Characterizing the evolution of the EU-US R&D intensity gap using data from top R&D performers

J. Stančík and F. Biagi

(IPTS-JRC, Seville)*

January 2012

Abstract

In this paper we look at the evolution of the R&D intensity gap between the EU and its major competitors (among which a special attention is given to the US) using data from the Industrial Scoreboard covering the period 2002-2010. We focus on R&D intensity as it is normally recognized as an important determinant of the competitiveness of firms, and hence of economic regions, and we assess whether the gaps between the EU and its major competitors arise from differences in industrial composition (structural component) or differences within sectors (intrinsic component). This is important from a policy perspective since modifying the industrial structure is much harder than implementing an R&D subsidy across the board. The paper is divided in two parts.

In the first part we first present the evolution of the R&D intensity gap between the EU and its major competitors (US, Japan, BRIC, Asian Tigers) and then we look more closely at the role and evolution of the structural and intrinsic components for each pair-wise comparison, by looking at four basic macro-sectors defined in term of their R&D intensity.

In the second part of our work we concentrate on the EU-US R&D intensity gap and, by applying firm level analysis, we test whether the results obtained by the statistical decomposition of aggregate R&D intensity are confirmed. The evidence provided by this exercise is especially important because it allows us to perform a comparison where the ceteris paribus condition is more likely to be satisfied. In particular, we test whether there is evidence of across-sector variability in R&D intensity and whether, within sectors, EU and US firms perform differently in terms of R&D intensity. For this we have to control for various factors such as size, cyclical effects, common macroeconomic shocks and company's age. Age is important for at least two reasons. First, young companies might have more problems in finding access to funds necessary in order to invest in R&D. Second, young companies might have to be especially aggressive in terms of innovation if they want to enter and succeed in markets where incumbents already exist. More generally, company age is important because it takes time to build, test and eventually change a given business model and there is plenty of evidence that young firms are those exhibiting the highest dynamism. Therefore, our aim here is also to document the age profile for R&D intensity and to verify whether the R&D intensity gap between EU and US companies is related to age of the firm. Finally, we check if R&D intensity is affected by the abundance of internal funds (as captured by the profit/sales ratio), if this relationship changes with the age of the company and if the latter effect shows across-regional variation.

Our results indicate that there is evidence of strong across-sector variation and some evidence of within-sectors-across-region variation, which –however- is not always in favour of the US. This allows us to conclude that firm level analysis confirms the results from the aggregate analysis. Moreover we find that R&D intensity tends to decrease as firm size increases (as measured by the number of employees), that the age profile for R&D intensity behaves very differently in the two regions and that young companies in the EU exhibit a much higher reactivity to lagged profits-to-sales ratio, when compared to their US counterpart. We believe that this is an indication that the conditions for accessibility and cost and availability of funds for young top R&D performers differ significantly across the two regions.

***Address** Edificio Expo, c/ Inca Garcilaso 3, E-41092 Seville, Spain.
Email: Juraj.Stancik@ec.europa.eu, Federico.Biagi@ec.europa.eu.

1. Introduction

R&D activity, defined by the Frascati manual as "*creative work undertaken on a systematic basis in order to increase the stock of knowledge and the use of this stock of knowledge to devise new applications*" has long been recognized as a source of productivity, at the micro-meso and aggregate level. This is true for business R&D as it is for Government R&D.

Business R&D activity directly impacts on firms' productivity because it leads to new goods and services, it improves the quality of existing goods and services, and leads to improved or new production processes.[1] Studies that use firm level data to look at the direct relationship between productivity and business R&D typically find an output elasticity of own R&D between 1% and 25%, but centered at around 8% (see Hall et al, 2009). However, the own/private effect of R&D on labour productivity (or TFP) is not the only relevant factor. In fact, the whole endogenous growth theory is based on the assumption that there exist relevant knowledge spillovers that are generated by R&D activity. Such spillovers are those that permit long run growth to be generated by market driven accumulation (as opposed to external forces driving technological change, as in a typical Solow-type macro model). The literature on the impact of such spillovers tends to see a higher degree of variability in the estimates (and this is partly to be explained by the nature of the problem), but there is consistent evidence that such spillovers exist and they are sizable.[2]

Within a macro context, Guellec and van Pottelsberghe (2001) find a long run elasticity of TFP (computed at the country level) to business R&D of about 0.13, which is considered to be high, especially since it captures mostly the social return to R&D. They also notice that the (social) return to R&D stock has been increasing with time, confirming the hypothesis that R&D is becoming an increasingly important activity in the knowledge-based economy. Their results also show that R&D intensity has an additional positive effect on the elasticity of TFP to R&D, pointing to some form of increasing return to scale to R&D investment, while the share of government spending has a

[1] R&D not only affects directly the ability of a given organization (or country) to innovate and hence improve productivity, but it also increase its "absorptive capacity" (see Choen and Levinthal, 1989), hence creating an environment more receptive to innovation stimuli coming from outside. Disentangling the direct from the indirect effect is far from easy, but there is empirical evidence that the indirect effect is sizeable (see Griffith et al. 2003: additional evidence in favour of the absorptive capacity role of R&D is coming from studies that document the positive correlation between foreign R&D and productivity, see Coe and Helpman, 1995).

[2] As argued by Hall et al. (2010), from a firm's perspective spillovers can come from R&D activity performed 1) by other firms in the same sector; 2) by firms in other sectors; 3) by public research laboratories and universities and 4) by firms' laboratories and government policies of other countries. From the domestic country perspective, the first three sources are components of the social or aggregate return to R&D, while only the fourth is a spillover. So "whether we label something a spillover depends on whether it is being created by the unit under the investigation or by an entity external to that unit". This also means that estimates of spillover effects tend to be larger when evaluated at the firm level, relative to the sector or country level. However, this does not mean that the social return should be lower at the country level (in fact, the opposite is true).

negative impact on the elasticity of TFP to R&D stock (but this seems mostly due to the military components), pointing to some form of crowding out.

While providing an exhaustive review of the relationship between R&D and productivity (at the micro, meso or macro level of aggregation) is outside the scope of this work, we believe that there is sufficient evidence supporting the hypothesis that R&D matters for productivity and growth.

This conclusion is important in the context of the EU-US productivity gap: if R&D is important for growth and there are both a productivity and a R&D gap, closing the R&D gap is a precondition for closing the productivity gap[3] (however, some studies cast some doubts on the role of R&D in accounting for the different evolution of TFP in the EU relative to the US, see Havik et al, 2008).

In order to account for differences in size, the R&D gap is often presented in terms of R&D intensity gap (R&D over GDP or value added). When comparing countries, focusing on the measured R&D intensity gap is interesting, especially if one of the countries is at the technological frontier, because it shows how far a given country is from the level of R&D that it "should" exhibit given its size (as measured by GDP). Hence, a positive US-EU gap means that, relative to its size, the EU is not spending enough in R&D, and this, given the positive relationship between R&D and productivity, implies that the EU is not achieving as much productivity as it could.

In fact, this is exactly the perspective that the EU Commission and the EU Council have taken when setting the 3% target for (public and private) R&D spending as a percentage of GDP.[4] A part from the fact that the option of determining a target based on R&D/GDP ratio is debatable,[5] the Lisbon target clearly indicates that knowledge led growth is among the main objectives of the EU (see also Mathieu and van Pottelsberghe, 2008).

