UvA-DARE (Digital Academic Repository)

The cyclicality of R&D investment revisited

van Ophem, H.; van Giersbergen, N.; van Garderen, K.J.; Bun, M.

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The cyclicality of R&D investment revisited

Hans C.M. van Ophem, Noud P.A. van Giersbergen, Kees Jan van Garderen and Maurice J.G. Bun

www.ase.uva.nl/uva-econometrics

Amsterdam School of Economics
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands
The cyclicality of R&D investment revisited

Hans van Ophem∗
Noud van Giersbergen †
Kees Jan van Garderen ‡
and
Maurice Bun §

December 2017

Abstract

In Fabrizio and Tsolmon (2014) and Barlevy (2007) it is concluded that R&D investments are procyclical. Fabrizio and Tsolmon (2014) utilize a model based on Barlevy (2007), but differs in some respects and allows for more heterogeneity. However, we doubt whether their implied trends are intended. Fabrizio and Tsolman also set missing values for R&D equal to zero leading to unrealistic jumps in investment and its first differences. We reconcile and replicate both the Fabrizio and Tsolmon and Barlevy papers by considering extensions that encompass both models. Furthermore, we treat missing values more appropriately to check robustness of the results. Procyclical is confirmed, but we find much less heterogeneity than Fabrizio and Tsolmon (2014) do. In particular obsolescence and patent effectiveness are no longer important but external financing is.

Keywords: R&D expenditures, innovation, cyclicality, linear panel data estimation

∗Amsterdam School of Economics, University of Amsterdam and Tinbergen Institute. Corresponding author. Address: Amsterdam School of Economics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. Email: j.c.m.vanophem@uva.nl.
†Amsterdam School of Economics, University of Amsterdam.
‡Amsterdam School of Economics, University of Amsterdam.
§Dutch Central Bank and Amsterdam School of Economics, University of Amsterdam.
1 Introduction

The relation between research & development investments (R&D) and the business cycle has been studied in a large number of papers. From a theoretical point of view, Schumpeter (1939) argues that investments in innovation are countercyclical since the opportunity costs in times of a recession are lower than in an upswing. Empirical evidence on the relation between R&D and economic growth, however, suggests procyclicality. For instance, Griliches (1990), Himmelberg and Petersen (1994), Geroski and Walters (1995), Fatas (2000), Rafferty and Funk (2004), Comin and Gertler (2006), Barlevy (2007), Ouyang (2011) and Fabrizio and Tsolmon (2014) all conclude that there is a positive relation between R&D and industry output growth. From a theoretical point of view, Francois and Lloyd-Ellis (2009), reconcile Schumpeter’s theory with the empirical findings by pointing out that innovation is more than only R&D and that other innovative activities might exhibit different business cycle properties. On the other hand, if we look at the empirical literature the evidence is much more heterogeneous than pointing only at procyclicality: Wälde and Woitek (2003), Aghion et al. (2005), Aghion et al. (2012) and Männsoo and Meriküll (2014) all allow for some kind of heterogeneity and find mixed results, whereas Francois and Lloyd-Ellis (2003) and Kraiczy et al. (2015) find no relation between R&D and the business cycle. At closer inspection, Fabrizio and Tsolmon (2014) also find mixed effects: although for most firms procyclical R&D investments are estimated, for some firms countercyclical R&D investments are found. Because in Fabrizio and Tsolman (2014) the marginal effects depend on the levels of patent protection, the rate of product obsolescence and the degree of external financing, cyclicity is heterogeneous.

There are several reasons why such ambiguous results are found. In general it can be stated that none of the studies above use the same model. Differences can be found in the definition of the dependent variable (R&D, log(R&D), R&D yes or no etc.), differences in the covariates (none, only lags, lags and leads, firm characteristics etc) and functional form (linear, nonlinear). Furthermore, the character of the data (macro, cross section, panel) and the estimation method (random effects, fixed effects, ignoring the panel character) can differ. In this investigation we will compare the methods used in two related papers to establish whether the conclusions of the papers hold in different settings. We will analyze the research of Fabrizio and Tsolmon (2014) deeper and compare it with Barlevy (2007). Fabrizio and Tsolmon (2014), henceforth denoted by FT, start from the model used by Barlevy (2007) and allows the effect of output growth to be different across the levels of patent protection, product obsolescence and external financing the industry that the firm belongs to experiences. Apart from using a different data set, FT use a different, but related, estimation technique and use controls that are different in one important aspect: they use first differenced variables instead of level variables in their specification. To make the analysis comparable we will estimate the model of Barlevy (2007), extend it according to the suggestions of FT using the same data and the same estimation method. The results are in section 2. In section 3, we will treat missing values on R&D expenditures differently than FT do, and reestimate the models to see whether the conclusions changes. In section 4 we discuss some related models and section 5 concludes.
2 The models of Fabrizio and Tsolmon (2014) and Barlevy (2007)

