Innovation and competition in the Netherlands: testing the inverted U for industries and firms

10

Michael Polder and Erik Veldhuizen

The views expressed in this paper are those of the author(s)
and do not necessarily reflect the policies of Statistics Netherlands

Discussion paper (201021)

Statistics Netherlands

The Hague/Heerlen, 2010

Explanation of symbols

.	= data not available
*	= provisional figure
**	= revised provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2008–2009	= 2008 to 2009 inclusive
2008/2009	= average of 2008 up to and including 2009
2008/’09	= crop year, financial year, school year etc. beginning in 2008 and ending in 2009
2006/’07–2008/’09	= crop year, financial year, etc. 2006/’07 to 2008/’09 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands - Grafimedia

Cover
TelDesign, Rotterdam

Information
Telephone +31 88 570 70 70
Telefax +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax +31 45 570 62 68

Internet
www.cbs.nl

ISSN: 1572-0314

600830201021 X-10

Innovation and competition in the Netherlands: testing the inverted U for industries and firms

Michael Polder and Erik Veldhuizen[a]

Abstract

Competition can be good or bad for innovation by firms. On the one hand it stimulates firms to innovate in order to escape competition, on the other hand it hampers firms to reap additional profits from innovation. The recent literature has embraced a model that describes an inverted U-shape relationship between competition and innovation at the industry-level. With the Price Cost Margin and Profit Elasticity as measures of competition, we find evidence supporting this prediction using industry data from the Dutch National Accounts. Moreover, we test the non-linear relation at the micro-level, with special attention for the role of the distribution of technology within industries. We find evidence that there is a threshold for this 'technology spread' where the (marginal) effect of competition on innovation activity by firms turns from positive to negative.

Keywords: competition, innovation, R&D

[a] Both authors are affiliated at Statistics Netherlands. We thank Dirk van den Bergen, George van Leeuwen, and Leo Soames for helpful discussions and comments on the draft version.

Corresponding author: jplr@cbs.nl, tel. +31 70 337 5138.

1. Introduction

The effect of competition on economic activity has been at the centre of debate for quite some time (see Ahn, 2002, for an overview). As noted by Nickell (1996), it is a common belief that competition stimulates firms to improve their performance and distinguish themselves from competitors to attract demand for their products. On the other hand, in line with the more traditional Schumpeterian view, a market where competition is very stiff may offer little room for innovative activities, so that in the long-run competition may negatively affect productivity growth. Thus, the relation between competition and innovativity or productivity may not be unequivocally positive.

In line with these two contrasting views on the impact of competition, Aghion et al. (2005) suggest a theoretical model where the relation between competition and innovation is non-linear at the industry level, which they refer to as an inverted U-shape.[1] They also present empirical results supporting the predictions of their model. Thus, if competition is intensified starting from a low level, innovation activity is stimulated, whereas starting from a high level of competition, innovation activity is discouraged by increased competition. Aghion et al. suggest that this is due to a composition effect in the dispersion of the production technology.

In this paper, we investigate the relationship between innovation and competition using firm- and industry-level data for the Netherlands.[2] The paper consists of two parts: a macro-study and a micro-study. In the macro part we test for an inverted U relationship between industry-level R&D expenditures and competition. In the micro part, we focus on whether the non-linear mechanism can be explained by differences in the distribution of the production technology as predicted by Aghion et al. model.

We make a few contributions. Firstly, since the argument for an inverted U relationship provided by Aghion et al. is for the industry-level, most of the existing empirics is also at the industry level. We test for a non-linear effect at both the industry- and the firm-level. Moreover, since the theoretical model predicts that the effect of competition on innovative activity should depend on the distribution of technology, we test explicitly for the role of this 'spread'. Secondly, besides the Price Cost Margin (PCM) we employ a relatively new competition measure called the Profit Elasticity (PE) as proposed by Boone (2008). Thirdly, our dataset is relatively large and rich compared to most of the existing empirical work, with micro-data being sourced from surveys by the statistical office, and industry-level variables being the official figures published in the Dutch National Accounts and R&D satellite accounts. Finally, our dataset includes services and other non-manufacturing industries. This meets the interest in measuring the effects of competition in industries outside manufacturing, on which this type of studies has usually focussed. Especially in the

[1] Additional argumentation and an overview of related papers can be found in Aghion and Griffith (2008).

[2] Our results serve as a complementary country specific analysis in the project 'Market incentives to innovation' of the Working Party for Industrial Analysis (WPIA) of the OECD.

Netherlands, but typically in Western economies, a shift towards services can be noted, making it necessary to go beyond manufacturing in empirical analyses.

The setup of the paper is as follows. Section 2 briefly discusses the theoretical arguments for the inverted U-shape, but also presents some caveats and possible counterarguments. It also reviews some related empirical literature. Section 3 puts down the estimation equations for both the industry- and firm-level estimation and discusses the econometric approach. Section 4 describes the data, discusses the measurement of competition and innovation, and presents summary statistics. Section 5 presents the estimation results. Finally, section 6 provides a summary of the conclusions and avenues for further research.

2. Theoretical background

2.1 Why should there be an inverted U-shape?

The frequently cited Aghion et al. (2005) study suggests that the relation between innovation and competition may be non-linear. The theory behind this is that the level of competition affects the distribution of production technologies in an industry.[3] We will refer to the dispersion in this distribution as the 'technology spread'. A 'levelled' industry has a low technology spread, whereas an 'unlevelled' industry has a high technology spread. The technology spread changes due to innovation. In the Aghion et al. (2005) model there are two possibilities for this: if a laggard in an unlevelled industry innovates, it catches up with the leader and the industry becomes levelled; if a firm in a levelled industry innovates, it becomes the leader, and the industry becomes unlevelled.

In a levelled industry, an increase in competition is positive for innovation since it stimulates firms to escape competition. In an unlevelled industry, innovation by the laggard decreases with competition, because higher levels of competition will prevent it from reaping the profits of innovation, the so-called Schumpeterian effect.[4]

---

[3] Chapter 3 of Aghion and Griffith (2008) presents a different model that also results in an inverted U relationship but under quite different assumptions. See also Vives (2004) and Rauch (2008) for other models giving rise to an inverted U. Each of these models has its own implications for empirical testing, so that our results do not necessarily relate to these models as well.

[4] By construction the leader does not innovate in an unlevelled industry (Aghion et al. 2005, p. 713). Or, more precisely, it can innovate but the laggard will take over the old leading technology. Since innovation occurs step-by-step in the model, the new situation is the same as the old, so the leader has no incentive to innovate. (The leader does not seem to take into account the possibility of catching-up of the laggard, which could provide an incentive to stay one step ahead. In fact, an increase in competition decreases the threat of a laggard catching up. A model in which such strategic innovation decisions are considered is presented in Aghion and Griffith, 2008, chapter 3, but in the context of an incumbent firm versus a potential entrant.)

The inverted U-shape relation between innovation and competition then results from the idea that under low competition industries will (most of the time) be levelled (so that increases in competition are good for innovation), whereas under high competition, industries will (most of the time) be unlevelled (so that increases in competition are bad for innovation).

In other words, the inverted U results from what Aghion et al. call a "composition effect" of levelled versus unlevelled industries. Low competition is associated to levelled industries, since when the industry is unlevelled, there is a relatively high incentive for laggards to catch up with the leader and level the industry (as mentioned above, by assumption, the leader does not innovate). The incentive to become the leader in a levelled is relatively small when competition low, however. So, when competition is low, an industry will be quick to leave the unlevelled state, but slow to leave the levelled state. On the other hand, under high competition, firms in a levelled industry have a relatively high incentive to innovate and escape competition. Laggards in an unlevelled industry will be discouraged to innovate in a high-competition industry, however. So, on balance, high competition industries will be unlevelled.