[3] For a study that uses micro data on the productivity gap see Ortega Argilés et al.(2011), where the effectiveness of R&D in the EU relative to the US is evaluated. For studies that look at the productivity gap at a macro level, using a sectoral decomposition, see O' Mahony and van Ark (2003) and Denis et al. (2004). See also Bassanini et al. (2000).
[4] In fact, the Lisbon agenda set a target of about 3% for the R&D intensity ratio for the year 2010, with two thirds of which to be funded by the business sector and the rest to be funded by governments.
[5] Using R&D intensity as a target as opposed to selecting a given nominal value for R&D expenditures has the obvious advantage that it does not depend upon inflation (as long as the same price index can be used for GDP and R&D, which is debated by some), It also has the advantage that across-countries-sectors comparability is possible. However, choosing the R&D intensity as a target for the overall EU has the clear disadvantage that it refers to an aggregate (the EU) which hides great across-country differences. If it is clear that some countries, such as Finland and Sweden, are very likely to satisfy the 3% target, others, such a Cyprus or Romania have values that are well below 1%. What this means is that, even if the overall target where to be satisfied at the EU level, a large across-country inequality is likely to remain. On this see van Pottelsberghe (2008). In fact, individual EU countries have selected their own year 2020 R&D intensity targets; see Innovation Union Competitiveness Report (2011). Finally, notice that since R&D expenditures, especially for the business sector, might not follow exactly the cycle, we have the paradox that the EU has more chances of reaching the target simply due to a decrease in GDP as opposed to an increase in R&D expenditures

From a research perspective, the first step is the analysis of the size and origin of the overall gap, where by "origin" we mean the role of the different economic sectors. In fact, there exists a consistent body of literature documenting the role of sectoral composition in accounting for the EU-US gap. By decomposing the overall gap into a structural component (which reflects the role of differences in sectoral composition for a given average sectoral intensity) and an intrinsic component (which reflects the with-sector R&D intensity gap, for a given average sectoral composition), it is possible to document which of the two components seem to be mostly responsive for the gap at a certain point in time or over a certain time horizon.[6]

The evidence on the role of the structural vs. the intrinsic component is not uncontroversial and this has very strong implications in terms of policy implications. In fact, if the intrinsic component is to be found dominant it means that –on average- the EU is showing a lower R&D intensity across the board, which also implies that there is a list of candidate policies to increases R&D spending (and hence intensity) that might be effective in the short run. These go from a tax credit to R&D activity to improving Intellectual Property rights to favouring foreign direct investment. If, on the contrary, the structural effect is dominant, the ability to reduce the gap in the short run are greatly reduced by the fact that the latter is arising from the structural composition of the economy, which is not likely to be altered by the policies typically considered to affect R&D intensity in the short run.[7]

Erken and van Es (2007), using the OECD ANBERD and STAN database over the period 1987-2003 for 15 countries (including the US) and 36 industries, find that the intrinsic component is dominant over the structural one, and they argue that this is in large part driven by differences in the R&D intensity within the service industry (so that, within the manufacturing sector, the role of the intrinsic component is reduced). These results appear in total contradiction with those obtained by Mathieu and van Pottelsberghe (2008), who – studying the evolution of the R&D intensity gap in the manufacturing sector using sectoral data from OECD ANBERD and STAN – find that a large share of across sector-country-time variation in business R&D intensity can be explained by across-sector variation, pointing to the role of industrial specialization as one of the major determinants of aggregate R&D intensity, hence casting even more doubts on the usefulness of an across-the-board target (while for public sector R&D intensity the bulk of the variation can be explained by country specific factors). It should be stressed that Mathieu and van Pottelsberghe look at industries within

[6] In general, the relevance of the structural effects tends to rise as the number of sectors used for the decomposition increases. This is to be expected given that the within sector differences in R&D intensity are likely to be reduced when we move to a more disaggregated sectoral composition. Moreover, the result of the decomposition is very sensitive to the measurement of R&D in the service industry, which, by itself, is a very complex task due to lack of homogeneity across countries.

[7] However, even horizontal policies might have a differential impact across sectors, since they affects differently marginal costs and benefits of R&D activity. We will come back to this later.

the manufacturing sector, hence excluding services. This is crucial in delivering the results since one of the drivers of the US-EU R&D intensity gap is the different R&D intensity in the service industry[8] (especially in trade service industries).

Moncada-Paternò-Castello et al. (2010), using firm level data obtained from the 2008 Industrial Scoreboard,[9] look at the structural vs. intrinsic component of the EU-US (and Japan) R&D intensity gap. From their analysis it is clear that the structural component is absolutely dominant since it accounts for 4/5 of the overall gap[10] (in fact the intrinsic effect is in reducing the EU gap relative to the US). The role of the structural component is stressed in the "Knowledge for Growth Expert Group" report.

A partially different story is contained in Uppenberg (2009), where the focus is on the three industries that show the highest R&D intensity (chemicals and pharmaceuticals, transport equipment and ICT and other non transport equipment), for which the evidence goes in favour of lower R&D intensity in the EU relative to US and Japan (at the same time showing that EU is more specialized in technology intensive manufacturing.[11])

A paper that is closely related to ours is Ortega-Argilés and Brandsma (2010), where they use the four waves of the 2006 Industrial Scoreboard to test whether there is evidence of across sector and across countries variation in R&D intensity. They find evidence of large across-sector variation and that the US tends to have higher R&D intensity across all sectors (however, this evidence is gathered only by separate regressions for the EU and the US with no test on the significance of the difference in the coefficients).

[8] This is mostly due to a different approach to sectoral classification of R&D activities across the Atlantic. In the EU most countries use a product-based classification, so that most of R&D are attributed to the manufacturing sector producing the product for which R&D activity was performed, while in the US the principal activity of the firm is a more widespread and used approach. In fact the NSF has estimated that roughly 93% of the R&D expenditures recorded under the US trade services industry should be allocated to the manufacturing industry. Duchene et al. (2010), using the NSF estimates re-evaluate the EU-US gap and get to a more balanced view: the gap is in both manufacturing and services, but the gap in services is now lower and the one in manufacturing is higher.

[9] Here are the main differences between Scoreboard and BERD data: Scoreboard data capture R&D investment by a given company, irrespective to the location where the R&D is performed, so that it captures the global corporate funding, while BERD data refer to R&D activities within a given territory. BERD includes non company sources of R&D while Scoreboard does not (so that the portion of R&D financed by public funding is excluded). Scoreboard collects data from audited financial accounts and reports, while BERD uses a stratified sample methodology. Moreover, Scoreboard provide information on sales, while BERD provides information on value added, so that R&D intensity can be defined in terms of R&D/sales for the former and R&D/Value added for the latter. Finally, BERD uses NACE classification while Scoreboard uses ICB classification.

[10] This result applies to the case in which the full set of EU companies is considered. The relevance of the structural gap is even higher when a reduced set of EU companies is used.

[11] The picture changes when instead of nominal value added the computations are made in real value added. Within the ICT and other non-transport equipment industry, US and Japan are particularly specialized in ICT equipment production, where (relative to other sector) prices tend to decline much faster, which drives up the contribution of the ICT production sector in the economy.

In the context of the R&D intensity gap, there have been some attempts to look more in depth at the sources of the structural gap, by considering firms demographics. This is related to the literature pointing to large EU vs. US differences in post-entry performance, and hence to the role of factors hindering firms' growth[12] and not just entry or exit (see Bartelsman et al 2003 and Bartelsman et al, 2004). If EU firms are not able to grow as fast as their US counterparts, and this is true even in high tech sectors, the effect of this in the context of the R&D intensity decomposition[13] would be that the structural component dominates, simply because EU economies are not able to expand in the high growth - high R&D intensive sectors.. The empirical counterpart of this literature in the context of the R&D intensity gap is the analysis of the gap across different age groups. This is exactly what Cincera and Veugelers do in their 2010 paper (Veugelers and Cincera, 2010) where, using data from the 2008 Industrial Scoreboard dataset find that: 1) a large fraction of leading innovators (34%) are young (i.e. born after 1975); 2) that young leading innovators (*yollies*) are especially R&D intensive; 3) that *yollies* tend to be especially present in high tech sectors (especially Internet, Biotechnology and Software); 4) that about 55% of the EU-US R&D intensity differential can be explained by the lower R&D intensity of EU *yollies* as compared to US *yollies*; 5) that 92% of this intensity differential can be explained by the different sectoral composition (i.e. EU *yollies* are not in high R&D intensive sectors), and in particular to the biotechnology, pharmaceutical semiconductors and Internet sectors.