The starting point of FT in their analysis of the cyclicality of R&D is the first-differenced model of R&D used by Barlevy (2007). It reads:

\[
\Delta R&D_{kt} = \beta_0 + \beta_1 \Delta X_{it} + \beta_2 \Delta M_{kt} + \beta_3 \Delta M_{kt-1} + \alpha_t + \omega_{kt},
\]

where \( \Delta \) is the first differencing operator, \( R&D_{kt} \) is the natural log of real R&D investments of firm \( k \) in year \( t \), \( X_{it} \) is the log industry output relating to the industry \( i \) to which firm \( k \) belongs in year \( t \) and \( M_{kt} \) firm-level controls, contemporary and one period lagged. \( \alpha_t \) are time fixed effects.\(^{2}\) This model is extended with some interaction terms that will be discussed later and then estimated by pooled OLS.

Note that the addition of the constant \( \beta_0 \) indicates that FT allow for the possibility that there is a trend in R&D expenditures. The model of Barlevy (2007) actually differs in some important aspects. First of all, Barlevy estimates the relation between the growth in R&D and the growth in output using a fixed effects or within estimator. Furthermore, Barlevy uses related but different controls. His model reads:\(^{3}\)

\[
\Delta R&D_{kt} = (\beta_0 +) \beta_1 \Delta X_{it} + \beta_2 M_{kt} + \beta_3 M_{kt-1} + \alpha_k + \alpha_t + \xi_{kt},
\]

where \( \alpha_k \) denote firm specific effects. The presence of the constant \( \beta_0 \) is questionable, but this bears no significance because it is not identified in a fixed effects estimation as employed by Barlevy. This means that if firm specific effects exist and if they correlate with one or more of the regressors, estimating eq. (1) with pooled OLS, i.e. what is actually done by FT, will yield biased estimates, whereas estimation of (2) with the fixed effects estimator still gives unbiased results. Finally, note that Barlevy uses level controls and not first differenced controls in his specification.

As revealed in footnote 4 of FT, the first differenced model is only introduced because the model in levels displays autocorrelation and heteroskedasticity. In fact, it also removes possible endogeneity due to higher-order fixed effects. To show this, rewrite eq. (1) in level variables:

\[
R&D_{kt} = \theta_0 + \beta_0 t + \beta_1 X_{it} + \beta_2 M_{kt} + \beta_3 M_{kt-1} + \eta_k + \eta_t + \zeta_{kt}.
\]

First differencing this model gives the model in eq. (1) where we need to define: \( \alpha_t = \eta_t - \eta_{t-1} \) and \( \omega_{kt} = \zeta_{kt} - \zeta_{kt-1} \). Note that first differencing takes care of the potential endogeneity due to the correlation between the firm specific effect \( \eta_k \) and \( X_{it} \) and/or \( M_{kt} \), and that \( \eta_k \) for sure creates autocorrelation and quite likely creates heteroskedasticity. If we compare the models in eqs. (1) to (3) they are clearly related and model (2) seems to be a special mix of models (1) and (3). It allows for what is known as incidental trends whereas model (1) does not.

\(^{1}\)This variable and some other controls are inflation corrected by using a GDP-deflator.

\(^{2}\)We changed the notation somewhat here. FT use the notation \( \sum \tau_t \) where \( \tau_t \) are year indicators. In Table 5 they use the term year fixed effects.

\(^{3}\)Barlevy (2007) does not provide a regression equation. Eq. (2) is deduced from the supplementary information available on the website of the American Economic Review. The actual code used is provided there, although the data is not made available.
Cyclicality of R&D expenditure is represented by a nonzero effect of industry output (growth): $\beta_1 \neq 0$. Procyclicality would be represented by $\beta_1 > 0$ and countercyclicality by $\beta_1 < 0$.

