., “Do *R&D* Subsidies Stimulate or Displace *R&D*? Evidence from Israel”, *Journal of Industrial Economics*, 2002, 369-390

[10] Lach, S. and R. Sauer, (2001), ”R&D, Subsidies and Productivity”, STE Working Paper 7-2001.

[11] Nelson, R. R. and S. G. Winter, *An Evolutionary Theory of Economic Change*, The Belknap Press of Harvard University Press, 1982.

[12] Wald, A., (1949), ”Note on Consistency of the Maximum Likelihood Estimate”, *Annals of Methematical Statistics* 20 (1949), 595-601.

[13] Wolpin, K., (1987) ”Estimating a Structural Search Model: The Transition from School to Work”, *Econometrica*, 55(4), (July 1987), 801-817.

A. Solving an R&D-search Problem

The firm's objective function for a single period, $\phi(k, \theta)$, is derived from solving the firm's profit maximization, given a Cobb Douglas production function, productivity level θ, k units of capital, and real wage rate w:

$$\max_{L} \left\{ \theta A k^{\gamma} L^{1-\gamma} - wL \right\}.$$

Substituting the resulting optimal labor input yields the following profit function:

$$\pi(k, \theta, w) = A^{1/\gamma} \gamma \cdot \left(\frac{1-\gamma}{w} \right)^{(1-\gamma)/\gamma} \cdot k\theta^{1/\gamma}. \tag{A.1}$$

The searcher controls only the last term,

$$\phi(k, \theta) = k\theta^{1/\gamma} \tag{A.2}$$

by choosing a search strategy, which affects the productivity and scale of of production in the current period.

Accordingly, the objective is to maximize the expected value of $\phi(\cdot, \cdot)$ by choosing a strategy that maps sampled technologies and remaining new capital into the binary decision "accept" or "reject". At each stage during the search process the searcher can choose a default technology option, z_0. Accepting a technology means stopping the search and operating that technology with all available capital, $q + K_0$. Rejecting a sampled technology means making at least one more draw, at the cost of α units of new capital. The search is conducted over draws from the distribution $H(\cdot)$. As noted in the text, we choose the Exponential distribution, $H(\theta) = 1 - \exp^{-\lambda\theta}, \quad \theta \geq 0$.

The Bellman equation that summarizes the optimal decision is:

$$V(q,\theta \mid K_0,\theta_0) = \tag{A.3}$$
$$Max\left\{\phi(q+K_0,\theta), \phi(q+K_0,(1-\delta)\,\theta_0), EV(q-\alpha,\tilde{\theta} \mid K_0,\theta_0)\right\},$$

Solving (A.3) yields the optimal search strategy, to be denoted $\theta^*(q,K_0,\theta_0)$, such that the search process is stopped, and the technology θ is utilized with $q+K_0$ units of capital, as soon as $\theta \geq \theta^*(q,K_0,\theta_0)$.

Since $\phi(k,\theta)$ increases in θ, the threshold technology level can be found by equating the first and last terms in the maximand in (A.3), utilizing the fact that when $\theta \geq \theta_0$:

$$V(q-\alpha,\theta \mid K_0,\theta_0) = \begin{cases} \phi(q-\alpha+K_0,\theta), \text{ if } \theta \geq \theta^*(q-\alpha,K_0,\theta_0) \\ \phi(q-\alpha+K_0,\theta^*(q-\alpha,K_0,\theta_0)), \quad \text{if } \theta < \theta^*(q-\alpha,K_0,\theta_0). \end{cases} \tag{A.4}$$

In particular, (A.4) implies:

$$EV(q-\alpha,\tilde{\theta} \mid K_0,\theta_0) = H\left[\theta^*(q-\alpha,K_0,\theta_0)\right]\cdot\phi\left(q-\alpha+K_0,\theta^*(q-\alpha,K_0,\theta_0)\right)$$
$$+\int_{\theta^*(q-\alpha,K_0,\theta_0)}^{\infty}\phi(q-\alpha+K_0,\theta)dH(\theta) \tag{A.5}$$

Equating the first and third terms in the maximand of (A.2), using (A.5), together with the particular specification of $\phi(\cdot)$ and $H(\cdot)$, we get $\theta^*(q,K;\theta_0)$ as the solution to the recursive relation:

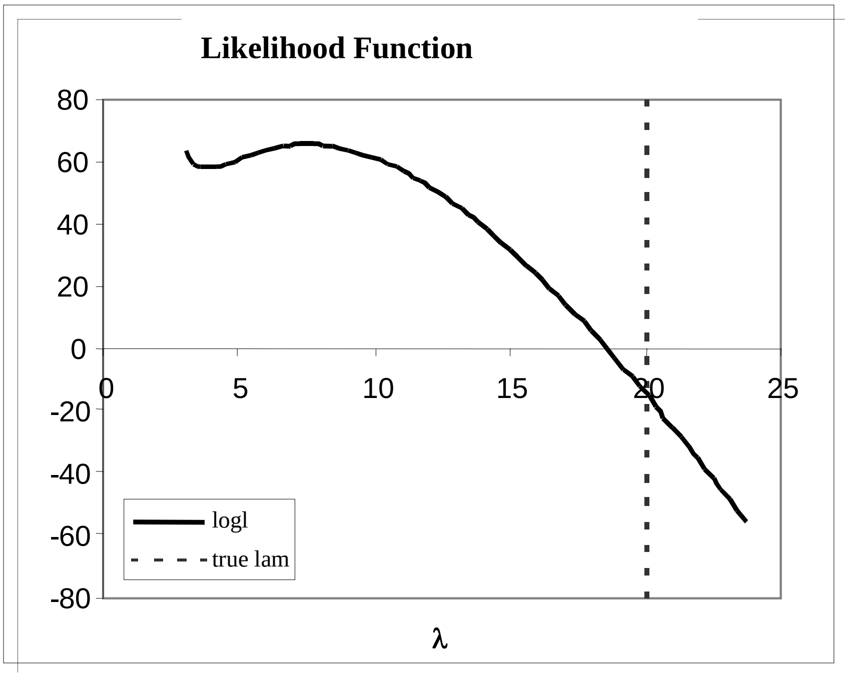
$$\theta^*(q, K_0, \theta_0) = \left(1 - \frac{\alpha}{q + K_0}\right)^{\gamma} \cdot \tag{A.6}$$

$$\left\{[1 - \theta^*(q-\alpha, K_0, \theta_0)^{-\lambda}]\theta^*(q-\alpha, K_0, \theta_0)^{1/\gamma} + \lambda \int_{\theta^*(q-\alpha, K_0, \theta_0)}^{\infty} \theta^{1/\gamma - \lambda - 1} d\theta \right\}^{\gamma}$$

Equation (A.6) allows us to find $\theta^*(q, K, \theta_0)$ recursively. Specifically, given any initial quantity of new capital, q, $N(q) = [\frac{q}{\alpha}]$ is the maximal number of possible draws from H that can be taken. We let $q_r \equiv q - \alpha N(q)$ be the residual amount of flexible resource that is left if the firm uses the maximal number of draws within the period. By definition, $q_r \leq \alpha$, so that $\theta^*(q_r, K_0, \theta_0) = \theta_0$. Using (A.6) we find $\theta^*(q_r + n\alpha, K_0, \theta_0)$ recursively for $n = 1, 2, ..., N(q)$. Thus, $\theta^*(q_r + N(q)\alpha, K_0, \theta_0)$ is the initial threshold which determines whether the firm takes at least one draw from H, or implements θ_0 with $q + K_0$.

B. The Bias Correction of the Proposed Estimation Procedure

Here we illustrate graphically the bias introduced by a standard maximum likelihood estimation of the parameter of the distribution of new ideas sampled by the firms, and how our proposed iterative procedure overcomes this problem.[22] The first picture depicts the likelihood function with threshold levels that vary with the distribution's parameter. The data was artificially created with a Pareto distribution with $\lambda = 20$.



The following picture displays the result of the likelihood function constructed

[22]This illustration was computed by our former graduate student, Sivan Frenkel, using a Pareto distribution with a single parameter, $H(\theta) = 1 - \theta^{-\lambda}$.

by the iterative procedure as described in the text. The threshold levels are being held fixed in each iteration, and are re-computed in each iteration until the change is small enough.