The policy implications arising from these results are immediate: in order to close the gap it is necessary to have policies directed at stimulating the growth of young firms in highly R&D intensive sectors (see also Veugelers 2009 and Schneider and Veugelers 2008).

In the first part of the paper we first present the evolution of the R&D intensity gap between the EU and its major competitors (US, Japan, BRIC, Asian Tigers) and then we look more closely at the role and evolution of the structural and intrinsic component for each pair-wise comparison, by looking at four basic macro-sectors defined in term of their R&D intensity (as proposed by the OECD, see Hatzichronoglou, 1998).

In the second part of our work we concentrate on the EU-US R&D intensity gap and, by applying firm level analysis, we test whether the results obtained by the statistical decomposition of aggregate R&D intensity are confirmed. The evidence provided by this exercise is especially important because it allows us to perform a comparison where the ceteris paribus condition is more

[12] Such factors are identified in the cost of funds, administrative costs, lack of fully integrated market, lack of appropriate skills and presence of labour and product market regulation, all of which are believed to generate a comparative disadvantage for the EU relative to the US.

[13] On the role R&D gap in the ICT sector see Lindmark et al, (2010).

likely to be satisfied. In particular we test whether there is evidence of across-sector variability in R&D intensity and whether, within sectors, EU and US firms are performing any different. To do this we have to control for various factors such as size, cyclical effects, common macroeconomic shocks and company's age. Age is important for at least two reasons. First, young companies might have more problems in finding access to funds necessary in order to invest in R&D. Second, young companies might are especially aggressive in terms of innovation if they want to enter and succeed in markets where incumbents already exist. More generally, company age is important because it takes time to build, test and eventually change a given business model and there is plenty of evidence that young firms are those exhibiting the highest dynamism: if they do not grow they are likely to disappear, but if they grow they do so much faster that older firms (in part due to the basic growth mechanics). Therefore, our aim here is also to document the age profile for R&D intensity and to verify whether the R&D intensity gap between EU and non-EU companies is related to the age of the firm. Finally, we check if R&D intensity is affected by the abundance of internal funds (as captured by the profit/sales ratio), if this relationship changes with the age of the company and if the latter shows across-regional variation.

On the one hand, our work extends the work by Moncada-Paternò-Castello et al. (2010) by looking at a much larger period (2002-2010) and this is very important because a long time span is needed if one wants to look at the evolution of the gap and of its components. Besides, drawing conclusions from a single snapshot, while tempting, might be incorrect, if anything because of the simple reason that the EU and the US might not be perfectly aligned in terms of cycles (which affect both R&D investment and Sales but not necessarily in the same way).

On the other hand, our work is close to Ortega Argiléz and Brandsma (2010), but we look at the whole series of Industrial Scoreboard waves (hence gaining on the time dimension) and we are able to test the differential role of the main explanatory variables in the EU and the US. Besides, we also provide evidence with respect to the relationship between R&D intensity and age, hence following the insights from Veugelers and Cincera (2010).

Our paper proceeds as follows. In Section 2 we discuss our approach and present the data. In Section 3 we look at the aggregate picture emerging from our exercise, while in Section 4 we perform the firm level analysis. Section 5 concludes our work.

2. Data

In this paper we use data from the *EU Industrial R&D Scoreboard* databases[14] (henceforth the *Scoreboard*) in which R&D investment data, and economic and financial data from the last four financial years are presented for the 1,000 largest EU and 1,000 largest non-EU R&D investors.[15] This database covers about 80% of all company R&D investments worldwide. All data within the *Scoreboard* are presented in EUR applying the last available exchange rate for the whole covered period (i.e., all data in year *T* Scoreboard use the exchange rate measured in year *T-1*, even if the observations refer to *T-2*, *T-3*, *T-4*). As a result, data from different *Scoreboards* are not directly comparable and they differ even for the same company in the same reference year.[16] Therefore, we first transform all data into nominal values applying the exchange rate from the reference year and then we apply to each observation in the reference year the correct exchange rate. Then, we create a panel by joining all available *Scoreboards*. Furthermore, we drop those reference-year observations, for which a base year is available and which are different from this base year. If the base year is missing and there are more than one reference-year observations available, we employ the one from the closest *Scoreboard*.[17]

The *Scoreboard* database collects information only about top R&D investors and hence omits small companies and this tends to bias the sample in favour of large firms. By applying the approach described above, when we fill in missing firm-year observations for companies below the R&D investment threshold by information from subsequent *Scoreboards*, we actually go against this bias and in fact we decrease it, e.g., we go towards the 'real' distribution.

Another concern with the *Scoreboard* is that it applies different thresholds for EU and non-EU companies. For instance, the 1000th EU company invested into R&D 4.5 mil. EUR in 2010. On the other hand, the 1000th non-EU company invested 32.7 mil. EUR in the same year. In order to make EU and non-EU samples more comparable, we choose to keep firms with at least 50 employees (so that the conditioning variable is not the same as R&D, the subject of our analysis).

---

[14] http://iri.jrc.ec.europa.eu/research/scoreboard_2010.htm

[15] The first issue, the 2004 Scoreboard (covering the period 2000-2003), comprises only the top 500 EU and 500 non-EU R&D investors. Later Scoreboard issues from 2005 to 2007 report info on the top 700+700 R&D investors. Only starting in 2008, it comprises the top 1000+1000 R&D investors.

[16] The T *Scoreboard* covers the period of four years, from T-4 till T-1. Year T-1 is called the base year, while T-2, T-3, and T-4 are reference years. The T-1 exchange rate has been applied to transform data from other currencies to EUR for all three remaining years.

[17] To illustrate this, let's consider a company from the 2007 *Scoreboard*, included with a base year 2006 and three additional reference years 2003-2005, thus covering the last four financial years. In the 2008 *Scoreboard* this company still belongs to the top R&D investors; hence it is there with a base year 2007 and year 2006 is as a reference year. Similarly, it is also in the 2009 and 2010 *Scoreboards*. We have four firm-year observations (1 base and 3 references) pointing to the same year, 2006. In this case, we drop all three reference years 2006, leaving this company only with the firm-year observation from its base year. At the same time, this company is not among the top R&D investors in any of the 2004, 2005 or 2006 *Scoreboards*. Therefore, we use its reference years coming from the 2007 *Scoreboard* to fill in missing firm-year observations for the period 2003-2005. A general rule is that first we use all base-year observations and only if they are missing, we employ reference-year observations.

The resulting panel consists of 3,034 unique companies from 32 countries covering the period 2002-2010. In total it comprises 19,207 firm-year observations. This panel is thus unbalanced (each company stays in our sample for about 6 years on average). In a typical year, there are over two thousand companies in the data and this amount varies between 1,651 (in year 2002) and 2,466 (in year 2006). Time distribution of our data is provided in Table 1.

Summary statistics of all firm-level observations used in the analysis are provided in Table 2. This is the dataset used when comparing aggregate R&D intensity across countries, using the structural vs. intrinsic decomposition (i.e., in Section 3).

When we look at the EU-US R&D intensity comparison using firm level data (Section 4), we also want to control for company's age. However, the *Scoreboard* does not provide information about age and so we had to match the *Scoreboard* with *Orbis database*[18] and gather this information there[19]. We were able to match only about 75% of *Scoreboard* companies. This means that the panel dataset with age is made up by 2,051 unique companies and 14,041 firm-year observations.