From rewriting the model in level variables it is also clear that the constant added to the model by FT, their $\beta_0$ in the first differenced specification, represents a time trend in R&D. FT then proceed by adding their most important variables to the model: measures of industry-level obsolescence ($Ob_i$), patent effectiveness ($PE_i$) and external financing ($EF_i$) both as interactions with growth in industry output and as level variables, although these last variables are not explicitly stated in the equation listed on page 666 (second column) of FT. However, their introduction is clear from the regression output in Table 5 of FT and from the supplementary material provided by the authors. As such adding the level variables is sound econometrics but, as we will see in a moment, it has some unintended consequences for the starting model that uses R&D in levels. The most elaborate model of FT is:

$$
\Delta R&D_{kt} = \beta_0 + \alpha_t + \beta_1 \Delta X_{it} + \beta_2 \Delta M_{kt} + \beta_3 \Delta M_{kt-1} + \\
\beta_4 \Delta X_{it} \cdot Ob_i + \beta_5 \Delta X_{it} \cdot PE_i + \beta_6 \Delta X_{it} \cdot EF_i + \delta_4 Ob_i + \delta_5 PE_i + \delta_6 EF_i + \epsilon_{kt}.
$$

What FT do is allowing for heterogeneous effects of industry output growth on the growth in R&D expenditures where the heterogeneity depends in $Ob$, $PE$ and $EF$. FT expect $\beta_4 > 0$, i.e with higher degree of obsolescence firms will be more sensitive to changes in demand, and $\beta_5 < 0$, i.e. with higher degree of patent protection firms will be less sensitive to changes in demand. No discussion of the expected signs of $\beta_6$, $\delta_4$, $\delta_5$ and $\delta_6$ is provided, in fact they are not mentioned at all by FT except in the estimation output. The cyclicality of R&D expenditure is now also heterogeneous. The marginal effect of industry output growth on R&D growth is now:

$$\frac{\partial E(\Delta R&D_{kt})}{\partial \Delta X_{it}} = \beta_1 + \beta_4 \cdot Ob_i + \beta_5 \cdot PE_i + \beta_6 \cdot EF_i$$

and depending on the sign of the parameters and the level of the covariates it can indicate procyclicality for some firms and countercyclicality for others.

The meaning of $\delta_4$, $\delta_5$ and $\delta_6$ becomes clear if we reformulate model (4) in level variables. Just like in the case of the added constant, adding the time constant variables to a first difference specification, gives rise to a trend in the level variable. In the present case, there is not only a pure time trend $\beta_0 \cdot t$ but also three heterogenous time trends: $\delta_4 \cdot Ob_i \cdot t$, $\delta_5 \cdot PE_i \cdot t$ and $\delta_6 \cdot EF_i \cdot t$. As such this can be a correct specification, but FT do not provide theoretical or empirical arguments why adding these variables to the specification is a good idea. On top of that, the effect of these variables are only identified if no incidental trends are present. If we would actually follow the estimation method of Barlevy (2007), i.e. using a fixed effects estimation, we would allow for firm specific effects in eq. (1) which might potentially be correlated with the explanatory variables of eq. (1). If there is

---

$^4$FT present their estimation results on the growth in R&D in Table 5 in seven different specifications. The specifications differ in whether year fixed effects are used and in the three different combinations of their measures of patent effectiveness, obsolescence and external financing. The interaction term and level variable of each of these three variables are used in combination every time.
indeed a correlation, then pooled OLS is not the correct estimation method to use, but a fixed effects estimation will not identify the constant and the parameters of $Ob$, $PE$ and $EF$. Consequently the estimation results of FT are only correct under an additional assumption that Barlevy (2007) is not prepared to make.

We will now proceed with a replication of the estimation results of FT. In the next section we present estimation results of some alternative models that are less deviant from Barlevy (2007) to investigate whether the results of FT hold under less restrictive assumptions and alternative specifications.

Data

FT combine four different US data sources. The Compustat data provides annual information on individual firms over the period 1975-2002. This information is matched to industry level information from the NBER Manufacturing and Productivity Data and the Carnegie Mellon survey. The Carnegie Mellon survey is a cross section from 1994 and provides indicators of the explanatory variables patent effectiveness and obsolescence.

The relation between the growth in R&D expenditures of a firm, measured by $\Delta R&D_{kt}$, are related to the growth of the output of the industry the firm belongs to and a large number of controls: among others they include growth rates of cash flow, total assets, liabilities, debt and capital stock. Very similar specifications are used in Barlevy (2007), Rafferty and Funk (2004) and Ouyang (2011). The contribution of FT is to allow for additional heterogeneity of the effect of industrial output growth on R&D-expenditures. They argue that the cyclicality of R&D investments is influenced by some characteristics of the industry: the degree of patent effectiveness, obsolescence and external financing. Patent effectiveness is measured as the average (across industry) survey response on the question on the percentage of products and process innovations for which patenting was effective in protecting the firm’s competitive advantage. The obsolescence indicator is distilled from the results to the question in the Carnegie Mellon survey on the speed with which new product and process innovations are introduced in the focal industry. The external financing indicator is taken from the Compustat data by summing the annual external funds needed to finance the investments of each firm over the sample years 1975-2002 and dividing it by the firm’s total capital expenditures. This is then transformed to a industry level variable by using the median across the industry. The growth of the industry output is calculated from the NBER Manufacturing and Productivity Data by calculating the first difference in the natural logarithm of the real gross output of the industry.