In summary, the escape competition ('positive') effect of a change in competition on innovation dominates when the initial level of competition is low, whereas the Schumpeterian ('negative') effect dominates at initially high levels of competition. This gives rise to an inverted U-shape relation between the two variables.

2.2 Why should there *not* be an inverted U-shape?

The Aghion et al. model is embraced as a synthesis between the Schumpeterian argument that on the one hand firms should have some form of market power to innovate, and the findings in many empirical studies that show that competition correlates positively with innovation on the other hand. It is worthwhile to emphasize, however, that there are reasons, both theoretical and empirical, why the *observed* relation between innovation and competition may not conform to an inverted U.

Even if the theoretical model is correct, we may not be able to identify an inverted U. The identification depends on the empirical distribution of the industries over the presumed inverted U. Most empirical studies are restricted to a single country. If industries in a certain country are concentrated on either the increasing or decreasing slope of the curve, we cannot identify whether there is a certain threshold for the level of competition that would reverse the effect of a change.[5] Thus, to be able to identify an inverted U, one needs to have a sample where competition is both high

---

[5] In this context it is interesting to look at the non-parametric evidence summarized in figure II of Aghion et al. When we look at the spline it looks more like an M than an inverted U. In particular, if one disregards the outer tails of the M – which have a relatively low number of observations according to the figure – the resulting shape is convex rather than concave.

and low.[6] A reason why all industries may be subject to either low or to high levels of competition, is that the amount of competition is affected by the institutional setting of a country. The attitude towards market competition is for example far different in countries with historically a more socialistic background (e.g. France) than countries that are more liberal with respect to their market policies (e.g. the U.S.).

On the other hand, we could identify an inverted U even if it is not there. This happens if the observed relationship is sufficiently concave to fit a negative quadratic term, but the implied 'turning point' (i.e. the maximum of the parabola) is not in the data. The conclusion of an inverted U is spurious in such a case. Therefore, as noted also by Lind and Mehlum (2010), even if a negative quadratic form can be fitted to a particular relationship, one should check if the implied threshold is within the range of the data. Moreover, note that the theory does not predict that the relation is actually quadratic. It tells us that there should be a positive and a negative slope, which does not rule out the possibility of other functional forms.

A final potential empirical pitfall is the usual lack of information on what is exactly the firm's market. In general, the market is set equal to the firm's industry, and the level of detail at which industries are classified is restricted by the number of observations. Not all firms in the same industry compete with each other, especially when higher levels of classification are used. Moreover, if firms with different levels of efficiency belong to the same industry but not to the same market, the industry may well be classified as unlevelled, while in fact the actual markets could be levelled.

There can also be some theoretical reasons why the inverted U may not arise. Although it is beyond the scope of this text to formalize these points, we provide some tentative caveats. As noted in footnote 4, the leader in an unlevelled industry does not innovate. This assumption is quite crucial in generating the predictions of the model. It does not allow for the possibility that firms may have strategic incentives to innovate. If the laggard catches up, the leader is likely to loose profits, and more so if competition increases. On the other hand, the probability that the laggard innovates decreases with competition. But if the latter effect is relatively small compared to the increase in the loss in profits to the leader when it *does* enter, a rise in competition actually increases the expected loss of profit. The leader then has an incentive for *strategic* innovation: because it knows that the laggard may catch up it will innovate to maintain its leading position and a larger profit. In this case the unlevelled industry will remain unlevelled, in contrast to the prediction.

Another argument that may contradict an inverted U is the existence of sunk costs. In the case of high competition, the model says that if one firm manages to innovate and escape competition, another firm will in general not find it worthwhile to catch up due to low post-innovation rents. However, prior to the innovation, the increase

[6] We focus on the identification from a cross-sectional or panel data set. Similar caveats apply to the identification of the inverted U by industry from a time-series, however. One then needs a sufficiently long time-series with observations for the industry distributed well over the inverted U. That is, periods of low and high competition must have occurred.

in competition provided *both* firms with an incentive to innovate and both firms will therefore pay the innovation cost (e.g. investment in R&D). When one of the firms succeeds in the innovation earlier than the other, the innovation costs become sunk for the non-innovator. These costs will not be taken into consideration by the firm in the decision to continue an innovation program or not. Assuming that the innovation by the competitor does not completely pre-empt the market, the lower the additional costs to complete the innovation are,[7] the more likely it is that the firm decides to complete its innovation program and level the industry, again in contrast to the prediction for this scenario. This argument is reinforced if abolishing the innovation program is actually costly, for example when external financers are involved or when stopping the program involves the displacement of workers.

A final caveat concerns the two-firm nature of the theoretical model. By construction, the unlevelled industry can only have one leader. In practice, an industry consists of many firms, and in an unlevelled industry some of them are leaders and some laggards. It is then possible that the leaders will still try to escape the competition of their fellow-leaders, in contrast to the line of reasoning in the model.

3. Some micro-econometric work on the inverted U

There is a vast literature on competition, and its relation to innovation and firm performance, see e.g. Ahn (2002) for an overview and references. To keep the discussion brief, we restrict ourselves to the empirical literature known to us that deals with the possibility of a non-linear relationship between competition and innovation or firm performance. After discussion of the main results by Aghion et al. (2005), we focus on the studies that investigate these relations at the micro-level.

The seminal work in this respect is the paper by Aghion et al. (2005). Based on an early suggestion by Scherer (1967), they find a hump-shaped relationship at the industry-level between competition (measured by 1 minus the industry average of the price-cost margin) and count data of (citation weighted) patents. They develop a model where escape competition dominates when firms are neck-and-neck (i.e. in a levelled industry), and Schumpeterian competition dominates when the technology spread is characterized by leaders and laggards (i.e. unlevelled industries). As explained above, they explain their findings by the argument that at a relatively low level of competition, the industry will most of the time be characterized by neck-and-neck firms, while under high competition it will be characterized by leaders and laggards. Thus, the empirical weight is at the positive slope when there is low competition, and at the negative slope when there is high competition.

Since the work by Aghion et al. the empirical literature has been hunting for the inverted U-shape. Studies using micro-data as in our case are yet relatively rare,

[7] In the IO models of competition by Salop (1977) and Dixit and Stiglitz (1977), for example, the innovation cost is fixed, so that additional outlays are in fact zero.

however. Poldahl and Tingvall (2006) use Swedish firm-level data for the manufacturing sector, and find evidence for an inverted U when measuring competition by the Herfindahl index. However, when using the Price Cost Margin (PCM) they find a negative relationship. In addition, when controlling for firm-specific effects, they do not arrive at a significant relationship between R&D and competition. Tingvall and Karpaty (2008) do a similar analysis for Swedish firms in the services sector. They use the Herfindahl and the Profit Elasticity as competition measures, and control for selectivity in the CIS R&D data, but do not use firm fixed effects. They find evidence for an inverted-U relationship, with the exception of non-exporting firms. Extramural R&D does not fit the inverted-U pattern, but intramural R&D and expenditures on the training of employees do.

Askenazy et al. (2008) suggest a role for the cost of innovation. They find evidence for an inverted U for large French firms, but not for the full sample. When controlling for innovation costs (e.g. the cost of patenting), the results suggests that the relation between innovation and competition becomes flatter as these costs increase. If innovation costs are relatively high compared to value added, changes in competition become less important in the innovation decision.

Friesenbichler (2007) reports evidence for a bell-shaped pattern for the Austrian mobile phone industry. The study uses the Herfindahl index as a measure for competition (obtained from the Austrian competition authorities), and R&D as well as a service innovation indicator as innovation measures. Although the sign on the quadratic Herfindahl is negative, pointing at an inverted U, the study only provides a test of the joint significance of the linear and quadratic term. Bos et al. (2009) find an inverted U for firms in the American banking sector using the PCM, and further conclude that in the time period considered, the level of competition of the industry exceeds the optimal level, considering the fact that deregulation has lowered innovation rates in this sector.