As for sectors, while in Section 4 we use the original IBC classification, in Section 3 (i.e., when we look at the decomposition of the aggregate R&D intensity gap) we adopt an aggregation that is in line with the OECD methodology,[20] hence grouping the original ICB sectors[21] according to their average R&D intensity (across regions and time). We generate the following macro-sectors:

- *High R&D intensity* aggregation (in which the average R&D intensity is above 5%): Biotechnology, Electronic office equipment, Health care equipment & services, Internet, Leisure goods, Pharmaceuticals, Semiconductors, Software, Telecommunications equipment
- *Medium to High R&D intensity* aggregation (in which the R&D intensity is between 2 and 5 %): Aerospace & defense, Alternative energy, Automobiles & parts, Chemicals, Commercial vehicles & trucks, Computer hardware, Computer services, Electrical components & equipment, Electronic equipment, General industrials, Household goods, Household goods & home construction, Industrial machinery, Support services
- *Medium to Low R&D intensity* aggregation (where R&D intensity is between 1 and 2%): Fixed line telecommunications, Food producers, Media, Oil equipment, services & distribution, Other financials, Personal goods, Tobacco

[18] Orbis, Bureau van Dijk global database, has information on over 85 million companies.
[19] Data on company's' age for the 2008 Scoreboard have been kindly made available by Michele Cincera.
[20] Hatzichronoglou, T. (1997).
[21] The Industry Classification Benchmark (ICB) is a definitive system categorizing over 70,000 companies and 75,000 securities worldwide, enabling the comparison of companies across four levels of classification and national boundaries. The ICB system is supported by the ICB Database, an unrivalled data source for global sector analysis, which is maintained by FTSE International Limited.

- *Low R&D intensity* aggregation (where R&D intensity is below 1%): Banks, Beverages, Construction & materials, Electricity, Food & drug retailers, Forestry & paper, Gas, water & multiutilities, General retailers, Industrial metals, Industrial metals & mining, Industrial transportation, Life insurance, Mining, Mobile telecommunications, Nonlife insurance, Oil & gas producers, Travel & leisure

This sectoral aggregation, together with broad ICB categories, is described in Table 3.

In terms of regional aggregation, in Section 3 we focus on five world regions – EU, US, Japan, BRIC, and Asian Tigers- while in Section4 we only consider the EU-US R&D intensity gap. A complete description is provided in Table 4.

3. R&D intensity gap decomposition on macroeconomic data: EU versus US, Japan, Asian Tigers and BRIC

In order to analyze the sources of the R&D intensity gap between main world regions, we decompose the aggregate gap between region A and region B into an intrinsic and a structural component:

$$RDI_t^A - RDI_t^B = \sum_i (RDI_{it}^A - RDI_{it}^B)\, w_{it}^{AB} + \sum_i (w_{it}^A - w_{it}^B)\, RDI_{it}^{AB} \qquad (1)$$

with

$$w_{it}^{AB} = \frac{w_{it}^A + w_{it}^B}{2};\, RDI_{it}^{AB} = \frac{RDI_{it}^A + RDI_{it}^B}{2} \qquad (2)$$

where RDI_{it}^A is the R&D intensity of sector *i* in year *t* in region *A* (defined as R&D investment over sales) and w_{it}^A denotes the share of sector *i* in year *t* sales within region *A* total sales. The first term in equation (1) represents the intrinsic effect while the second term is the structural one.

In practice, region *A* is always the EU[22] (which is our reference region) while region *B* is, one at a time, the US, Japan, BRIC countries and Asian Tigers, respectively.

The structural component measures the difference in R&D intensity due to industrial composition. In other words, it basically tells us weather the EU is more or less specialized (relatively to any of

[22] We also look at a within-EU decomposition and we find that the firms from the newly added countries tend to have an R&D intensity that is smaller than the one observed for companies belonging to the EU15 countries. However, given the small relative size of such companies, the time profiles for R&D intensity for the EU15 and the EU27 are basically identical. Hence, in the paper we concentrate on the whole EU.

the other regions) in R&D intensive sectors. On the other hand, the intrinsic component captures the across-regional differences in within sector R&D intensity.

The results of these decompositions are depicted in Fig. 1. Since the EU is always the reference region, a negative value means that a given region is more R&D intensive than the EU, while a positive value would indicate a relatively higher R&D intensity for the EU.

Looking at the first of those four sub-figures, our analysis shows that the EU is on average less R&D intensive than the US (by about 2 percentage points) and that this gap has had an increasing trend over time (the blue line). The structural component (the green line), shows a very similar pattern and matches almost perfectly the total gap also in terms of its value. It follows that the intrinsic component is almost negligible, so that - within sectors - the EU is on average as intensive as the US. In fact, at the beginning of the analyzed period, the intrinsic component was more favorable the EU, but there has been a decreasing pattern up to 2007, with some catching up thereafter.

The broad picture that emerges from this simple decomposition is that there is a gap, that such gap has not decreased and that it is fundamentally a structural gap: within sectors EU companies do not exhibit a lower R&D intensity relative to their US competitor, however the EU economy seems to be mostly specialized in sectors characterized be a relative low R&D intensity (i.e. sectors that have low R&D intensive play a much bigger role in the EU economy than in the US). Overall, this results in the negative R&D intensity gap.

The comparison between the EU and Japan exhibits a similar pattern. As a consequence of a different sectoral composition between these two regions, Japan is more R&D intensive than the EU (however, the R&D intensity gap with respect to Japan appears to be lower than the one relative to the US). While the EU is more R&D intensive within sectors, this positive contribution to the gap is completely offset by the structural component, which is even bigger than the total R&D intensity gap. Likewise, although the EU is more R&D intensive within sectors compared to Asian Tigers, its different sectoral composition is such that Asian Tigers are overall more R&D intensive than the EU.

The BRIC is the only region not corresponding to above mentioned results. The EU is more R&D intensive than the BRIC and this gap is caused both by its higher R&D intensity within sectors and by its sectoral composition.

Fig. 2 describes the relationship between the R&D intensity gap and its structural/intrinsic components from a slightly different angle.[23] Again, it depicts four graphs where the EU is compared against its four major world competitors. The first two graphs (EU vs US and Japan) show almost identical patterns and the structural component follows closely the evolution of the total gap. During the period when the total gap becomes more negative (i.e. it increases in absolute value), the role of structural component decreases while, symmetrically, the one of the intrinsic component grows. The increase in the gap is due to either an increase in R&D intensity by US (or Japanese) firms or to a drop in R&D intensity by EU firms, but such a change is –on average- similar across sectors (hence perhaps reflecting the impact of some public policy directed at stimulating R&D in the US or Japan or simply the fact that economic cycles are not aligned across countries). Once such short-run phenomenon disappears, the series for the R&D intensity gap goes back to its previous values and the structural effect becomes once again the dominant force.

In the case of the BRIC region, the structural component oscillates between 0.55 and 0.7, leaving thus a lot of 'room' for the intrinsic component (which is likely to respond –once again- to short term variation in R&D intensity within sectors[24]).

Finally, the graph for the Asian Tigers region shows an overall rather negligible total R&D intensity gap. Nevertheless, values for the ratio of the structural component highly above 1 (except year 2002) suggest a strong and opposite role of the intrinsic component, which is also confirmed by our previous results.

Overall, this analysis shows that the total R&D intensity gap relative to any of the other four regions, exhibits a decreasing trend up to years 2005-2006. This means that, relative to the regions that on average have a higher R&D intensity (US, Japan, and Asian Tigers), in the period 2002-2005, the gap between the EU and these competitors increases, while –at the same time- relative to the regions that have a lower R&D intensity (BRICS), the gap is reduced. In both cases this work at the disadvantage of the EU. After years 2005-2006 there is some evidence that the position of the EU is improving relative to Japan and especially the Asian Tigers, but it is worsening relative to the US and to the BRICS (especially after 2008). Moreover, we see that the total gap and the share of the structural component over the total gap seem to be very strongly correlated (i.e. as the total gap increases the role of the structural component is reduced) when comparing the EU to Japan and the

[23] The intrinsic component is omitted here since its share would be just a mirror image of the structural component.
[24] It is also possible that this variation depends upon the sample composition (i.e. the small number of firms belonging to the BRIC region).

US[25], while the same does not happen when the EU is compared to BRIC and to the Asian Tigers (there is no clear correlation between the two series).

Now we look more closely inside both the intrinsic and the structural component.

