Nordic Journal of Political Economy

Volume 25 1999 Pages 87-93

# The Effect of Firm Heterogeneity on R&D Competition

Pedro P.Barros Tore Nilssen

Pedro P. Barros * and Tore Nilssen **

# The Effect of Firm Heterogeneity on R&D Competition ***

There is a large and growing literature providing empirical evidence to the stylized facts that firms' R&D efforts relate positively with their sizes, and that R&D productivity declines with size; see Cohen and Klepper (1996) for a recent account. There are only a few theoretical analyses of these issues and they seem mainly preoccupied with the former relationship: How does differences in size transform into differences in levels of R&D activity? Within the context of various non-tournament models of R&D competition, Rosen (1991), Poyago-Theotoky (1996), and Yin and Zuscovitch (1998) all interpret differences in size as corresponding to differences in *ex-ante*, or pre-R&D, production costs, with low marginal costs implying a large firm. They go on to explore how such a heterogeneity in costs affects firms' R&D efforts in the ensuing R&D competition.

In this paper, we take issue with this way of modelling firm heterogeneity in non-tournament R&D competition. While differences in *ex-ante* costs may be important, we would like to emphasize here the importance of differences among firms in their technological opportunities. Such differences have been discussed at great length in the previous literature. On one hand, Rosenberg (1963) and Scherer (1965) have pointed out the importance of differences in the cost of producing knowledge in order to understand differences in R&D activities among firms, while on the other hand, Schumpeter (1942) emphasizes differences in entrepreneurial ability; see Pakes and Schankerman (1984) for further discussion.

In order to capture the breadth of the above arguments, we introduce heterogeneity in firms' R&D technology in *two* dimensions: Firms differ in their costs of doing R&D, in line with the view of Rosenberg and

* Faculdade de Economia, Universidade Nova de Lisboa, Travessa Estêvão Pinto, P-1070 Lisboa, Portugal, Fax: 351-1-388 60 73, Email: ppbarros@fe.unl.pt

** Department of Economics, University of Oslo, P.O.-BOX 1095 Blindern, N-0317 Oslo, Norway, Fax: 47-22 85 50 35, Email: tore.nilssen@econ.uio.no

*** We are grateful to Tor Jakob Klette for valuable comments on a previous draft. Our collaboration has been facilitated by travel grants from INOVA and Praxis XXI/PCSH/C/CEG/9/96. Nilssen's research is in part financed by the Research Council of Norway through the Foundation for Research in Economics and Business Administration in Bergen.

Scherer. In order to capture the view of Schumpeter, we let, in addition, firms differ in the rate at which R&D activity transforms into cost-savings. Below, we explore how such a double heterogeneity in R&D technology affects the R&D competition among firms. In our companion paper, Barros and Nilssen (1998), we bring our analysis further to explore the consequences of this heterogeneity in a government's optimum R&D policy towards such firms.

It is noticeable that differences in R&D productivity are a key ingredient in one of the stylized facts on which the empirical literature has settled: "Among R&D performing firms, the number of patents and innovations per dollar of R&D decreases with firm size and/or the level of R&D, and among all firms, smaller firms account for a disproportionately large number of patents and innovations relative to their size." [Cohen and Klepper, 1996, Stylized Fact 4, p. 930]

The usual interpretation of this stylized fact is the existence of decreasing marginal returns in R&D activity. Considering two sources of heterogeneity in R&D technology which are not perfect substitutes with each other, we can offer a different explanation. In fact, we are able to argue that the observed regularity is consistent with constant marginal returns to innovation combined with such a heterogeneity among firms in terms of their technological opportunities.

The paper is organized as follows: We present our model in Section 2. Thereafter, we explore its consequences for differences among the firms in R&D efforts in Section 3. We then conclude in Section 4, where we also return to the stylized fact referred to above.

## The model

We set our analysis in the simplest framework possible. Consider a market with demand given by a linear inverse demand function, $P = a - Q$, where $P$ is price and $Q$ is total quantity supplied in the market. On the supply side, there are 2 firms. Let $q_i$ be the quantity produced by firm $i$.

Production technology is characterized by constant returns to scale. The level of the unit production cost depends on the R&D activity performed by the firm. In particular, the unit production cost of firm i is given by: $c_i = \bar{c}_i - \theta_i x_i$, where $\bar{c}_i$ is the initial level of unit production cost and $\theta_i x_i$ is the reduction in cost obtained by firm $i$, depending on its R&D investment, $x_i$, and its productivity, measured by $\theta_i$. A higher $\theta_i$ means a more efficient firm in doing R&D activities, or a firm endowed with more innovative possibilities. As indicated by the subscripts on $\bar{c}_i$ and $\theta_i$, we do not restrict firms to be equal.