4. Empirical approach

4.1 The industry-level analysis

To test for a non-linear relationship between R&D expenditures at the industry-level and competition, we firstly estimate the following equation:

(1) $$\log(RnD_{jt}/VA_{jt}) = \alpha_j + \lambda_t + \beta_1 \cdot COMP_{jt} + \beta_2 \cdot COMP_{jt}^2 + \varepsilon_{jt},$$

where j indexes the industry and

RnD_{jt}	total R&D expenditures (see section 5.1);
VA_{jt}	value added of industry j;
$COMP_{jt}$	competition in industry j (see section 5.2);
α_j	industry dummy (i.e. industry-specific constant)

λ_t year dummy;

ε_{jt} disturbance.

If the relationship is an inverted U, we should find $\beta_2 < 0$. The 'optimal' level of competition $COMP^*$ (i.e. the point where the effect of a change in competition on R&D turns from positive to negative) can be found by solving

$$\partial\log(RnD/VA)/\partial COMP = \beta_1 + 2\beta_2 COMP = 0 \Rightarrow COMP^* = -\beta_1/2\beta_2.$$

Note that since $COMP > 0$, we also need $\beta_1 > 0$ for the $COMP^*$ to be positive. In addition, to exclude possible spurious conclusions about at a turning point, $COMP^* \in (COMP_L, COMP_H)$, where $COMP_L$ and $COMP_H$ are respectively the minimum and maximum level of competition in the sample.

4.2 The firm-level analysis

At the firm-level, to test more directly the mechanism that produces the non-linearity as suggested by Aghion et al. and described above, we estimate

$$(2) \qquad \log(RnD_{it}/VA_{it}) = \alpha_i + \beta_1 COMP_{jt} + \beta_2 \text{SPREAD}_{jt} \times COMP_{jt} + \beta_3 X_{it} + \varepsilon_{it},$$

where SPREAD_{jt} is a measure for the distribution of the production technology (see section 5.3), and X_{it} is a vector including firm size (log employment, EMP_{it}), capital intensity ($\log(K_{it}/EMP_{it})$), firm distance-to-the-frontier (DTF_{it}) and year dummies.[8]

The interaction term of the technology spread with competition allows for a non-linear marginal effect of competition.[9] If $\beta_1 > 0$ and $\beta_2 < 0$, the sign of the marginal effect changes from positive to negative as the spread in the industry becomes larger, i.e.

$$\partial\log(RnD/VA)/\partial COMP = \beta_1 + \beta_2 \text{SPREAD}_j > 0 \Leftrightarrow \text{SPREAD}_j < -\beta_1/\beta_2$$

$$\partial\log(RnD/VA)/\partial COMP < 0 \Leftrightarrow \text{SPREAD}_j > -\beta_1/\beta_2,$$

with $-\beta_1/\beta_2 > 0$. In words, if the spread is low the value of the interaction term is not high enough to neutralize the positive effect of the single competition term in (2), and so there is an increase in innovation on balance. When the spread is bigger and exceeds the threshold value of $\tau \equiv -\beta_1/\beta_2$, the marginal effect turns from positive to negative. Note that a difference with the threshold that can be calculated from the quadratic industry-level equation is that τ concerns the value of the spread, rather than the level of competition.

[8] Industry dummies are not included explicitly in X since we use within-regression. In principle, this implicitly takes account of industry-specific effects (i.e. the industry dummies drop out in the within-transformation). Note that, if included, the coefficients on the industry dummies are identified through observations where firms have switched industry only, so that the pertinent industry dummy has some time variation.

[9] This measure for spread is calculated at the 3-digit level. Although not pursued in this study, it is also possible for the macro regression to interact competition with a measure of spread at the P42 level.

We test the prediction of the model with respect to the role of the technological spread more directly than is the case in the empirical work cited in section 3, where the notion of the inverted U-shaped relationship has been translated quite literally from the Aghion et al. industry-level model to the firm-level. The non-linearity there is modelled by a quadratic competition term, yielding an equation like (1) for firms. Because the role of the technology spread in the reaction to a change in competition is ignored, evidence for non-linearity is consistent with Aghion et al. but not necessarily due to the mechanism they suggest.[10]

4.3 Estimation

In both the estimation of (1) and (2) we take into account the panel structure of the data. We estimate the industry-level specifications by OLS, controlling for industry fixed effects using industry dummies. The firm-level specifications are estimated by fixed effects (within) regression to allow for a firm-specific effect α_i. Accounting for fixed effects controls for any omitted variables that are constant over time, and also for any time-invariant measurement and other additive specification errors. If a variable is endogenous in the sense that it is correlated with the overall error term $\varepsilon_{it} = \alpha_i + u_{it}$, it can be argued that controlling for the fixed effect mitigates the endogeneity bias in the estimated coefficients if the correlation is mainly through the firm-specific effect (see Mundlak, 1963, and Olley and Pakes, 1996).[11] In addition, to check the sensitivity of our results to endogeneity bias, we also estimate specifications with lags for potentially endogenous variables.

5. Measurement issues and data

Our firm data are sourced from various surveys by Statistics Netherlands, and span the years 1997 to 2006. The availability of independently measured time-series and the restriction to use only final year data limits the data for the industries to the years 1999 to 2006. Those are taken from the Dutch National Accounts. In both micro- and macro-data, manufacturing industries (13 in total) as well as non-manufacturing industries (21) are selected.[12] For the macro-dataset this results in 234 observations. From the micro-data we have an estimation sample with well over 6000 firms and around 14,000 observations, depending on the specification used.

---

[10] As mentioned in footnote 3 there are other models yielding an inverted U-shape relation.

[11] Technically speaking, this is true if $\mathrm{Cov}(W_{it},\varepsilon_{it})$ is small relative to $\mathrm{Cov}(W_{it},\alpha_i)$.

[12] The selection of industries corresponds to what in the Dutch Growth Accounts is called the ‘commercial sector’ (NACE codes 01 to 67, 72-74, 804, and 85-93), see CBS (2009). Research and development (NACE 73) is excluded since it is highly atypical in this context. We also exclude Mining and quarrying (NACE 10-14) because the data do not allow to calculate meaningful competition measures for this industry.

Firm-level costs of capital (*K*), labour (*L*), and intermediate inputs (*IV*), as well as the number of employees (*EMP*, in full-time equivalents) are taken from the Production Statistics (PS).[13] Some industry variables (PCM, PE and average DTF) were calculated using the full PS dataset, as described below. The estimation sample is smaller, mainly because of the linking to the R&D data. The latter data are taken from the R&D and innovation surveys (the former for odd years, the latter for even years). Given that the sampling provides full coverage of larger firms, but only partial coverage of smaller ones, our results are possibly biased towards larger firms as is common in micro-data studies.

Compared to other studies our dataset is quite rich. In particular, the Aghion et al. study uses the average number of (citation weighted) patents by industry to measure innovation. Their number of observations calls into question whether the industry patent variable can be representative, however. Without knowledge about the statistical properties of the sample with respect to population totals, it is difficult to obtain representative industry averages. By contrast, we use as our industry innovation measure the official R&D figures sourced from the R&D satellite accounts, which are consistent with other National Accounts data and derived from underlying micro-data drawn from a stratified sample. Moreover, the Aghion et al. data are sourced from the London Stock Exchange (LSE). This sample only contains quoted firms, of which the innovative and competitive behaviour is likely to be different from that of non-quoted companies, due to for example the fundamental differences in the type of ownership and management of these companies. In addition, since the statistical properties of the sample are again unknown, it is not clear a priori if industry measures derived from these data are representative. Our study uses Structural Business Survey data from Statistics Netherlands, which consists of a census for firms above 50 employees, and a stratified sample (including raising factors) for smaller firms. This allows us to create appropriately weighted industry figures.