We start with the intrinsic component. In Fig. 3 we plot the R&D intensity gap separately for each of the four macro-sectors described at the end of Sect. 2 (all countries are compared to the EU within each macro-sector). From the upper left graph it is evident that there are almost no differences in R&D intensities among our five world regions within *low* R&D intensive sectors (those with R&D intensity below 1%) as all four lines are very flat and close to the zero line (all of them are in negative values, which means that the EU has lower R&D intensity). The graph for the medium to low R&D intensive sectors (those with R&D intensity between 1 and 2%), however, already shows some diversity. In this macro-sector, firms in the BRIC region tend to be less R&D intensive than EU firms, while the opposite holds for Japanese firms. Firms in Asian Tigers and in the US do not perform very differently from EU firms (the gap is close to zero but there is some variation for Asian Tigers). An interesting result comes from the bottom-left graph about *medium to high* R&D intensive sectors (those with R&D intensity between 2 and 5%). Except when compared to Japanese firms (relative to which the difference is close to zero), EU firms tend to be more R&D intensive than those belonging to the remaining three regions. Finally, the biggest differences are shown in the bottom -right graph, which focuses on *high* R&D intensive sectors (those with R&D intensity above 5%) and documents a very contrasting result for the EU: the EU is much more R&D intensive than Japan, Asian Tigers or the BRIC (by about 3-5%), but it is lagging behind the US by about 2 percentage points.

These results suggest that although it might not be visible from the global economy perspective[26] (Fig. 1), there are some relevant across-country-within-sector differences (so that the intrinsic component cannot be entirely forgotten). This seems to be particularly evident when comparing the EU with the US and Japan: the US is more R&D intensive than the EU especially in *high* intensive sectors, while Japan (and Asian Tigers) are more R&D intensive mainly in *low* and *medium to low* intensive sectors.

In the last part of this sub-section we document the evolution of the difference in sectoral weights (across the five regions) for each macro-sector (Fig. 4). This exercise is meant to clarify the

[25] With the exception of the comparison with the US for the years after 2007, in which both the total R&D intensity gap and the share of the structural component are increasing, hence documenting that the UE is moving in the wrong direction.

[26] This happens because the across regions R&D intensity gap is weighted by the average size of the high-tech macro-sector across any two regions.

evolution of the sectoral composition (and hence bring some light on the sources of the structural component of the gap). This simple analysis tells us that the EU is much more specialized into the *low* intensive sectors than the three regions with higher overall R&D intensity (US, Japan and Asian Tigers). Additionally, the economic importance of *medium to high* R&D intensive sectors is lower in the EU than in those three previously mentioned regions. Finally, the EU is clearly at a disadvantage in the *high* R&D intensive sectors when compared to the US and Japan (but not when compared to BRIC or Asian Tigers). In other words, regions like the US, Japan or Asian Tigers are more oriented in sectors that require higher R&D intensity which, in sum, naturally results in the overall smaller R&D intensity in the EU. Similarly, the BRIC is more oriented in *low* intensive sectors than the EU and less oriented in *medium to high* and *high* intensive sectors, which makes it a less R&D intensive region than the EU.

When we put together the pictures emerging from Fig. 3 and Fig. 4 we notice that the position of the EU relative to the US is of special interest. On the one hand the EU is not able to shift from low to high R&D intensive sector at a satisfying rate (relative to the US the positive gap in the *low* and the negative gap in *high* intensive sectors are both increasing), and, at the same time, when we look within the more R&D intensive sectors, EU firm are less R&D intensive than their American counterpart. Overall, the first effects greatly dominate over the second (the structural component dominates over the intrinsic component), but the problem is that they both work at the disadvantage of the EU relative to the US.

In the next section we focus on the EU-US comparison and we verify whether these findings are confirmed when we use firm level data. This is very important because aggregate data[27] comparisons hide many confounding factors. Only with micro data we have some hope to satisfy the ceteris-paribus condition, which allows us to capture the role of sectoral differentiation while controlling for some of the other explanatory variables affecting R&D intensity.

4. R&D intensity gap using firm level data: EU versus US

In this section we focus on the EU-US comparison in R&D intensity, using firm level data obtained by the Scoreboard unbalanced panels with company's age among the relevant variables (as described in Section 2). Firm level data are necessary if one wants to move beyond the aggregate characterization and capture the impact of each explanatory variable while controlling for the others

[27] Even aggregate data obtained by aggregating firm level data.

(ceteris paribus hypothesis). For instance, if R&D intensity is responsive to business cycles and if the cycles are not perfectly aligned along calendar years, the cycle effect would impact on the comparison of R&D intensity between the EU and the US. Hence, if we want to compare R&D intensity within sectors and across regions we need to control for cycle components. Similarly, if R&D intensity changes with age and size of the firm, controlling for these two variables is essential if one wants to characterize R&D intensity across sectors or regions (or both).

The underlying hypothesis of our work is that technology shapes the way in which R&D generates (product and process) innovation and that this is very much a sector specific effect (i.e., the relationship between R&D and the technological frontier is sector specific). However, we also want to test whether –once we look within sectors- relevant aggregate differences among countries persist. In fact, we do not expect to observe the same within-sector R&D intensity across countries or regions, since actual benefits and costs of R&D activity are likely to be very much country-region specific. Let's think of a fictitious firm in the ICT sector that is considering investing in R&D to enhance its platform. The technology for this enhancement is available irrespective of the geographical location of the company, however the actual costs for implementing such a measure are largely dependent upon country (and, perhaps, local) conditions, such as the availability and cost of appropriate human resources, the availability and cost of funds, the size of the market in which the platform is meant to work and so on. From our perspective, it is interesting not only to capture across-sectoral differences, but also to verify if, within sectors, there is also evidence of regional variation. If we find evidence that –once controlling for the other covariates- R&D intensity varies a lot across sectors, but –within sectors- there is no sign of across regional variation, we can conclude that the structural component is indeed at the source of the R&D intensity gap. If, on the contrary, we find that there is large across-regional variation but not much across-sectoral variation, we can conclude that the intrinsic component is dominant. To test this we need to have both sector dummies and interaction terms between those dummies and the regional dummies (i.e., the EU/US dummy). However, we also have to control for the other (exogenous) factors that might affects R&D intensity, such as cycle effects, employment (a proxy for size), age and, eventually, some proxy for the availability of internal funds.
Notice that in this section we use a finer sectoral partition, since we want to have a more precise idea of the relationship between R&D intensity and sectors (especially in light of the evidence on the role of the ICT and biotech sectors).

In this section we present the results for two separate sets of regressions. In the first set we run three regressions. First, in Model (1), we regress the log of R&D intensity (measured at the firm level) on a regional dummy (EU vs. US), sector dummies, interactions between the sector dummies and the regional dummy, year dummies, a country (but not sector) specific variable capturing cycle effects and the log of the age of the company (not interacted with the regional dummy). The equation for Model (1) is:

$$ln(RDI_{icst}) = \beta_0 + \beta_1 D_{US} + \sum_{s=1}^{n} \beta_{2s} D_s + \sum_{s=1}^{n} \beta_{3s} D_s D_{US} + \sum_{t=2002}^{2010} \beta_{4t} D_t + \beta_5 GDPcycle_{jt} + \beta_6 \ln{(age_{it})} + \varepsilon_{icst}$$

where $\ln{(RDI)}$ is the natural logarithm of R&D intensity for a given firm, D_s are sector specific dummies, D_{US} is dummy value taking a value of 1 for US firms, D_t are year dummies, $GDPcycle$ is the variable representing GDP cycles and $\ln{(age)}$ is the natural logarithm of the age of the company. Subscripts *i*, *c*, *s* and *t* denotes company, country, sector and time, respectively. We allow the error terms to be correlated across observations for the same firm.[28]

In the second regression - Model (2) - we add the log of employment, used as a proxy for the firm size, but we do not interact it with the regional dummy. This specification is interesting because allows us to check whether age, per se, has some explanatory value once we control for firm size. Finally, in the third regression (Model (3)), we allow for log age and log employment interactions with the regional dummy, in order to test whether the relationship between these two variables and R&D intensity varies significantly across regions.