Denote R&D costs by $\varphi(x_i, \gamma_i)$, where $\gamma_i$ is a parameter describing the efficiency level of firms. R&D costs are not restricted to be equal across firms. A higher value of $\gamma_i$ means a less efficient firm, with $\varphi(0; \gamma_i) = 0$, $\partial\varphi/\partial\gamma_i > 0$ for $x_i > 0$, and $\partial^2\varphi/\partial x_i \partial\gamma_i > 0$. That is, both total costs and marginal costs increase with $\gamma_i$. Thus, firms' R&D technologies may vary in two different ways: differences in R&D productivity and differences in R&D costs.

Firm $i$ has a profit function given by

$$\Pi_i = (P - c_i) q_i - \varphi(x_i, \gamma_i) \quad (1)$$

One reason for insisting on two dimensions of heterogeneity in R&D technology can be seen by considering the following reparameterization of our model: Let $y_i = \theta_i x_i$ be firm $i$'s R&D decision variable, rather than $x_i$; *i.e.*, let firms decide on R&D output rather than R&D input. Firm $i$'s profit function can now be written (Barros and Nilssen, 1998):

$$\Pi_i = (P(Q) - \bar{c}_i + y_i) q_i - \varphi\left(\frac{y_i}{\theta_i}, \gamma_i\right), \quad (2)$$

where the definition of $c_i$ has been used. Thus, our two sources of heterogeneity in R&D technology can be interpreted as two different sources of heterogeneity in R&D costs: one multiplicative (given by $\frac{1}{\theta_i}$) and another having any form. If, for example, the R&D cost function takes the quadratic form $\varphi = \gamma_i^2 x_i^2 = (\frac{\gamma_i}{\theta_i})^2 y_i^2$, then the effect of $\gamma_i$ cannot be distinguished from that of $\frac{1}{\theta_i}$. In general, however, the two sources of heterogeneity in R&D costs will have different implications. This, of course, also holds for our preferred interpretation of the model, which is in terms of heterogeneity in both R&D productivity and R&D costs, with R&D efforts as firms' choice variables.

We consider the following two-stage game: first, firms decide on R&D investments; and, second, after R&D investments have been made and become common knowledge, firms decide on their production levels. We will be looking for the subgame-perfect equilibrium of this game. This structure of the game is present in most of the papers in this area of research. Thus, our results can be confronted with well-known previous findings in an easy way.

For the sake of completeness, we start out by considering whether initial low-cost firms perform more R&D, or not. This exercise has also been done by Rosen (1991) and Poyago-Theotaky (1996). By repeating the exercise here, we are able to better relate our model to theirs in the case when, like in their models, firms' R&D technologies are equal. We have:[1]

***Remark 1*** *When firms are equal, except with respect to initial costs, the initial low-cost firm does more R&D than the initial high-cost firm.* This remark shows that cost asymmetries in the production stage result in a higher incentive to invest in R&D by the lower cost firm, i.e., R&D gives rise to divergence, with different firms becoming even more different. This result is also obtained by Poyago-Theotoky (1996), under a similar demand structure, a linear cost-reduction function (set as a convex function of R&D effort), and quadratic R&D costs. We have a simpler R&D productivity function but a more general R&D cost function. The intuition behind this result is the following. The low-cost firm produces more in equilibrium. Since R&D reduces the constant marginal cost of producing the final good, an equal marginal R&D outcome is applied to a greater mass of production by the low-cost firm. Therefore, this firm has a higher marginal benefit from R&D, which leads to a higher equilibrium R&D investment and to an increase in production cost asymmetries.

To focus on the role played by R&D heterogeneity, we assume, from now on, no initial asymmetry in the basic cost parameter: $\bar{c}_i = \bar{c}, \forall i$. On the other hand, firms are allowed to differ in their costs of performing a given level of R&D activity and in their abilities to put R&D effort to use. That is, both $\theta_i$ and $\gamma_i$ may differ across firms.