5.1 Measuring innovation

Innovation is a broad and abstract concept. Measures that are used in the literature include patent behaviour, Research and Development (R&D), and survey questions from innovation questionnaires as for example the estimated percentage of sales attributed to innovative products or binary variables for whether a particular type of innovation was carried out by the firm or not. In this paper we focus on R&D investments, especially because of the availability of this variable in the R&D satellite accounts. Hence, we are focussing on innovation efforts (or input), rather than on the outcome of the innovation process (output). We therefore investigate the effect of competition on R&D investment, but not its effect on the *success* of these investments. It seems more natural to consider competition primarily as an incentive to employ innovative activities rather than as a driver of innovative success.

[13] Costs of capital are proxied by total depreciation costs.

Annual R&D micro-data are sourced from two surveys: the Dutch version of the Community Innovation Survey (CIS, even years) and the R&D survey (odd years). The measure we use is the sum of intramural and extramural R&D expenditures. In even years we only observe R&D for technological innovators.[14] The firms that state to have own R&D personnel in the even years are questioned on their R&D in odd years. Thus, in odd years we only observe R&D for firms with own R&D (in the previous year).

R&D macro-data are obtained from the R&D satellite accounts of the Dutch National Accounts. These have recently been developed to anticipate the capitalisation of R&D as recommended by the new System of National Accounts (United Nations, 2008). The main sources of the R&D macro-data are the CIS and the R&D survey. R&D capital expenditure is obtained by translating the gross expenditure on R&D (by producer and by supplier of funding) from these surveys. The translation process comprises several steps including the revaluation of the R&D expenditure data in order to obtain R&D output according to SNA guidelines and the elimination of overlaps with software development. A more detailed description of the methods used to estimate R&D capital expenditure in the Dutch R&D satellite accounts is given by van Rooijen-Horsten et al. (2007, 2008) and Tanriseven et al. (2007).

Since we use the ratio of R&D to value added, we do not use a deflator to convert R&D into real terms (thereby implicitly assuming that the deflator is – approximately – equal to that of value added).

5.2 Competition indicators

We discuss various indicators of competition in an earlier paper (Polder et al. 2009). In this study we use two of these indicators, namely the Price Cost Margin (PCM) and the Profit Elasticity (PE, sometimes loosely referred to as the Boone indicator, see Boone, 2008). In short, the idea behind the PCM is that under competitive pressure firms are less profitable, so that profit margins are negatively related to competition. Because of this inverse relation, we use one minus the PCM as an indicator of competition. The PE is a measure of the relation between profitability and inefficiency. If firms are punished more harshly for being inefficient in competitive markets, we can derive a measure of competition from the strength of the negative effect of inefficiency on profitability. Boone et al. (2007) show that the measure is monotonously related to various competition parameters, unlike other commonly used measures among which the PCM.

As in Polder et al. (2009), we use three versions of the PCM. Introducing *i* to index firm-variables, and *j* for industries,

(PCM-1) $PCM_{it} = (Y_{it} - VC_{it})/Y_{it}$,

[14] An innovator is a firm that has developed a product or a process innovation, or has a project aimed at this that is still ongoing or that has been abandoned at an earlier stage. The innovation can have been developed only by the firm itself, in cooperation, or entirely by third parties.

where

$VC_{it} = L_{it} + IV_{it}$	total variable cost, the sum of labour cost and that of intermediate inputs
Y_{it}	total value of production.

(PCM-2) $PCM_{jt} = \sum_{i \in j} w_{it} PCM_{it}$

where $w_{it} = Y_{it}/Y_{jt}$ and $Y_{jt} = \sum_{i \in j} \alpha_{it} Y_{it}$, α_{it} being the raising factor associated with firm i.

For both (PCM-1) and (PCM-2) on the basis of micro-data we use the Production Statistics (PS). For the industry PCM calculated from the firm data (PCM-2) we include imputations and corrected observations, where as we normally use only data that is based on actual response by firms. The reason is that these adjustments are made by the statistical office to come to the proper industry totals, which is what we want in this case as well. We use (PCM-1) and (PCM-2) in the firm-level regressions. Note that (PCM-2) is the measure used by Aghion et al (2005).

For the industry-level regressions, we calculate the PCM directly from National Accounts industry data:

(PCM-3) $PCM_{jt}^{(\mathrm{NR})} = (Y_{jt}^{(\mathrm{NR})} - VC_{jt}^{(\mathrm{NR})})/Y_{jt}^{(\mathrm{NR})}$.

Following Boone et al. (2007), we estimate the PE from the regression:

$$\ln(\pi_{it}) = \alpha_i - \beta_{jt} \cdot \ln(VC_{it}/Y_{it}) + \lambda_t + \mu_j + \varepsilon_{it} \qquad \forall j,t$$

where π_{it} is $Y_{it} - VC_{it}$, λ_t and μ_j are respectively year dummies and industry dummies, and β_{jt} is the PE. The equation is estimated using only data based on response. We perform a separate within-estimation for all firms in each industry j to control for the fixed effects and also include the log of employment to control for size. To allow for β_{jt} to vary per year, the explanatory variables are interacted with the year dummies. For the use in micro-data regressions we estimate the PE at the three-digit level; for the use in the industry-level analysis we estimate the PE at the publication level of the Dutch Growth Accounts (which divides the economy into 42 industries). The estimations are done with data from the PS, using only observations based on actual response. For each industry, a minimum of 20 observations per year is used, otherwise the PE is set to missing for that industry in that year.

For industries that are not covered by the PS (e.g. some services industries like financial institutions, and agriculture) we cannot estimate the PCM or PE. Note that the variables underlying these indicators are not deflated, since the idea of the indicators is that the development of competition is partly reflected in changes of prices over time.

5.3 Distance-to the frontier

We follow the index approach in Jorgenson and Griliches (1967) to define the level of Total Factor Productivity (TFP) as

(3) $TFP_{it} \quad = RY_{it}/RX_{it}$

$$= RY_{it}/(RK_{it} + RL_{it} + RIV_{it})$$

where X_{it} is total factor costs and K_{it} is capital cost, which we proxy by depreciation. These variables are again taken from the PS. The R denotes that these nominal values are deflated with the appropriate deflators.[15] This definition of the TFP level is consistent with the Dutch Growth Accounts in that the formula for the growth rate of TFP used there equals the growth rate implied by equation (3).

The distance-to-the-frontier is then defined as the difference between the TFP of a firm and the (national) technological leader in its industry (i.e. the most productive firm). To avoid sensitivity to outliers we use the 95th percentile instead of the actual observed maximum by 3-digit industry. That is,

$$DTF_{it} = (TFP_{Fjt} - TFP_{it})/TFP_{Fjt},$$

where TFP_{Fjt} is the 95th percentile of the TFP distribution in industry j in year t. For the observations in the upper 5% of the distribution, where $DTF_{it} < 0$, DTF_{it} is set to zero and are thus considered to be on the frontier.

We also use the DTF to measure the spread in technology within an industry by taking industry averages:

$$\text{SPREAD}_{jt} = \frac{1}{n_{jt}} \sum_{i \in j} DTF_{it} ,$$

where n_{jt} is the number of firms in industry j in year t.

5.4 Summary statistics

The tables 1a and 1b at the back of the paper provide summary statistics for the main variables in the analysis. These tables use the industry classification used in the Dutch Growth Accounts, which is also the level of the industry analyses. It is important to note that the summary statistics refer to the estimation samples.[16] This explains why the averages of the industry variables calculated from the micro-data (based on the PS full sample) are somewhat different from the firm averages here. The estimation sample is substantially smaller because of the availability of the R&D data. The difference is never large enough to raise any concerns about comparability however. Note also that the ratio of R&D to value added is in general a lot higher in the firm data. The reason is that in the micro-data we only look at R&D

---

[15] We use deflators for gross output, value added, and intermediate inputs from the EUKLEMS-database (see www.euklems.net). Deflators for capital and labour cost are derived at the 2-digit level from labour costs and user-cost of capital figures in current and constant prices (unpublished estimates made for purposes of the Dutch Growth Accounts, recalculated to the 2-digit level, which is the lowest level available for these price indices). In all cases base year 2000 was chosen.