In the second set of regressions we introduce the log of the lagged value of the profit-to-sales ratio, which might turn out significantly affecting R&D intensity if firms are someway constrained in their access to funds. We expect this variable to affect firms differently depending on their age, and so we run the same regression separately for three age categories (Model (4), Model (5) and Model (6)): *Young* (less than 15), *Middle age* (between 15 and 60) and *Old* (above 60) companies.

The equation for Model (4), (5) and (6) is:

$$lnRDI_{isct} = \beta_0 + \beta_1 D_{US} + \sum_{s=1}^{n} \beta_{2s} D_s + \sum_{s=1}^{n} \beta_{3s} D_s D_{US} + \sum_{t=2002}^{2010} \beta_{4t} D_t + \beta_5 GDPcycle_{ct} + \beta_6 ln(psratio_{it-1}) + \beta_7 D_{US} ln(psratio_{it-1}) + \beta_8 ln\,(empl_{it}) + \beta_9 D_{US} ln\,(empl_{it}) + \varepsilon_{isct}$$

[28] We do not run fixed effects estimation because we are very interested in the coefficients on the sector and country dummies, which would disappear if we differenced the data.

where $ln(psratio)$ is the natural logarithm of the profits-to-sales ratio and $ln(empl)$ is the natural logarithm of firm employment (again, we allow the error terms to be correlated across observations for the same firm).

The choice of selecting the age of 15 as a threshold between *Young* and *Middle age* companies arises from the realization that Venture Capital funding usually has an investment cycle no longer than 15 years. This means that, if EU and US firms face very different conditions in the market for funds, this is likely to show up even more for the "Below 15" group.

The variable used to control for the cycle is obtained by HP filtering the series for GDP and then using the ratio of the cycle to the trend component for each country (so that the cycle effect is measured with comparable units across all countries). We also control for year dummies that, given the presence of the country specific cycle components, capture common macro shocks.

Controlling for employment is important because there is evidence that R&D investment does not growth linearly with firm size (see Cohen and Klepper, 1992; Cohen and Klepper, 1996; Ortega Argilés and Brandsma, 2010) and smaller firms tend to have larger R&D intensity.

Age is important for at least two reasons. First, young companies might have more problems in finding access to funds necessary in order to invest in R&D. Second, young companies have to be especially aggressive in terms of innovation if they want to enter and succeed in markets where incumbents already exist. The second effect would lead young firms to be more R&D intensive (relative to old companies), while the first one would lead them to have lower R&D intensity (again, when compared to older companies). More generally, there is vast evidence documenting that young EU firms do not grow sufficiently fast when compared to US firms (see Bartelsman et al. 2003 and Bartelsman et al, 2004). Finally, there is abundant evidence showing that the availability of internal funds matter for R&D intensity and we suspect that this relationship might be age and region specific (where regional variation reflects, among other things, institutional differences).

We now look at the first set of results, where, while controlling for firm’s size and cyclical effects, we are able to estimate the across-sector and across-country variation in R&D intensity, its (common across sectors) age profile and to verify whether such age profile is significantly different in the US relative to the EU.

When it comes to the structural vs. intrinsic debate, our results from Model (1) (see Table 5) show strong evidence of significant within-region-across-sector variation in R&D intensity, but less evidence of within-sector-across-region variation. In the EU, the most R&D intensive sectors (the reference sector is *Oil & Gas*) are: *Technology Hardware and Equipment*, *Pharmaceutical and Biotechnology*, *Software and Computer Services*, *Health Care Equipment and Services*, *Aerospace and Defense* and *Electronic and Electrical Equipment*. When comparing across regions and within sectors we find that the US are less intensive in the *Telecommunication* and *Aerospace and Defense*, while they are more intensive in the *Utilities* and *Financials* sectors. Notice that these results are confirmed in Models (2) and (3), the only change being the ordering concerning the *Aerospace* vs. the *Health Care* sector.[29]

With the exception of the *Financials* sector, for which the difference between the EU and the US is very large, the within-region-across-sector variation, when significant, tends to be larger than the within-sector-across-region variation. Moreover, while in two sectors EU firms perform less R&D than their US counterparts, the opposite is true in other two sectors. In our opinion this confirms the prevalence of the structural effects over the intrinsic effect: EU firms are on average not less R&D intensive when compared to US firms.

Coming to the age variable (see Table 6), the results of Model (1) indicate that log R&D intensity tends to decrease as (the log of) age increases, which is not unexpected given the type of selection process we are dealing with (only high R&D performers are selected); however this result is not robust to the introduction of the log of employment (i.e., firms size): see Model (2). When we allow for region-age and region-employment interactions -Model (3)- we have the following results: the age profile for R&D appears to be quite different in the EU relative to the US, since the coefficient on log age is positive for the EU and negative for the US,[30] while there is no evidence of a different profile with respect to firm size: in both countries log R&D intensity is negatively related to log employment and there is no significant interaction between size and the US dummy. This result is interesting because it seems to indicate that young companies in the EU perform differently from young firms in the US with respect to R&D decisions, perhaps because of the different functioning of financial markets in the two regions.

Notice that we have some evidence that R&D intensity varies positively with the cycle (hence showing that R&D grow more than sales in boom times and they contract more in recessions), even when controlling for common macro shocks. However, the cycle component variable is not statistically significant (we still keep it since there might be some collinearity with year dummies)

To get a better understanding of the determinants of R&D intensity, in the second set of regressions we test whether log R&D intensity is affected by the abundance of internal funds (as captured by

[29] For brevity sake we do not present the coefficients on the sector dummies for Model (2) and (3). However, they can be provided upon request to the authors.

[30] We have also tried a specification in which the log of firm's age is interacted with sector dummies, in order to verify whether differences between the US and the EU are mainly due to a different sectorial composition in the two regions. Our results, not reported here for sake of brevity, indicate that this is not the case: (i.e. there is no evidence of sector specific age profiles).

the log of the profit/sales ratio) and if the relationship between R&D and internal funds changes with the age of the company. Our estimates (Table 6, Models (4), (5) and (6)) show that there are relevant across-region and across-age group differences.

On the one hand, when looking at EU companies, we find that the proxy for internal funds enters with a positive and significant coefficient only for the *Young* and the *Middle age* group, and that the size of the estimated coefficient is declining as we move from the *Young* to the *Middle age* group[31], which is expected, since –on average- access to external funds might become easier when companies growth and accumulate both reputation and assets.

On the other hand, when we compare EU to US firms within each age group, we notice that *Young* firms in the US react to internal funds in a very different way: the interaction between the US dummy and the log of profit-to-sales ratio is negative and significant and the value of the overall coefficient on the log of profit-to-sales ratio for US *Young* firms (the sum of the coefficient on log of profit-to-sales ratio and the interaction term just mentioned) is not significantly different from zero, documenting that US firms do not really seem to be constrained by the availability of internal funds. When we look at the *Middle age* and *Old* groups, there is really no evidence of an across-region significant difference in the relationship between our dependent variable and lagged log profit-to-sales ratio, so that, for these groups, US firms do not behave differently from EU firms (hence *Middle age* US firms show a positive and significant relationship -at 90% confidence level- between the log of R&D intensity and the log of profit-to-sales ratio, while older firms do not).

We have also tested if and how these results are dependent upon the age categorization, by trying different specifications, changing the upper limit of the *Young* group, both adding and subtracting 5 years,[32] and we have found that they remain substantially unchanged.

ROBUSTNESS CHECKS (TO BE DONE)

We are now working on two issues:

1) Given the criteria imposed to firms in order to be selected in the *Industrial Scoreboard* sample, we intend to employ Heckman two-stage sample selection correction to test if our results are confirmed once we explicitly take into account the selectivity issue.

2) Our results reflect the behaviour of an average firm. However the distribution of firms is not symmetric and hence we want to verify if and how our results are confirmed one we use Quantile regressions.