The assumption of a common $\bar{c}$ allows us to write equilibrium quantities produced by each firm, for a given pair of R&D efforts $x_1$ and $x_2$, as:

$$q_i = \frac{\alpha + 2\theta_i x_i - \theta_j x_j}{3}, \tag{3}$$

where $\alpha := a - \bar{c}$. By differentiation of 3, the following comparative-statics results can be seen to hold in the quantity subgame:

1. Here and throughout, we assume that the R&D cost function is sufficiently convex that stability conditions on the equilibrium are satisfied. See Barros and Nilssen (1998) for details on this and for a proof of Remark 1.

$$\frac{\partial q_i}{\partial x_i} > 0, \frac{\partial q_i}{\partial \theta_i} > 0, \frac{\partial q_i}{\partial x_j} < 0, \frac{\partial q_i}{\partial \theta_j} < 0,$$

These results are intuitive ones, as they say that increases in R&D investment or in R&D productivity increase own production and reduce the other firm's production (in the quantity-subgame equilibrium).

Note that these are only the direct effects and they should not be taken as changes in equilibrium values of the full game in the case of the R&D productivity parameter $\theta_i$ (or $\theta_j$), as it is necessary to include the effect through equilibrium choices of R&D efforts.

We are now ready to characterize the R&D competition stage. Without specifying a functional form for $\varphi$, it is not possible to solve explicitly for $x_i$, $i$=1,2. Nonetheless, the following comparative statics can be obtained.[2]

***Remark 2*** *The effects of changes in R&D parameters are:*

$$\frac{\partial x_i}{\partial \theta_i} > 0; \frac{\partial x_i}{\partial \theta_j} < 0; \frac{\partial x_i}{\partial \gamma_i} < 0; \frac{\partial x_i}{\partial \gamma_j} > 0,$$

$i = 1,2; j = 1,2; i \neq j$

From this remark, it is also easy to obtain that

$$\Delta = x_1 - x_2$$

is increasing in $\theta_1$ and decreasing in $\theta_2$.

All these effects are, again, in line with what economic intuition predicts. Increases in R&D productivity (or cost savings) stimulate own R&D investment and reduce investment by competitors (in equilibrium).

## Two sources of heterogeneity in R&D

To explore the implications of two sources of heterogeneity in R&D activities, it is instructive to look first at two special cases. The first case keeps R&D productivity constant across firms ($\theta_1=\theta_2$), while the second assumes the same R&D cost function for both firms ($\gamma_1=\gamma_1$).

Take symmetry as a starting point ($\theta_1=\theta_2;\gamma_1=\gamma_2$). Consider a small increase in $\gamma_1$ (firm 1 becomes less cost efficient in R&D). The productivity of R&D is unchanged and equal for both firms. The marginal cost of R&D is higher for firm 1, which reduces its R&D effort. Strategic substitutability results in a lower R&D effort by firm 1 and an increased effort by firm 2. The difference $\Delta$ becomes negative. Thus,

***Remark 3*** *Holding R&D productivity equal across firms, the high-R&D-cost firm, in the sense of having higher R&D costs for the same level of R&D effort, performs less R&D and has a higher unit cost of producing the final good.*

The first part follows directly from Remark 2. On the second part, it is easy to check that

$$\frac{\partial \tilde{\Delta}}{\partial \gamma_1} < 0, \quad \frac{\partial \tilde{\Delta}}{\partial \gamma_2} > 0 \tag{4}$$

where, in general,

$$\tilde{\Delta} = \theta_1 x_1 - \theta_2 x_2,$$

although, under the condition of remark 3 that $\theta_1=\theta_2=\theta$, we have: $\tilde{\Delta}=\theta\Delta$.

Turn now to the other case, where firms differ only in R&D productivity.