[16] Because the various specifications for the firm-level regressions do not exactly have the same number of observations we use the largest estimation sample, which is for the specification with 1-PCM in year t as competition measure.

performers, whereas in the industry data we normalize on industry value added, which also includes that of the non R&D performers.

6. Results

6.1 Industry-level regressions

Table 2 presents the OLS regression results on industry-level data for equation (1). As explained above, competition is measured as 1 minus the industry PCM (equation (PCM-3) in section 5.2) and the PE. Besides results with competition measured in the same year as the dependent variable, we also estimate the model using a one year lag to check the robustness with respect to the timing assumption. It may be that changes in competition affect R&D with a lag. Moreover, while as mentioned in section 4 the inclusion of industry dummies to some extent protects the results from endogeneity bias, the lagged specification is less vulnerable to this type of bias and we therefore use it also as a robustness check.[17]

Table 2. Industry-level regression results.

	year t		year $t{-}1$	
$1{-}PCM$	-20.90		-31.20	
	(20.2)		(20.3)	
$(1{-}PCM)^2$	15.47		19.73	
	(12.7)		(12.3)	
PE		0.747^{**}		0.747^{**}
		(0.35)		(0.32)
PE^2		-0.0469^{**}		-0.0524^{***}
		(0.021)		(0.021)
No. of obs.	234	164	234	159
R^2	0.88	0.92	0.88	0.92

***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.10$.

Estimation method is OLS. Robust standard errors are reported in parentheses. Dependent variable is log R&D over value added. The regressions also contain year and industry dummies not reported.

For the year t results, competition is not significant when 1−PCM is used as the indicator of competition. With the PE, however, both the linear and the quadratic term of the PE are significant. Moreover, since $\beta_1 > 0$ and $\beta_2 < 0$ the results are consistent

[17] We have experimented with instrumental variable techniques (using lagged values of competition as instruments for current values), but these exercises did not yield satisfactory results. Most coefficients turned insignificant in this case.

with an inverted U-shaped relationship between competition and R&D. The implied optimal competition level in the full sample (derived in section 4) is $-\beta_1/2\beta_2$ = $-0.747/(2\cdot(-0.0469)) = 7.963$. This means that the point where the effect of a change of competition turns from positive to negative fulfils the requirement of being within the range of the observed data. In fact, by looking at the distribution of the PE, we see that 76% of the observations in the estimation sample are on the positive slope, and 24% of the observations are on the negative slope. Examples of industries where increases in competition are bad for aggregate innovation (i.e. those with observations predominantly on the negative slope) are the manufacture of paper and paper products, the manufacture of basis metals, manufacture of machinery and equipment, retail trade and repair, and legal and economic activities. Figure 1 illustrate the hump-shaped pattern, by plotting the fitted R&D ratio against the PE.

The two right columns of table 2 present the results. The results for the lagged PE measure are nearly identical to the contemporaneous model. The implied optimal competition level slightly decreases to $-\beta_1/2\beta_2 = -0.747/(2\cdot(-0.0524)) = 7.128$. The results with the PCM remain insignificant. Thus, we conclude that, for both the PCM and the PE, the results do not seem to change when using lags. We view this as an indication of the robustness of our results, in particular to the potential endogeneity of the competition variable.

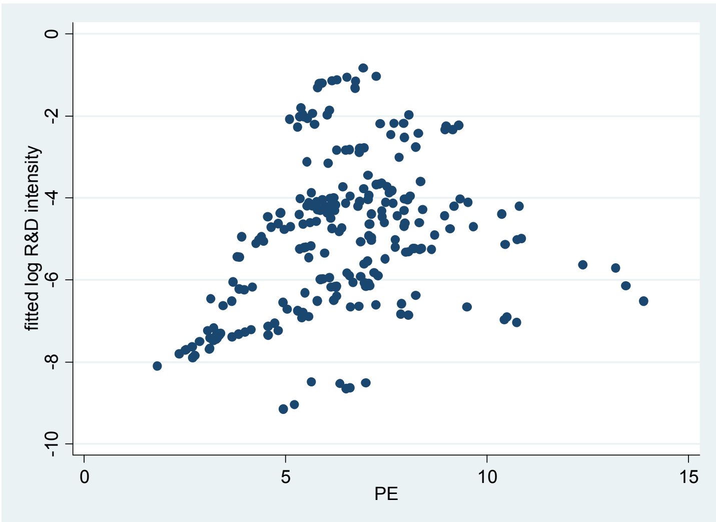

figure 1. Fitted R&D intensity based on the specification with Profit Elasticity.

6.2 Firm-level regressions

In this section the estimation results for equation (2) are given. We use various specifications: using 1–PCM or the PE as competition measure, and with capital and labour variables in the same year as R&D as well as one year lagged. In addition, it is useful to estimate the model separately for manufacturing and non-manufacturing to identify possible differences by sector, since the manufacturing sector is on average much more R&D intensive than other sectors.[18] Recall that for an inverted U, it is required that $\beta_1 > 0$ and $\beta_2 < 0$. In addition, the tipping point lies at $-\beta_1/\beta_2$, the effect of a change in competition being positive if the measure of spread is lower than this value, and negative if it is bigger.

Table 3a presents the results using (one minus) the industry PCM based on firm-level data as competition indicator (i.e. (PCM-2) in section 5.2). This is the competition indicator used in Aghion et al. (2005). Columns (1) and (2) show the results when using the whole sample. The difference between the two is that lagged capital and labour is used in the latter. We use the latter specification to check whether the possible bias due to the endogeneity of capital and labour (to the extent that this not captured by the firm and year effects) affects the estimates for β_1 and β_2.[19] For both specifications we find $\beta_1 > 0$ and $\beta_2 < 0$, in line with the predictions from the Aghion et al. model. The differences between the estimates for the coefficients of interest are moderate, which we consider an indication that the result is not biased by endogeneity of the employment or capital intensity variable. The implied estimate of the threshold for the spread where the effect of a change in competition turns from positive to negative is 0.26 in the first specification (around 80% of the observations below the threshold) and 0.31 in the second (around 92% of the observations below the threshold). In columns (3) and (4) we also see that there is evidence for the non-linear relation between competition and R&D. The implied threshold is higher in the non-manufacturing sector, however, with 0.21 for manufacturing (86% of the observations below the threshold) versus 0.67 in non-manufacturing (all observations below the threshold). This means that the spread in non-manufacturing should be much higher before a change in competition has a negative effect on innovation; in fact, although the value for the threshold is feasible (between 0 and 1), it is not observed in the sample so that the observed effect is always positive. Figure 2 gives the distribution of the technology spread in both subsamples.

---

[18] We also estimated the industry regression making this sample split, but the number of observations appeared too low to produce satisfactory results.

[19] As in the macro regressions we also estimated specifications with lagged competition (and distance-to-frontier) variables but in this case we could not identify any significant effects. This could be due to the fact that the timing assumption is not appropriate. When using instrumental variable estimation with lagged competition and distance-to-frontier as instruments, the results are again mostly insignificant. Thus, we cannot rule out the possibility that the firm-level results are subject to endogeneity bias. On the other hand, as argued above, taking account of the fixed-effect through within-estimation mitigates the effect of endogeneity if the correlation of the endogenous variable(s) with the fixed effect is relatively large compared to the correlation with the idiosyncratic error. Moreover, it can be argued that simultaneity of competition and firm-level R&D, as well as possible feedback effects of R&D on competition, are less of an issue when competition is measured at the industry level, especially in larger industries.