[31] Similarly, the coefficient is higher for the *Middle age* group than for the *Old* group, for which, however, it is not statistically different from zero at a 90% confidence level.

[32] Given that we change the upper limit of the interval for Young firms, it follows that the lower limit of the interval for Middle age firms is affected as well. The group of Older firms is not affected by the rescaling.

5. Conclusions

In this paper, using data from the *Industrial Scoreboard* for the years 2002-2010 we look at the R&D intensity gap between the EU and its major competitors from two different but complementary perspectives. In the first part of the paper we study the evolution of the overall R&D intensity gap and of its structural and intrinsic components. Our analysis reveals that the EU is less R&D intensive than the US, Japan or Asian Tigers. This gap is, however, not caused by an overall lower intensity within sectors but rather by different sectoral compositions. In all three comparisons with these regions, the structural components play a strong role, leaving just a little room for the intrinsic components, showing that, on average, the EU is specialized in sectors with low or medium R&D intensity, while US, Japan and Asian Tigers are more oriented on higher R&D intensive sectors. In fact, when we look within sectors, we find that sometimes the EU is even more R&D intensive than these three regions. When compared to the BRIC region, the EU is more R&D intensive. But in this case, it is a consequence of higher R&D intensive activities within sectors as well as of a sectoral composition in favour of highly R&D intensive industries.

In the second part of the paper we move to firm-level analysis and we find a confirmation that across-sector differences dominate over within-sector differences. In particular we find that in the EU, (the reference sector is *Oil & Gas*), the highest R&D intensity is found in the *Technology Hardware and Equipment, Pharmaceutical and Biotechnology, Software and Computer Services, Health Care Equipment and Services, Aerospace and Defense* and *Electronic and Electrical Equipment* sectors. When we compare the EU to the US we find that US companies are less R&D intensive in the *Telecommunication*, *Aerospace and Defense* while they are more intensive in the *Utilities* and *Financials* sectors. However, the within-sector-across-region variation, when significant, is smaller in size than the within-region-across-sector variation. We also find that the age profile is radically different across the two regions: positive in the EU and negative in the US, even after controlling for a proxy for firms size (the log of employment, which enters with a negative and significant coefficient, relative to which EU and US firms do not seem to perform differently). Finally, when we look at the role of the log of lagged profit-to-sales ratio, we find that this variable is highly and positively correlated to the dependent variable for the young and middle aged firms in the EU, with a coefficient declining as we move from the first to the second group. However, for US firms, we find no evidence that the proxy for internal funds affects log R&D intensity for the group of *Young* and *Old* firms, while it does so for the *Middle age* group (for which the coefficient is not significantly different from the one registered for EU companies).

Overall, our results show that –for top R&D investing young firms- the relationship between R&D intensity and the availability of internal funds is very different in the US and in the EU. However, such differences disappear when we look at the remaining age groups. This seems to point out at the crucial role of early funding: in the US, young (top R&D investing) companies have access to external funding via Venture Capital or other sources of seed funding (private and public) and hence can detach the relative size of R&D investment from their ability to generate profits. For EU young (top R&D investing) companies this does not appear to be the case. The disappearance of significant across-regional differences when we move beyond the *Young age* group confirms that early Venture Capital financing and/or public policies directed at favouring early stage R&D investment are crucial.

References

Acharya R. C. and Coulombe S. (2005), Research and Development Composition and Labor Productivity Growth in 16 OECD Countries, mimeo, 2005

Bartelsman, E.J., Haltiwanger J. and Scarpetta S. (2004), Microeconomic evidence of creative destruction in industrial and developing countries, IZA Discussion paper n.1374

Bartelsman, E.J., Scarpetta S. and Schivardi F. (2003), "Comparative analysis of firms demographics and serviva: Micro-level evidence for the OECD countries", OECD Economics Department Working Papers, n. 348, OECD Publishing.

Bassanini A., Scarpetta S. and Visco I. (2000), "Knowledge, technology and economic growth: recent evidence from OECD countries", NBB Working papers n. 6, May.

Coe D. and Helpman (1995), "International R&D spillovers", European Economic Review, 39:5, pp. 859-887.

Cohen, W.M. and Klepper S. (1992), "The Anatomy of Industry Intensity Distributions", The American Economic Review, Vol. 82, n.4, pp. 773-799.

Cohen, W.M. and Klepper S. (1996), "A Reprise on the size and R&D", The Economic Journal, vol. 106, n.437, pp. 925-951

Denis C., Mc Morrow K. and Rogersw W. (2004), "An analysis of the EU and US productivity developments (a total economy and industry perspective)" European Economy Economic Papers, n. 208.

Duchene V., Lykogianni E. and Verbeek A. (2010), "R&D in service industries and the EU-US R&D investment gap", Science and Public Policy, vol. 37, n.6, July.

Griffith R., Redding S. and Van Reenen J. (2003), "R&D and absorptive capacity: theory and empirical evidence", Scandinavian Journal of Economics, 105 (1), 99-118.

Griliches, Z. (1995), "R&D and Productivity: Econometric results and Measurement Issues", in P. Stoneman (ed.) Handbook of Economics of Innovation and Technical Change, Blackwell Handbooks in Economics.

Guellec D. and Van Pottelsberghe (2001), “R&D and productivity growth: panel data analysis of 16 OECD countries”, OECD Economic Studies, n. 33, 2001/II.

Hall. B. H. , Mairesse J, and Mohen P. (2010),”Returns to R&D and productivity”, in Hall B. H. and Rosenberg N. eds. “Handbook of the Economics of Innovation”, vol. 2, Elsevier, 2010

Hatzichronoglou, T. (1997), “Revision of the High-Technology Sector and Product Classification”, OECD Science, Technology and Industry Working Papers, 1997/92, OECD Publishing

Havik K., Mc Morrow K., Roger W. and Turrini A. (2008), “The EU-US total factor productivity gap: An industry perspective”, European Economy Economic Papers 339.

Inklaar R., O’Mahony M. and Timmer M. (2005), “ICT and Europe’s Productivity Performance: Industry Level Growth Account Comparisons with the United States”, Review of Income and Wealth, vol. 51 (4), pp.505-536.

Lindmark S., Turlea G. and Ulbrich M. (2010), "Business R&D in the ICT sector: examining the European ICT R&D deficit", Science and Public Policy, 37 (6), p. 413-428

Mathieu A. and van Pottelsberge B. (2008), "A note on the drivers of R&D intensity", CEB Working paper n. 008/002

Mc Morrow and Rogers W. (2009), “R&D capital and economic growth: The empirical evidence”, EIB Papers, vol. 14 n.1, 2009.

Moncada Paterno'Castello P., Ciupagea C. Smith K., Tubke A., Tubbs M. (2010), "Does Europe perform too little corporate R&D? A comparison of EU and non-EU corporate R&D performance", Research Policy 39(2010), p. 523-536

Mowery D. (2009), Plus ca change: Industrial R&D in the “third industrial revolution” Industrial and Corporate Change, Volume 18, Number 1, pp. 1–50 doi:10.1093/icc/dtn049 Advance Access published January 12, 2009

Ortega-Argilés R. and Brandsma A. (2010), “EU-US differences in the size of R&D intensive firms: do they explain the overall R&R intensity gap?” Science and Public Policy, vol. 37, n. 6, July.

Ortega-Argilés R., Piva M. and Vivarelli M. (2011), “The Transatlantic Productivity Gap: is R&D the main culprit?”, IZA DP n.5586, 2011

O’Mahony, M., van Ark B. (2003), “EU Productivity and Competitiveness: An Industry perspective. Can Europe Resume the Catching Up Process?”, Luxembourg, Office for Official Publications of the European Communities.