Data

The data are taken from the Review of Economics and Statistics replication database. URL: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/27467. In particular, we used RD_analysis.csv. A file created by FT that already combines all different data sources.

See FT for more detailed information on the creation of the data and definition of the variables.
Replications

We were able to reproduce the 'B. R&D Estimation'-panel of Table 2 of FT, containing descriptive statistics, exactly. The results of the pooled OLS-estimation of eq. (4), are given in Table 1, columns (1) to (3). They correspond to columns (2), (5) and (6) of Table 5 of FT.

Table 1: OLS estimates of growth in R&D expenditures (eq. 4)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆Output</td>
<td>0.272∗∗∗</td>
<td>-0.777∗∗</td>
<td>-0.747∗</td>
<td>0.143∗∗∗</td>
<td>-0.125</td>
<td>-0.053</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.262)</td>
<td>(0.292)</td>
<td>(0.022)</td>
<td>(0.156)</td>
<td>(0.181)</td>
</tr>
<tr>
<td>∆Output × Patent Effectiveness</td>
<td>-0.054</td>
<td>-0.178</td>
<td>-0.279</td>
<td>-0.313</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.306)</td>
<td>(0.314)</td>
<td>(0.212)</td>
<td>(0.213)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patent Effectiveness</td>
<td>0.183∗∗∗</td>
<td>0.101∗∗∗</td>
<td>0.070∗∗∗</td>
<td>0.052∗</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.026)</td>
<td>(0.018)</td>
<td>(0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆Output × Obsolescence</td>
<td>0.341∗∗∗</td>
<td>0.341∗∗∗</td>
<td>0.114∗</td>
<td>0.091</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.088)</td>
<td>(0.048)</td>
<td>(0.057)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obsolescence</td>
<td>0.048∗∗∗</td>
<td>0.033∗∗</td>
<td>0.029∗∗∗</td>
<td>0.026∗∗∗</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.007)</td>
<td>(0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆Output × External Financing</td>
<td>0.039</td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.045)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External Financing</td>
<td>0.014∗∗∗</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.038</td>
<td>-0.189∗∗∗</td>
<td>-0.115∗∗</td>
<td>-0.037∗∗</td>
<td>-0.156∗∗∗</td>
<td>-0.142∗∗∗</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.040)</td>
<td>(0.042)</td>
<td>(0.014)</td>
<td>(0.028)</td>
<td>(0.030)</td>
</tr>
</tbody>
</table>

Data used

<table>
<thead>
<tr>
<th>Data used</th>
<th>FT</th>
<th>FT</th>
<th>FT</th>
<th>OGGB</th>
<th>OGGB</th>
<th>OGGB</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (observations)</td>
<td>71264</td>
<td>71264</td>
<td>71264</td>
<td>40922</td>
<td>40922</td>
<td>40922</td>
</tr>
<tr>
<td>R²</td>
<td>0.224</td>
<td>0.225</td>
<td>0.226</td>
<td>0.122</td>
<td>0.123</td>
<td>0.123</td>
</tr>
<tr>
<td>Average marginal effect output</td>
<td>0.272</td>
<td>0.168</td>
<td>0.157</td>
<td>0.143</td>
<td>0.074</td>
<td>0.089</td>
</tr>
<tr>
<td>%Avg marginal effect &lt;0</td>
<td>0.0</td>
<td>8.1</td>
<td>14.9</td>
<td>0.0</td>
<td>5.0</td>
<td>12.0</td>
</tr>
</tbody>
</table>

Robust standard errors, clustered by firm in parentheses. * p<0.05, ** p<0.01, *** p<0.001.
Additional controls as in FT, i.e. first differenced, including year fixed effects.