2. The proof of this remark can be found in Barros and Nilssen (1998).

**Figure 1: Relative position of R&D (I)**

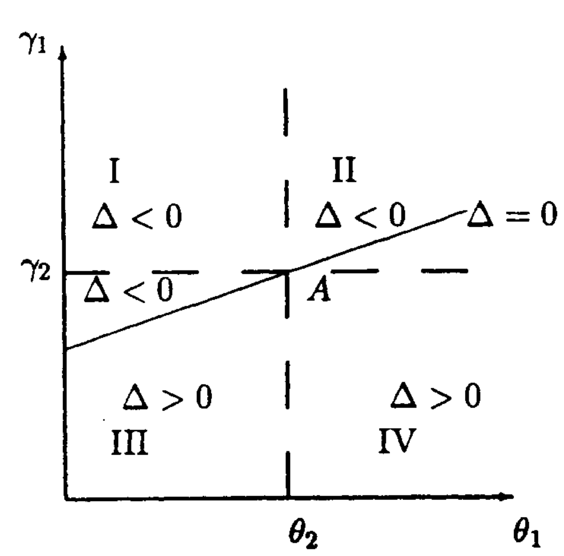
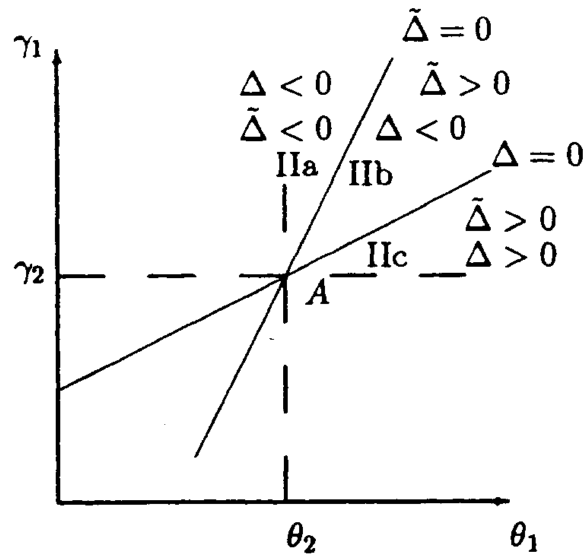



**Figure 2: Relative position of R&D (II)**




***Remark 4*** *Holding the R&D cost function constant across firms, the more R&D-efficient firm (the one with the higher $\theta_i$) performs more R&D, and has lower unit costs of producing the final good.*

The firm that is more efficient in seizing the results of R&D activities invests more in R&D since it gets a stronger cost reduction for the same level of activity.

Putting the two results together, we can say that the more efficient firm in both R&D productivity and costs does more R&D. However, we cannot make a precise statement when a firm is more efficient in one respect but less efficient in the other. This is because heterogeneity of firms in the two dimensions does not allow a well-defined ranking of efficiency across firms. This is illustrated in Figure 1, where we show, conditional on values for $\theta_2$ and $\gamma_2$, the combinations of $\theta_1$ and $\gamma_1$ where a greater effort is done by firm 1, *i.e.*, where $\Delta > 0$.

In region I of Figure 1, firm 1 is less efficient than firm 2, while the reverse is true in region IV. In regions II and III, we cannot make a ranking of firms according to efficiency. The curve $\Delta = 0$ slopes upwards and passes through the symmetry point *A*. To see that this must be the case, start from point *A*. Increasing $\gamma_1$ leads to $\Delta < 0$ and an increase in $\theta_1$ is required to restore $\Delta = 0$. A similar exercise can be done holding $\gamma_1$ constant and increasing $\theta_1$, which leads to $\Delta > 0$.[3]

To answer the question whether, or not, the more efficient firm does a higher level of R&D it is necessary to introduce some notion that combines R&D productivity and cost function differences into a single measure. A natural way out, but by no means the only one, is to look at the achieved cost reduction. This is an *ex-post* efficiency notion, which allows us to state whether the firm with the lower marginal cost of producing the final good is also the one that invests more in R&D, or not. To this end, Figure 2 depicts

3. To show that the slope of the schedule $\Delta = 0$ is positive, take $d\Delta = 0$, keeping all parameters but $\gamma_1$ and $\theta_1$ constant, and rearrange to see that $d\gamma_1 / d\theta_1 > 0$.

the locus of $\tilde{\Delta}$=0. It follows from Remark 3 that this locus is steeper at the symmetry point $A$ than that of $\Delta$ =0.

In Figure 2 we concentrate attention on region II, as region III can be interpreted in a similar way. Of course, the analysis is only valid for combinations $(\gamma_1,\theta_1)$ that are sufficiently close to $A$. Region II is divided into three sub-regions. In the first sub-region, IIa, firms are close to symmetry in R&D productivity but firm 1 has lower R&D costs. The latter inefficiency dominates, and firm 1 both conducts less R&D and has higher unit costs of producing the final good than firm 2. Sub-region IIc is characterized by the opposite situation. The relative efficiency of firm 1 in R&D productivity is high and the disadvantage in R&D costs is small. This leads firm 1 to produce more R&D than firm 2, having also a lower unit cost in the production of the final good.