Table 3a. Firm-level regression results with industry PCM (within-regression).

	whole sample		manufacturing	non-manufacturing
	(1)	(2)	(3)	(4)
1−*PCM* (industry)	1.17***	1.269***	1.125***	2.901**
	(0.33)	(0.38)	(0.33)	(1.30)
SPREAD×(1−*PCM*)	-4.557***	-4.108***	-4.872***	-4.335***
	(0.36)	(0.45)	(0.39)	(0.85)
DTF	3.496***	3.302***	3.555***	3.552***
	(0.20)	(0.25)	(0.21)	(0.45)
log(*EMP*)	-0.0687***	0.00083	-0.0674**	-0.0893**
	(0.021)	(0.030)	(0.026)	(0.037)
log(*K*/*L*)	-0.343***	-0.14**	-0.349***	-0.459***
	(0.046)	(0.066)	(0.056)	(0.088)
No. of obs.	14490	6272	9182	5308
No. of firms	6442	2372	3342	3130
R^2	0.06	0.06	0.07	0.06
implied threshold	0.26	0.31	0.23	0.67

***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.10$.

Dependent variable is log R&D over value added. The regressions also contain year dummies not reported.

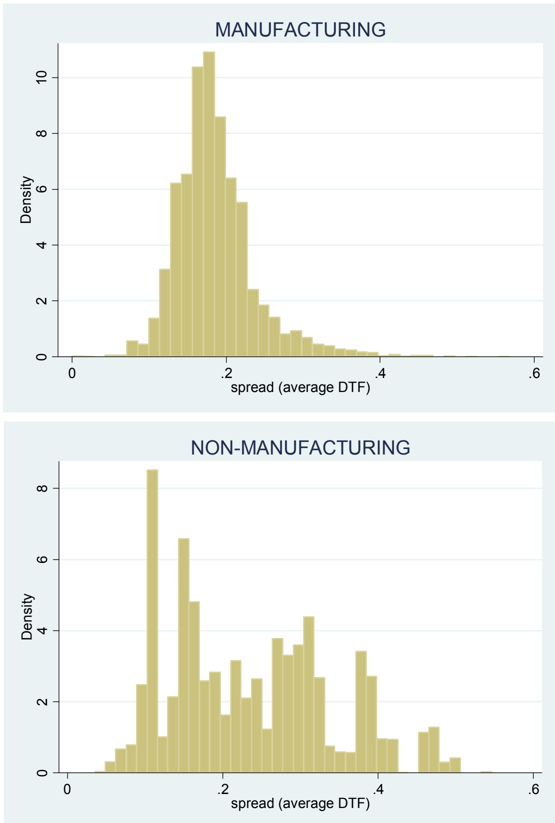

figure 2. Histograms of technology spread (average distance-to-frontier) per industry.

In each specification we find a positive effect for the distance-to-frontier. This means that a larger gap with the technological frontier is associated with a higher R&D intensity. At first glance, this may seem counterintuitive because a technological leader can be expected to invest more R&D. On the other hand, firms require a basic level of R&D to be able to absorb new technologies and catch up to the frontier (see e.g. Griffith et al. 2004). If lagging firms are relatively small, their R&D budget will form a relatively large part of their value added, which could explain the positive sign we find.

The conclusions when using the PE as an indicator of competition are slightly stronger, as can be seen in table 3b. As with the PCM, again we find that $\beta_1 > 0$ and $\beta_2 < 0$, both are significant, and the results are not affected when lagged capital and employment are used. The implied thresholds are slightly lower than in the case of the industry PCM, however, being somewhat over 0.2 for each specification. This comes down to around 70% (45%) of the observations in the estimation sample for manufacturing (services) being below the threshold, where the effect of a change in competition is positive. Unlike with the PCM, the results with the PE do not point at a different threshold for manufacturing compared to non-manufacturing. Moreover, there are firms in the non-manufacturing sector on both sides of the threshold, where with the PCM the effect was always positive. The effect of the distance-to-frontier is again positive and significant.

Table 3b. Firm-level regression results with Profit Elasticity (within-regression).

	whole sample		manufacturing	non-manufacturing
	(1)	(2)	(3)	(4)
PE	0.11***	0.112***	0.119***	0.14***
	(0.011)	(0.013)	(0.011)	(0.029)
SPREAD×*PE*	-0.55***	-0.512***	-0.586***	-0.677***
	(0.050)	(0.061)	(0.053)	(0.13)
DTF	3.32***	3.259***	3.572***	3.451***
	(0.19)	(0.24)	(0.22)	(0.42)
log(*EMP*)	-0.0539**	-0.00715	-0.0624**	-0.0758**
	(0.022)	(0.031)	(0.027)	(0.038)
log(*K*/*L*)	-0.344***	-0.136**	-0.33***	-0.458***
	(0.047)	(0.068)	(0.058)	(0.088)
No. of obs.	13724	5844	8464	5260
No. of firms	6227	2261	3147	3108
R^2	0.06	0.07	0.07	0.06
implied threshold	0.20	0.22	0.20	0.21

***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.10$.

Dependent variable is log R&D over value added. The regressions also contain year dummies not reported.

Since we have the availability over firm-level production data, it is possible to use a third indicator of competition, namely the firm-level (one minus) PCM (i.e. equation (PCM-1) in section 4.2). Arguably, this indicator is not subject to the criticism that the market of a firm is equated to its industry. However, as noted in footnote 19, the variable is more likely to be endogenous than its counterparts at the industry level. Moreover, it can be expected to contain more noise.

The results are presented in table 3c. Although we find $\beta_1 > 0$ and $\beta_2 < 0$, the most important result is that the implied thresholds for the spread of the industry where the effect of a change in competition changes from positive and negative, are well outside the admissible range for our measure of spread (between 0 and 1). Thus, in contrast to the previous results with industry competition measures, we find that the effect of competition is in effect always positive when using the firm-level version of 1−PCM as competition indicator.[20]

Table 3c. Firm-level regression results with firm PCM (within-regression).

	whole sample		manufacturing	non-manufacturing
	(1)	(2)	(3)	(4)
1−*PCM*	2.94***	3.583***	3.093***	2.662***
	(0.26)	(0.34)	(0.34)	(0.47)
SPREAD×(1−*PCM*)	-1.175**	0.0695	-1.26*	-1.539
	(0.58)	(0.69)	(0.68)	(1.29)
DTF	0.803**	-0.0826	0.700	1.242
	(0.40)	(0.48)	(0.49)	(0.83)
log(*EMP*)	-0.0147	0.0224	0.00341	-0.0538
	(0.022)	(0.030)	(0.028)	(0.039)
log(*K*/*L*)	-0.327***	-0.129**	-0.325***	-0.44***
	(0.046)	(0.065)	(0.056)	(0.088)
No. of obs.	14490	6272	9182	5308
No. of firms	6442	2372	3342	3130
R^2	0.07	0.08	0.07	0.06
implied threshold	2.50	(-51.55)	2.45	(1.73)

***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.10$.

Dependent variable is log R&D over value added. The regressions also contain year dummies not reported.

[20] Note that in the second column, the effect of competition is positive when it is *above* (rather than below) the threshold since we find that $\beta_2 > 0$. Because the implied threshold is negative, the DTF is always above the threshold. Moreover, it can be argued that for columns (2) and (4), the implied threshold is in fact infinity because β_2 does not differ significantly from 0.