The results presented in these columns are not an exact reproduction of the results presented in Table 5 of FT but they are both in the size of the estimated coefficients and their significance very similar. All conclusions drawn by FT remain valid. In most cases R&D expenditures are procyclical, depending on patent effectiveness, obsolescence and the external financing within the industry. As noted before the inclusion of these variables introduces heterogeneity and for some firms this will result in a negative marginal effect, indicating countercyclicality. The average marginal effects indicate

---

7We asked two of our research master students to replicate independently and they found exactly the same results as we did.
8In their Table 5, FT have the entry Output × .... (3 times), but this should read ∆Output × ...., as becomes clear from their Stata-program available on the database of the Review of Economics and Statistics.
that for most model specifications and most firms R&D expenditures are procyclical. Patent effectiveness and external financing do not appear to have an effect through growth in output but do have a positive significant themselves. Note that, if anything, patent effectiveness has a negative indirect effect in combination with output growth. This is in line with the procyclicality of R&D investments, as argued by FT. Also confirming the hypothesis of procyclicality of R&D expenditures, obsolescence has both a direct positive impact as an indirect effect through output growth. The average marginal effect of output growth is about 16%, when allowing for interaction effects.

We now turn to estimating the model of Barlevy (2007) on the data of FT. The specification of the model is given in eq. (2) and the model is estimated using a fixed effects estimator. The results can be found in Table 2, columns (1) to (4).

| Table 2: Fixed effects estimates of growth in R&D expenditures eq. (2) |
|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|                 | (1)            | (2)            | (3)            | (4)            | (5)            | (6)            | (7)            | (8)            |
| ΔOutput         | 0.241***       | 0.230***       | -0.026         | 0.019          | 0.357***       | 0.263***       | 0.133***       | 0.101**        |
| (0.044)         | (0.055)        | (0.042)        | (0.050)        | (0.027)        | (0.034)        | (0.028)        | (0.033)        |
| Data used       | FT             | FT             | FT             | FT             | OGGB           | OGGB           | OGGB           | OGGB           |
| Year fixed effects | No            | Yes            | No             | Yes            | No             | Yes            | No             | Yes            |
| Balance sheet controls | No          | No             | Yes            | Yes            | No             | Yes            | No             | Yes            |
| N (observations)| 85604          | 85604          | 75807          | 75807          | 52870          | 52870          | 44144          | 44144          |
| $R^2$ (overall) | 0.003          | 0.000          | 0.237          | 0.236          | 0.007          | 0.004          | 0.114          | 0.111          |

Robust standard errors, clustered by firm in parentheses. * p<0.05, ** p<0.01, *** p<0.001.
Balance sheet controls in levels as in Barlevy (2007).

In the fixed effects estimation the significance of output growth is lost if the control variables are added to the specification. Barlevy (2007) found strongly significant positive results pointing at procyclicality of the growth in R&D expenditures. We find no relation with the business cycle if we control for firm characteristics. Although the data come from partly the same source (Compustat) there are differences in the data: here we use about twice the number of observations, the time period considered is slightly different and a different number of industries are distinguished. However, the definition of the variables is the same. It should be noted that the Barlevy estimation results are quite heterogeneous: his estimates of the effect of output growth on growth in real R&D, where the data used vary across industry codes, included time period and the way output is measured, ranges from 0.21 to 0.92.

As concluded before, the estimations of FT are based on stronger assumptions than those of Barlevy. Before estimating both specifications using the same estimation technique and investigating some alternative models, we will first correct an important mistake in the preparation of the data.

3 Treating missings differently

A simple diagnostic scatter graph of the dependent variable and the fitted values reveals that there is an interesting problem. The scatterplots in Figure 1 displays the dependent variable in the regressions
in Table 1 (x-axis) and the fitted values from the regression of model (1) (y-axis). In the left figure, on both the y- and x-axis there is an unusual concentration of points. Closer inspection of the files supplied by FT in the data archive of the Review of Economics and Statistics reveals that, missing observations on R&D expenditures and some other variables are replaced by 0’s. This results in a large number of zero R&D observations (38.2% of the 71264 observations) and this explains the concentration of points on the y-axis. Another consequence of putting R&D expenditures to 0 is hugely negative and positive changes in the R&D growth rates: if for a certain firm R&D is not observed in a certain year and it is observed in the next year the growth rate is calculated to be huge if R&D is set to 0. The other way around generates very negative growth rates. We corrected this by leaving the missings untouched and this resulted in 40922 observations that are plotted in the right hand scatter in Figure 1. Although the growth rates are quite large in absolute value in some cases, the problem appears to have been solved to a large extent. We will use this restricted data set in the remaining part of this paper and indicate it by OGGB in Tables 2 and 3.