Sub-region IIb remains, where the R&D cost disadvantage of firm 1 is sufficiently large for this firm to do less R&D than the other firm ($\Delta$<0), despite its higher R&D productivity. However, since firm 1 is more efficient in translating R&D efforts into cost reduction, it has a lower unit cost of producing the final good ($\Delta$>0). We have, therefore, that the *ex-post* efficient firm does less R&D.

Collecting the above results, the following is established.

***Proposition*** *With heterogeneity in both R&D productivity and R&D costs, whenever ordering of the efficiency of the firms' R&D technologies is possible, the more efficient firm does more R&D. If there is no ordering of efficiency available, the firm doing more R&D may, or may not, be the ex-post larger firm.*

## Concluding remarks

The main purpose of this note has been to emphasize how differences in R&D technologies may help explaining differences in R&D activities. We have shown that, in an industry where firms differ in their R&D technologies, the firm doing the more R&D may or may not end up as the more cost efficient firm.

This result should be contrasted with previous theoretical work on R&D competition among heterogeneous firms, like Rosen (1991) and Poyago-Theotoky (1996). These authors stress differences in *ex-ante* production costs and find that the firm that is low-cost *ex-ante* also ends up being the low-cost firm *ex-post*. Although the above Proposition is stated under the condition that initial costs are the same, it follows by continuity from this result that differences in R&D technology may turn the above result of the previous literature on its head: The *ex-ante* high-cost firm may, because of differences in R&D technology, or technological opportunities, across firms, end up as the low-cost firm; and this may happen even if it does less R&D than the *ex-ante* low-cost firm.

Let us now return to the stylized fact quoted in the Introduction. First, note that R&D output per dollar invested, in our model, equals $\theta x/\varphi$, where subscripts have been dropped for convenience. The stylized fact states that this ratio varies with the level of R&D, which is $x$ in our model. The implication of the stylized fact, in terms of our model, is thus that the ratio $\theta/\varphi$ differs across firms. Whenever this is the case, we would argue, it is not appropriate to model firms as having identical R&D technologies, as has been done in the above-mentioned work of Rosen (1991) and Poyago-Theotoky (1996).

Suppose differences in the $\theta/\varphi$ ratio is mostly due to differences in the R&D-productivity parameter $\theta$, and recall that we have established, also for our model, that firms with lower initial production costs perform more R&D (Remark 1). The stylized fact of Cohen and Klepper (1996) relates differences in the $\theta/\varphi$ ratio to differences in size. If large firms tend to have a low R&D productivity and/or higher costs of performing R&D, the empirical regularity they report can be obtained in our model with constant marginal returns to innovation and heterogeneous firms. The regularity then captures points of different curves (associated with distinct technologies) relating $\theta x/\varphi$ with $x$ (or size), and not movements along the same curve, as has been considered in previous literature. Thus, diminishing returns to R&D may have been largely over-estimated, due to omission of the role of R&D heterogeneity.

This explanation for the stylized fact clearly shows the potential role of firm heterogeneity, at least, as an alternative to significant decreasing marginal returns to innovation effort. It places a call for future empirical research aimed at distinguishing the two explanations.

## References

Barros, P.P. and T. Nilssen, 1998, Industrial policy and firm heterogeneity, Discussion Paper 1986, CEPR, London.

Cohen, W.M. and S. Klepper, 1996. "A reprise of size and R&D", *Economic Journal,* 106, 925-951.

Pakes, A. and M. Schankerman, 1984. "An exploration into the determinants of research intensity", in: *R&D, Patents, and Productivity* (Z. Griliches, ed.), Chicago: University of Chicago Press, 209-232.

Poyago-Theotoky, J., 1996. "R&D competition with asymmetric firms", *Scottish Journal of Political Economy* 43, 334-342.

Rosen, R.J., 1991. "Research and development with asymmetric firm sizes", *RAND Journal of Economics* 22, 411-429.

Rosenberg, N., 1963. "Technological change in the machine tool industry, 1840-1910", *Journal of Economic History* 23, 414-443.

Scherer, F.M., 1965. "Firm size, market structure, opportunity, and the output of patented inventions", *American Economic Review* 55, 1097-1125.

Schumpeter, J.A., 1942. *Capitalism, Socialism, and Democracy,* New York: Harper.

Yin, X. and E. Zuscovitch, 1998. "Is firm size conducive to R&D choice? A strategic analysis of product and process innovations", *Journal of Economic Behavior and Organization* 35, 243-262.