6.3 Discussion of the results

Overall, our results point into the direction of a non-linear marginal effect of competition on innovation. At the industry-level we find evidence for an inverted U-shape when using the PE as competition indicator. At the micro-level we find that as the spread of the technology within an industry becomes larger the marginal effect of competition turns from positive to negative (when using either the industry PCM and the PE). In both the micro and macro results, observations are to the left and to the right of the inflection point, meaning that in some cases a change in competition is positive for innovation while in other cases it is negative. Another result that emerges from both the micro and macro results is that the mass of the observations tends to be on the positive slope. In other words, in most cases the effect of competition is positive. This can possibly explain the fact that empirical work in this area using only a linear competition term in the regression tends to find a positive relationship (e.g. Nickell, 1996). Our results point out, however, that there is also a significant portion where the effect of competition is negative. Nonetheless, when using the firm-level PCM as a measure of competition, we find that the effect is always positive. It can be argued, however, that endogeneity is likely to play a bigger role for this variable compared to its counterpart at the industry level. Also, the firm-level PCM may be a noisier indicator of competition because it is essentially a measure of a firm's profitability, which is influenced by many other (idiosyncratic) factors besides competition. Taking the industry average may be more appropriate in such a case, since it cancels out firm-specific factors going in opposite directions. Thus, we are more inclined towards the results from using the PE and the aggregated PCM.

Although the micro and macro results are generally in line, we do not find a significant relation when using the PCM in the industry regressions, while the (industry) PCM based on firm data does produce significant results in the micro regressions. A reason for this difference could be that the level of aggregation in the industry regressions is quite high (42 industries), whereas for the measure in the firm-level regressions we use a 3-digit classification. As mentioned in section 2, the higher level of aggregation, the more heterogeneous an industry becomes, making it more likely that firms are taken together that do not actually compete with each other. Apparently, the PE is less sensitive to aggregation issue, since we find similar results in the firm- and industry-level regressions.

7. Conclusion

In this paper we investigate the relationship between innovation and competition for the Netherlands. The model by Aghion et al. (2005) predicts that at the industry-level this relationship can be characterized as an inverted U-shaped curve. In our macro analysis, we test for such a relationship using industry-level data from the Dutch National Accounts. Two competition measures are used: one minus the Price Cost Margin (PCM) and the Profit Elasticity (PE). We find evidence for an inverted

U-shaped relationship at the industry-level using the latter measure, but not with the PCM measure possibly due to the high level of aggregation.

The model further predicts that the effect of a change in competition on a firm's innovative activities depends on the spread of technology within its industry. We test this prediction by estimating a firm-level regression relating innovation to the interaction of measures for competition and technological spread. We find evidence for a non-linear marginal effect of competition for both competition measures, being positive when the spread is low and turning to negative when it exceeds a certain threshold. This threshold can be determined from the estimation results. The calculation yields sensible results with industries spread out to the left and to the right of the threshold. Thus, for firms in some industries, an increase in competition will encourage innovation, while for other industries it will discourage innovation. For both the firm- and industry-level results the mass of the observations tends to be on the positive slope, so that the effect of competition is mostly positive. This can possibly explain the evidence for a positive linear relationship in existing empirical work. However, when using a third measure of competition – the firm-level PCM – we do not find evidence for a non-linear relation, although we argued that this measure is possibly more influenced by other factors than competition and the industry measures are preferred.

An important policy message from these results is that when thinking about stimulating innovation through competition, one should consider the heterogeneity of industries. There is no single answer to the question whether competition will stimulate innovation or not. From our results it appears that competition induces innovation when the industry is levelled. In such an industry, an increase in competition motivates firms to escape competition. Competition policy should therefore be designed to 'push the frontier'. However, stimulating innovation by increasing competition in an unlevelled industry is likely to yield counterproductive results. Rather than pushing the frontier, it is better to focus at enabling lagging firms to catch up to the frontier.

References

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005), 'Competition and innovation: an inverted-U relationship', *Quarterly Journal of Economics*, May 2005, pp. 701-728.

Aghion, P. and R. Griffith (2008), *Reconciling Theory and Evidence*, MIT Press, Cambridge (MA), USA.

Ahn, S. (2002), 'Competition, Innovation and Productivity Growth: A Review of Theory and Evidence', *OECD Economics Department Working Papers*, No. 317, OECD Publishing, doi:10.1787/182144868160.

Askenazy, P., C. Cahn, and D. Irac (2008), 'Competition, R&D, and the cost of innovation', Working Paper No. 2008-32, Paris School of Economics.

Boone, J., J. van Ours, and H. van der Wiel (2007), ‘How (not) to measure competition’, working paper, CENTER.

Boone, J. (2008), ‘Competition: theoretical parameterizations and empirical measures’, *Journal of Institutional and Theoretical Economics*, 164, pp. 587-611.

Bos, J.W.B., J.W. Kolari, and R.C.R. van Laroen, ‘Competition and innovation: evidence from financial services’, Tjalling C. Koopmans Research Institute, Discussion Paper 09-16.

CBS (2009), *De Nederlandse Groeirekeningen 2008*, CBS, Voorburg/Heerlen, http://www.cbs.nl/NR/rdonlyres/F2630959-6267-4A67-AF0D-D8A8EDEE3007/0/2008p40pub.pdf (in Dutch)

Dixit, A. and J. Stiglitz (1977), ‘Monopolistic competition and optimum product diversity’, *The American Economic Review*, vol. 67, no. 3 (Jun. 1977), pp. 297-308.

Friesenbichler, K.S. (2007), ‘Innovation and market concentration in Europe’s mobile phone industries: evidence from the transition from 2G to 3G’, WIFO Working Papers, no. 306, November 2007.

Griffith, R., S.J. Redding and J. van Reenen (2004), ‘Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries’, *The Review of Economics and Statistics*, vol. 86, issue 4, pp. 883-895.

Jorgenson, D. and Z. Griliches (1967), ‘The explanation of productivity change’, *The Review of Economic Statistics*, Vol. 34, No. 3, (Jul., 1967), pp. 249-283.

Lind, J.T. and H. Mehlum (2010), ‘With or without U? The appropriate test for a U-shaped relationship’, *Oxford Bulletin of Economics and Statistics*, 72(1), pp. 109-118.

Mundlak, Y. (1963), ‘Estimation of Production and Behavioral Functions from a Combination of Cross-Section and Time Series Data’, in: *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*, ed. by C. Christ et al. Stanford, California: Stanford University Press.

Nickell, S.J. (1996), ‘Competition and corporate performance’, *Journal of Political Economy*, vol. 104(4), pp.724-746.

Olley, S.G. and A. Pakes (1996), ‘The dynamics of productivity in the telecommunications equipment industry’, *Econometrica*, vol. 64(6), pp. 1263-1297.

Poldahl, A. and P.G. Tingvall (2006), ‘Is there really an inverted U-shaped relation between competition and R&D?’, *Economics of Innovation and New Technology*, 15(2), pp. 101-118.

Polder, M., E. Veldhuizen, D. van den Bergen, and E. van der Pijll (2009), ‘Micro and macro indicators of competition: comparison and relation with productivity change’, CBS Discussion Paper 09024, http://www.cbs.nl/NR/rdonlyres/B399FD48-78D2-42AD-B30C-C4CB6E14049B/0/200924x10pub.pdf

Rauch, F. (2008), ‘An explanation of the inverted-U relationship between competition and innovation’, Working Paper No. 0813, Department of Economics, University of Vienna.

Rooijen-Horsten, M. van, D. van den Bergen and M. Tanriseven (2008), ‘Intangible capital in the Netherlands: A benchmark’, Discussion Paper 08001, Statistics Netherlands, Voorburg/Heerlen.

Rooijen-Horsten, M. van, M. Tanriseven and M. de Haan (2007), ‘R&D sattelite accounts in the Netherlands: A progress report’, paper prepared for the OECD Working Party on National Accounts, Paris 3-5 October 2007.

Salop, S. (1977), ‘The noisy monopolist: imperfect information, price dispersion, and price discrimination’, *Review of Economic Studies*, vol. 44, pp. 393-406.

Scherer, F. (1967), ‘Market structure and the employment of scientists and engineers,” *American Economic Review*, vol. 57, pp. 524-531.