![Figure 1: Growth in R&D expenditures and fitted values](image)

Columns (4) to (6) in Table 1 and columns (5) to (8) in Table 2 contain the same estimations as in the earlier columns but using the corrected data. If we only look at the estimated signs of the coefficients, by and large the conclusions with respect to the FT-specification do not change. The estimated coefficients are usually closer to zero and the significance appears to have reduced somewhat despite of the calculated standard errors being smaller. We also find smaller $R^2$s. Obsolescence and patent effectiveness remain to have a significant impact. If we concentrate on the full model as displayed in column (6), actually only obsolescence and patent effectiveness themselves have a significant im-
pact. The significance of the interaction term is lost. Note that none of the variables related to output growth are significant. For most firms there is procyclicality of the R&D expenditures and this is stronger if the corrected data are used and if we judge this by the number of firms with procyclical R&D. However the average marginal effect of output growth has reduced to 9%. Table 2, columns (4)-(8) displays larger and more significant results than Table 2, columns (1)-(4). Like in Barlevy (2007) also the estimates of the effect of output growth while controlling for balance sheet variables are now strongly significant.

4 Reconciliation

Despite claiming to base their model on Barlevy (2007), the model specified and estimated by FT differs in three respects:

- FT use additional controls relating to patent effectiveness, obsolescence and external financing. This is the key contribution of FT: they allow for heterogeneous effects of output growth.

- Both models relate the growth of R&D to the growth in output but use different estimation methods. FT apply OLS and the resulting estimates are in fact a result of a first difference estimation on R&D. By first differencing they remove the firm specific effect and according to FT (their footnote 4) also autocorrelation and heteroskedasticity. Barlevy estimates using a fixed effects estimation. Consequently he allows for second-order fixed effects or incidental trends.

- FT use first differenced balance sheet controls in the growth in R&D equation, whereas Barlevy uses level controls.

We will now proceed with estimating some alternative models to make the analyses more comparable. We start with rewriting the model estimated by FT in levels and then proceed estimating this equation with a fixed effects estimation. A crucial observation here is that FT introduce a constant and three time constant explanatory variables (Patent Effectiveness, Obsolescence and External Financing) in the model they estimate. If we rewrite the model actually estimated by FT, i.e. the model in eq. (4), in level variables we obtain:

\[
R&\overline{D}_{kt} = \beta_0t + \beta_1X_{it} + \beta_2M_{kt} + \beta_3M_{kt-1} + \beta_4X_{it} \cdot Ob_i + \beta_5X_{it} \cdot PE_i + \beta_6X_{it} \cdot EF_i + \delta_4Ob_i \cdot t + \delta_5PE_i \cdot t + \delta_6EF_i \cdot t + \eta_k + \eta_t + \varsigma_{kt}.
\]

As such this model can make perfect sense, but in their discussion of R&D investments, indeed as a level variable - the growth in R&D investment is only introduced when discussing their empirical

---

9 We will disregard the difference between a fixed effects estimation and the first difference estimation. As discussed in Hsiao (1986) and Baltagi (2001) both estimation methods yield consistent estimates under the same assumptions, but the fixed effects estimator is more efficient.

10 A constant is present in the model FT written down on page 666, the three time invariant variables are not, although they are presented in Table 5 on page 672.
model, FT do not discuss time trends in general and time trend related to patent effectiveness, obsolescence and external financing in particular. If we review the results in Table 1, we can conclude that there is in general a negative trend in R&D expenditures but that this negative trend is less negative or even positive for firms with relatively high degrees of patent effectiveness, obsolescence and/or external financing.

The model in eq. (5) is, as far as we are aware, never used in the literature on R&D investments. Furthermore, as becomes clear from footnote 4 of FT, the authors are actually interested in estimating R&D in levels but resort to using first-differenced variables to mitigate the problems of autocorrelation and heteroskedasticity.\footnote{The authors do not actually mention the potential existence of firm specific effects in footnote 4 as a reason for using first differences. However, note that the autocorrelation and heteroskedasticity experienced by FT might be due to these firm specific effects.} If we estimate eq (5) with a first difference estimation\footnote{From an efficiency point of view a fixed effect estimation is better. The estimation results using the fixed effect estimator are very similar.}, the results of FT as they presented in Table 1 will be found, apart from the standard errors that are calculated somewhat differently. Now consider a model which is in line with the discussion of FT:\footnote{We added Ob, PE and EF as well, although not discussed in the theoretical part of FT, they are included in their estimations.}

\[
R&D_{kt} = \beta_1X_{it} + \beta_2M_{kt} + \beta_3M_{kt-1} + \beta_4X_{it} \cdot Ob_{i} + \beta_5X_{it} \cdot PE_{i} + \beta_6X_{it} \cdot EF_{i} + \delta_4Ob_{i} + \delta_5PE_{i} + \delta_6EF_{i} + \eta_k + \eta_t + \varsigma_{kt},
\]

(6)

the difference being that we deleted all variables related to the trend. We estimate this model using the fixed effects estimator. Due to using this estimator the effects of the time constant explanatory variables cannot be estimated.\footnote{This deviates with estimation on first differences employed by FT. However, the fixed effects estimator is more efficient and more observations are actually used in the estimations. Note that also the first difference estimator does not make it possible to identify time constant effects. The differences in the estimation results however are marginal.} The relevant estimation results are presented in Table 3, column (1).\footnote{Stata reports an estimate of a constant even in a Fixed Effects Estimation (XTREG, FE). This constant is the average fixed effect calculated from the residuals. We do not report these estimates in Table 2.} Note that the estimations now are performed on 45061 observations.

The conclusions we can draw from the estimation results in Table 3 column (1), are quite different from the conclusions from Table 1 column (6). The effect of industry output growth remains insignificant, although the sign has changed. The effect of patent effectiveness and obsolescence on the impact of industry output growth also changed: using a fixed effect estimation reverses the sign and maintains insignificance. The signs are now contrary to the theoretical predictions of FT. This can signal two things. Either the model estimated is not the correct model and the model estimated by FT represent reality better or there is a tendency in the direction of acyclical R&D expenditures for firms in industries with higher patent effectiveness and with lower levels of obsolescence. Finally, the effect of external financing remains positive but has become strongly significant. From the average marginal effects of output growth on R&D expenditures we can conclude that the procyclicality of R&D is even stronger in this alternative model: procyclicality is found for 95.2% of the firms whereas it was 88% in Table 1.
Table 3: Alternative estimates of (growth in) R&D expenditures, OGGB data

<table>
<thead>
<tr>
<th></th>
<th>(1) R&amp;D</th>
<th>(2) ΔR&amp;D</th>
<th>(3) ΔR&amp;D</th>
<th>(4) ΔR&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>0.055</td>
<td>(0.148)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output × Patent Effectiveness</td>
<td>0.066</td>
<td>(0.234)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output × Obsolescence</td>
<td>-0.005</td>
<td>(0.053)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output × External Financing</td>
<td>0.105***</td>
<td>(0.029)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔOutput</td>
<td></td>
<td>-0.069</td>
<td>-0.098</td>
<td>-0.159</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.160)</td>
<td>(0.299)</td>
<td>(0.279)</td>
</tr>
<tr>
<td>ΔOutput × Patent Effectiveness</td>
<td>0.004</td>
<td>-0.068</td>
<td>-0.112</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.199)</td>
<td>(0.279)</td>
<td>(0.262)</td>
</tr>
<tr>
<td>ΔOutput × Obsolescence</td>
<td>0.076</td>
<td>0.074</td>
<td>0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.049)</td>
<td>(0.090)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>ΔOutput × External Financing</td>
<td>0.142***</td>
<td>0.038</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.032)</td>
<td>(0.053)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.452***</td>
<td>-0.118****</td>
<td>-0.526***</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>(0.356)</td>
<td>(0.023)</td>
<td>(0.068)</td>
<td>(0.015)</td>
</tr>
</tbody>
</table>

Estimation method | FE | OLS | FE | FE | Level | Level | Level | Δ | Level | Level | Level | Level | Level |
R²                | 0.451 | 0.124 | 0.092 | 0.088 |
Average marginal effect output | 0.135 | 0.238 | 0.107 | 0.103 |
% Avg marginal effect < 0 | 4.8 | 2.5 | 1.2 | 3.5 |

Robust standard errors, clustered by firm in parentheses.
* p < 0.05, ** p < 0.01, *** p < 0.001.
We now turn to using the same controls as Barlevy (2007). Instead of controls specified as first differences, Barlevy uses level variables. The results, using pooled OLS like FT do, are presented in Table 3, column (2). The results are far from significant, apart from the interaction term relating to external financing (6). The effect of obsolescence in combination with output growth is again corresponding to the theoretical expectations of FT, although far from significant. The average marginal effect of output growth for this model is quite large, almost 24%, and the number of firm with countercyclical R&D is very small (2.5%).

If we now estimate the same model using a fixed effects estimation, this is the estimation method used by Barlevy (2007) except that three interactions are used as additional regressors, we find Table 3, column (3). All significance is lost and incidental trends appear to be nonexistent. Note that the $R^2$ is in line with the previous results, indicating that there is significance of the control variables.