Tanriseven, M., M. van Rooijen-Horsten, D. van den Bergen and M. de Haan (2007), ‘Research and Development Statistics’ Paper prepared for the Intellectual Capital Congress at the Inholland University of Professional Education, the Netherlands, Haarlem, 3-4 may, 2007.

Tingvall, P.G. and P. Karpaty (2008), ‘Service-sector competition, innovation and R&D’, SSE/EFI Working Paper No. 702, October 2008.

United Nations (2008), ‘System of National Accounts’, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, World Bank, Brussels/Luxembourg, New York, Paris, Washington D.C.

Vives, X. (2004), ‘Innovation and competitive pressure’, CEPR Discussion Paper 4369.

Table 1a. Summary statistics industry-level data (estimation sample).

		R&D/VA		PE		1–PCM	
NACE	industry	mean	sd	mean	sd	mean	sd
01-05	Agriculture, forestry and fishing	0.007	0.002	.	.	0.699	0.021
15-16	Man. of food products, beverages and tobacco	0.025	0.007	7.151	0.596	0.878	0.010
17-19	Man. of textile and leather products	0.015	0.006	5.497	0.780	0.890	0.008
21	Man. of paper and paper products	0.009	0.003	8.889	0.863	0.900	0.008
22	Publishing and printing	0.003	0.002	6.771	0.432	0.819	0.005
23	Man. of petroleum products	0.014	0.017	.	.	0.912	0.027
24	Man. of basic chemicals and chemical products	0.130	0.009	5.503	0.277	0.850	0.008
25	Man. of rubber and plastic products	0.019	0.003	7.634	0.939	0.902	0.017
27	Man. of basic metals	0.022	0.022	8.892	1.333	0.871	0.039
28	Man. of fabricated metal products	0.011	0.002	6.837	0.735	0.906	0.013
29	Man. of machinery and equipment n.e.c.	0.104	0.018	8.467	0.740	0.905	0.018
30-33	Man. of electrical and optical equipment	0.321	0.050	6.609	0.452	0.983	0.036
34-35	Man. of transport equipment	0.057	0.015	6.902	0.990	0.906	0.007
20;26;36;37	Other manufacturing	0.006	0.001	6.451	0.694	0.881	0.006
40-41	Electricity, gas and water supply	0.007	0.004	.	.	0.803	0.021
45	Construction	0.002	0.001	6.497	0.605	0.876	0.010
50	Trade and repair of motor vehicles/cycles	0.000	0.000	.	.	0.790	0.008
51	Wholesale trade (excl. motor vehicles/cycles)	0.005	0.001	7.819	0.419	0.730	0.010
52	Retail trade and repair (excl. motor vehicles/cycles)	0.002	0.001	9.021	1.469	0.765	0.029
55	Hotels and restaurants	0.000	0.000	5.015	0.881	0.749	0.007
60	Land transport	0.000	0.000	6.040	0.771	0.792	0.015
61	Water transport	0.002	0.003	3.116	0.275	0.776	0.013
62	Air transport	0.001	0.001	.	.	0.954	0.045
63	Supporting transport activities	0.005	0.002	4.151	0.281	0.804	0.015
64	Post and telecommunications	0.003	0.003	3.707	0.343	0.710	0.050
65	Banking	0.004	0.002	.	.	0.794	0.043
66	Insurance and pension funding	0.003	0.002	.	.	0.711	0.062
67	Activities auxiliary to financial intermediation	0.004	0.002	.	.	0.729	0.002
72	Computer and related activities	0.013	0.005	5.756	0.461	0.849	0.013
74	Other business activities n.e.c.	0.004	0.001	12.423	1.343	0.868	0.007
85	Health and social work activities	0.000	0.000	.	.	0.809	0.008
90	Sewage and refuse disposal services	0.003	0.000	.	.	0.799	0.009
92	Recreational, cultural and sporting activities	0.000	0.000	.	.	0.843	0.008
804;91;93	Other service activities n.e.c.	0.001	0.001	3.508	0.705	0.789	0.006

'.' = data not available

Table 1b. Summary statistics firm-level data (estimation sample).

		R&D/VA		PE		1–PCM		1–PCM (industry)		DTF		DTF (industry)		K/L		employment	
NACE	#firms	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
15-16	116	0.041	0.103	7.805	1.731	0.908	0.072	0.907	0.057	0.179	0.082	0.173	0.051	10.913	9.498	325.26	506.90
17-19	40	0.042	0.043	6.894	1.953	0.902	0.087	0.915	0.027	0.214	0.112	0.206	0.073	5.247	4.830	123.76	112.19
21	43	0.028	0.054	8.582	1.414	0.885	0.077	0.890	0.023	0.153	0.067	0.155	0.019	11.084	8.701	193.79	157.64
22	39	0.037	0.121	6.518	0.902	0.861	0.094	0.876	0.040	0.186	0.080	0.194	0.030	6.658	5.201	196.41	274.11
23	6	x	x	.	.	x	x	x	x	x	x	x	x	x	x	x	x
24	93	0.127	0.281	6.053	1.604	0.878	0.094	0.886	0.030	0.189	0.091	0.196	0.041	15.789	32.307	259.97	406.51
25	62	0.048	0.111	7.053	1.213	0.888	0.080	0.884	0.018	0.186	0.069	0.177	0.029	7.854	6.708	148.94	165.10
27	22	0.037	0.049	8.453	1.809	0.906	0.072	0.912	0.030	0.204	0.094	0.194	0.064	8.453	8.444	429.42	1371.24
28	109	0.042	0.060	7.569	1.635	0.902	0.084	0.909	0.029	0.185	0.080	0.175	0.042	5.784	5.251	124.94	123.28
29	153	0.115	0.369	8.562	1.267	0.913	0.082	0.915	0.040	0.174	0.074	0.170	0.028	4.253	4.870	151.36	232.52
30-33	90	0.167	0.364	7.072	2.099	0.885	0.104	0.890	0.087	0.225	0.118	0.224	0.069	5.341	6.050	258.09	743.56
34-35	53	0.102	0.268	8.440	2.753	0.927	0.075	0.932	0.035	0.195	0.092	0.184	0.060	4.734	7.326	348.49	788.71
20;26;36;37	100	0.045	0.127	7.514	2.326	0.883	0.097	0.887	0.049	0.207	0.107	0.201	0.060	7.985	7.987	176.57	206.64
45	83	0.027	0.100	7.046	0.826	0.947	0.089	0.954	0.019	0.255	0.066	0.221	0.028	3.323	5.030	349.12	521.51
51	193	0.083	0.180	9.045	1.525	0.924	0.072	0.964	0.012	0.126	0.066	0.132	0.035	5.249	6.177	164.07	418.82
52	60	0.039	0.147	9.397	3.271	0.931	0.053	0.956	0.022	0.156	0.074	0.135	0.047	3.698	5.298	1986.26	7672.43
55	26	0.010	0.014	4.576	1.657	0.845	0.113	0.887	0.067	0.320	0.084	0.261	0.037	3.096	2.568	675.97	1107.52
60	34	0.023	0.094	5.877	0.928	0.910	0.074	0.926	0.017	0.199	0.062	0.184	0.023	6.238	6.831	326.27	718.92
61	4	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
62	1	x	x	.	.	x	x	x	x	x	x	x	x	x	x	x	x
63	28	0.063	0.424	4.464	1.041	0.859	0.138	0.868	0.128	0.259	0.100	0.266	0.048	11.034	13.875	381.35	530.71
64	9	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
71	9	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
72	74	0.164	0.329	5.935	0.743	0.872	0.123	0.898	0.028	0.311	0.124	0.306	0.055	5.886	8.391	248.87	878.98
74	171	0.077	0.200	5.042	1.118	0.884	0.106	0.926	0.028	0.396	0.124	0.347	0.078	3.767	8.325	440.72	1771.97

‘x’ = data not published due to confidentiality; ‘.’ = data not available.
See table 1a for the industry names corresponding to the NACE codes.