The same conclusion follows if we estimate the model of FT as in Table 1 and therefore use controls in differences, but not applying pooled OLS but the fixed effects estimator. The results are in Table 3, column (4). As explained before, the effects of the time constant explanatory variables cannot be estimated.

Due to the lack of significance, we conclude that it is not necessary to allow for second-order fixed effects as suggested by Barlevy (2007). If we concentrate only on the signs of the estimated coefficients, the conclusions are very similar to the ones following from Table 1. There is almost a one-to-one correspondence of the sign (14 out of 16 cases). However, as noted, almost all significance is lost. Only the external financing cross term is significant but this is precisely the variable that is not emphasized in FT. In fact it is hardly discussed by them apart as entries in regression tables.

Although not very well known in economics\textsuperscript{16}, it is actually possible to estimate the effect of time constant explanatory variables while allowing for fixed effects. In a complete linear setting this was first discussed by Mundlak (1978). He proved that if the fixed effects are linearly related to the time averages of the time varying explanatory variables, adding these time averages to the model and estimating with either random effects estimation or the pooled regression estimation will yield the fixed effect estimates. These are the estimated effects of the time varying explanatory variables. The between estimates can also be retrieved directly: it is the sum of the estimated coefficient of the time varying explanatory variable and the estimated coefficient of the corresponding time averaged additionally added explanatory variable. What appears to be less well known, is that Mundlak’s assumption of linearity is superfluous. The result actually holds without specifying the relation between the fixed effects and the explanatory variables. Furthermore, even the standard errors calculated in the random effects estimation are correct i.e. are the same as the standard errors calculated if the fixed effects estimator would have been used.\textsuperscript{17} Because we can actually use a random effects estimator to calculate the fixed effects estimator, we can also add time invariant explanatory variables and their ef-

\textsuperscript{16}No mention of this more general results can be found in e.g. Baltagi (2001). In e.g. sociology and biology, adding the time averages of the time varying explanatory variables is quite common, see e.g. Allison (2009); Burnett and Farkas (2009); Goetgeluk and Vansteelandt (2008); Kaufman (1993); Neuhaus and Kalbfleisch (1998); Ousey and Wilcox (2007); Phillips (2006); Schunck (2013); Scott and Holt (1982); Teachman (2011).

\textsuperscript{17}A proof of these claims is available on request.
fects are indeed identified. However, for consistency we need to make the assumption that these time invariant explanatory variables are not correlated with the fixed effects. Note that this assumption is implicitly also made by FT. The model in the Mundlak specification reads:

\[
R&D_{kt} = \beta_1 X_{it} + \beta_2 \Delta M_{kt} + \beta_3 \Delta M_{kt-1} + \beta_4 X_{it} \cdot Ob_i + \beta_5 X_{it} \cdot PE_i + \beta_6 X_{it} \cdot EF_i + \delta_2 Ob_i + \delta_3 PE_i + \delta_4 EF_i + \beta_1^* \bar{X}_i \cdot Ob_i + \beta_2^* \bar{X}_i \cdot PE_i + \beta_3^* \bar{X}_i \cdot EF_i + \eta_k + \eta_t + \zeta_{kt},
\]

The random effect estimator gives the following effects of time constant variables Obsolescence, Patent Effectiveness and External Financing on R&D: 8.993 (se = 1.498), -0.562 (se = 0.427) and -0.282 (se = 0.241). Only the effect of Patent Effectiveness is significant and has the same sign as reported in FT.

Our conclusion is that if the same explanatory variables and appropriate estimation methods are used, the heterogeneity of cyclicality of R&D investment found by FT is reduced considerably. Heterogeneity appears to be dependent especially on the external financing of the firm. We do find procyclicality for the vast majority of firms.

5 Conclusion

In this paper we replicated Fabrizio and Tsolmon (2014) and Barlevy (2007). The results of Fabrizio and Tsolmon were retrieved almost precisely whereas significance was lost in the Barlevy specification. Significance of the Barlevy estimates was restored after having corrected the data, whereas the overall conclusion of the Fabrizio and Tsolman paper remained the same after this correction. If we use suitable (fixed effects) estimation methods, no significance was found for the key factor of Fabrizio and Tsolmon (2014) obsolescence and patent effectiveness. Heterogeneity of the cyclicality of R&D expenditures remained, however, due to the external financing interaction term. Despite significant differences, overall procyclicality was confirmed.
References