O' Sullivan M. (2007), The EU's R&D Deficit and Innovation Policy, Expert Group on Knowledge for Growth, 2007

Romer P. (1986), ‘Increasing Returns and Long-Run Growth’, Journal of Political Economy, 94: 1002-1036

Schneider C. and Veugelers R. (2008), On young innovative companies: why they matter and how (not) to policy support them, OR 0807, Department of managerial economics, Strategy and Innovation, Katholieke Universiteit Luven

Uppenberg K. (2009), “Innovation and Economic Growth”, EIB paper, vol. 14 n.1, 2009

van Ark B., Inklaar R., McGuckin R. H. (2002) ‘Changing Gear: Productivity, ICT and Services Industries: Europe and the United States’ EPWP #02-02, The Conference Board

van Ark B., Inklaar R., McGuckin R. (2003) ‘ICT and productivity in Europe and the United States. Where do the differences come from?’ Paper for the SOM PhD Conference

van Pottelsberge B. (2008), "Europe's R&D: missing the wrong targets?", Bruegel Policy Brief, n. 03/2008

Veugelers R. (2009), “A lifeline for Europe’s young radical innovators”, Bruegel Policy Brief, Issue 2009/01

Veugelers R. and Cincera M. (2010), “Young leading innovators and the EU’s R&D intensity gap”, Bruegel Polciy Controbution, Issue 2010/09

Appendix

Table 1: Number of companies by year.

	N	%
2002	1 651	8.60
2003	1 854	9.65
2004	2 139	11.14
2005	2 444	12.72
2006	2 466	12.84
2007	2 450	12.76
2008	2 257	11.75
2009	2 077	10.81
2010	1 869	9.73
Total	19 207	

Table 2: Summary statistics.

	N	mean	sd	min	max
R&D (mil. EUR)	19 207	177	565	0	7 610
sales (mil. EUR)	19 207	5 573	15 936	0	329 760
R&D intensity (R&D/sales)	19 207	3	187	0	18 478
profit (mil. EUR)	19 207	499	2 257	-91 960	59 356
employment	18 067	20 558	50 131	50	1 600 000
age	14 041	49	47	0	440

Table 3: Sector division by ICB1 categories and R&D intensity.

	N	%
Oil & Gas	445	2.32
Basic Materials	1 577	8.21
Industrials	4 929	25.66
Consumer Goods	2 609	13.58
Health Care	2 888	15.04
Consumer Services	777	4.05
Telecommunications	249	1.30
Utilities	410	2.13
Financials	580	3.02
Technology	4 743	24.69
Low R&D intensive	2 531	13.18
Medium to Low R&D intensive	2 312	12.04
Medium to High R&D intensive	6 960	36.24
High R&D intensive	7 404	38.55

Table 4: Regional division.

Country	Region	N	%
Austria		302	3.01
Belgium		366	3.65
Czech Republic		33	0.33
Denmark		408	4.07
Finland		595	5.94
France		1 131	11.29
Germany		1 955	19.51
Greece		61	0.61
Hungary		24	0.24
Ireland		143	1.43
Italy		481	4.80
Latvia		4	0.04
Luxembourg		73	0.73
Malta		10	0.10
Netherlands		510	5.09
Poland		44	0.44
Portugal		52	0.52
Slovakia		4	0.04
Slovenia		22	0.22
Spain		241	2.41
Sweden		744	7.43
UK		2 815	28.10
	EU	**10 018**	**52.16**
USA		5 621	29.27
	US	**5 621**	**29.27**
Japan		2 410	12.55
	Japan	**2 410**	**12.55**
Brazil		64	17.49
China		135	36.89
India		142	38.80
Russia		25	6.83
	BRIC	**366**	**1.91**
Hong Kong		57	7.20
Singapore		44	5.56
South Korea		229	28.91
Taiwan		462	58.33
	Asian Tigers	**792**	**4.12**

Figure 1: Decomposition of the R&D intensity gap into structural and intrinsic components.

This figure presents the decomposition of the R&D intensity gap into structural and intrinsic components for four major world competitors of the EU. R&D intensity is defined as a ratio of sector R&D over Sales. The reference region is always the EU. A negative value means that a region is more R&D intensive than the EU, while a positive value would indicate a higher R&D intensity of the EU.

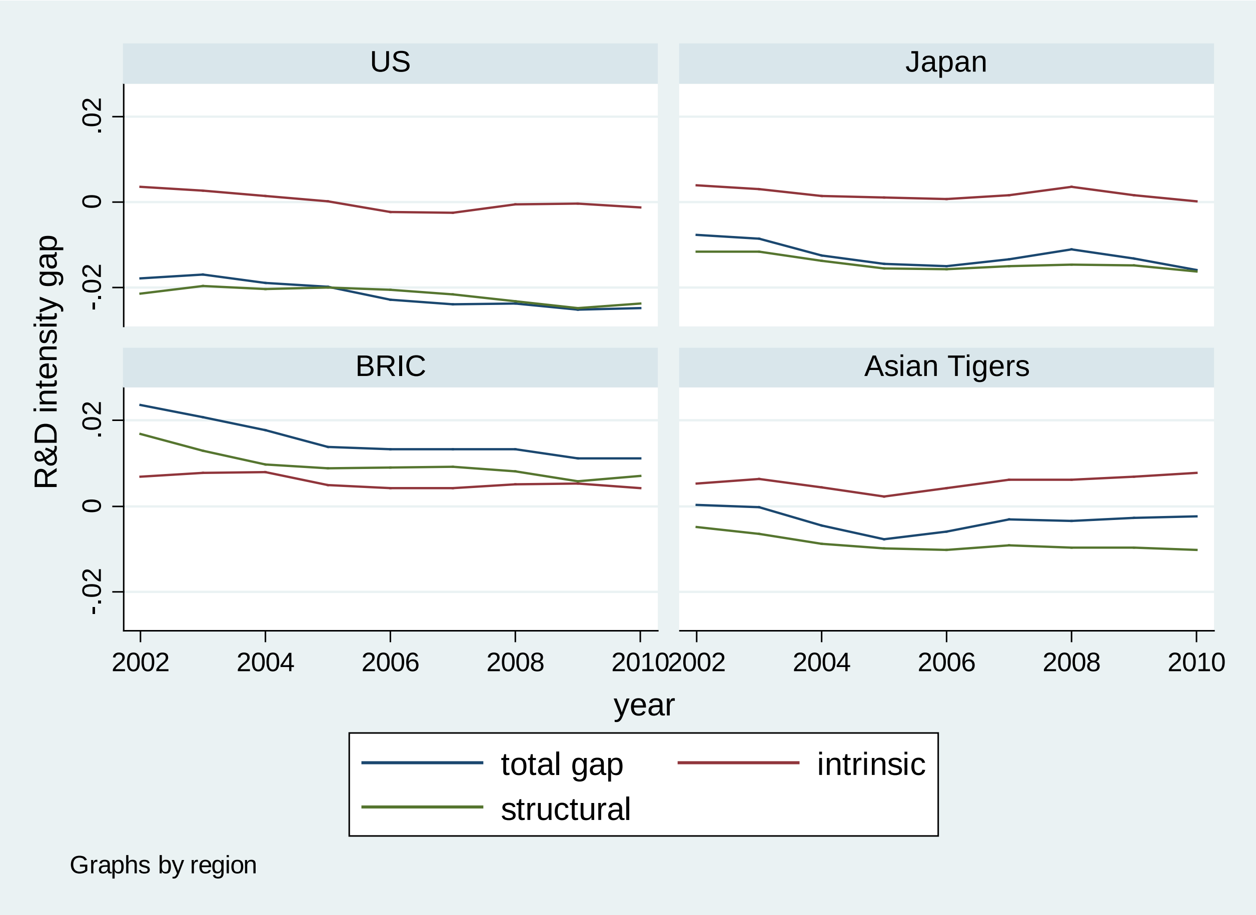

Figure 2: The share of structural component in the total R&D intensity gap.

This figure presents the evolution of the total R&D intensity gap and the share of structural component in it for the four major world competitors of the EU. The reference region is always the EU. The right hand side axis represents the ratio of the structural component in the total gap (red line).

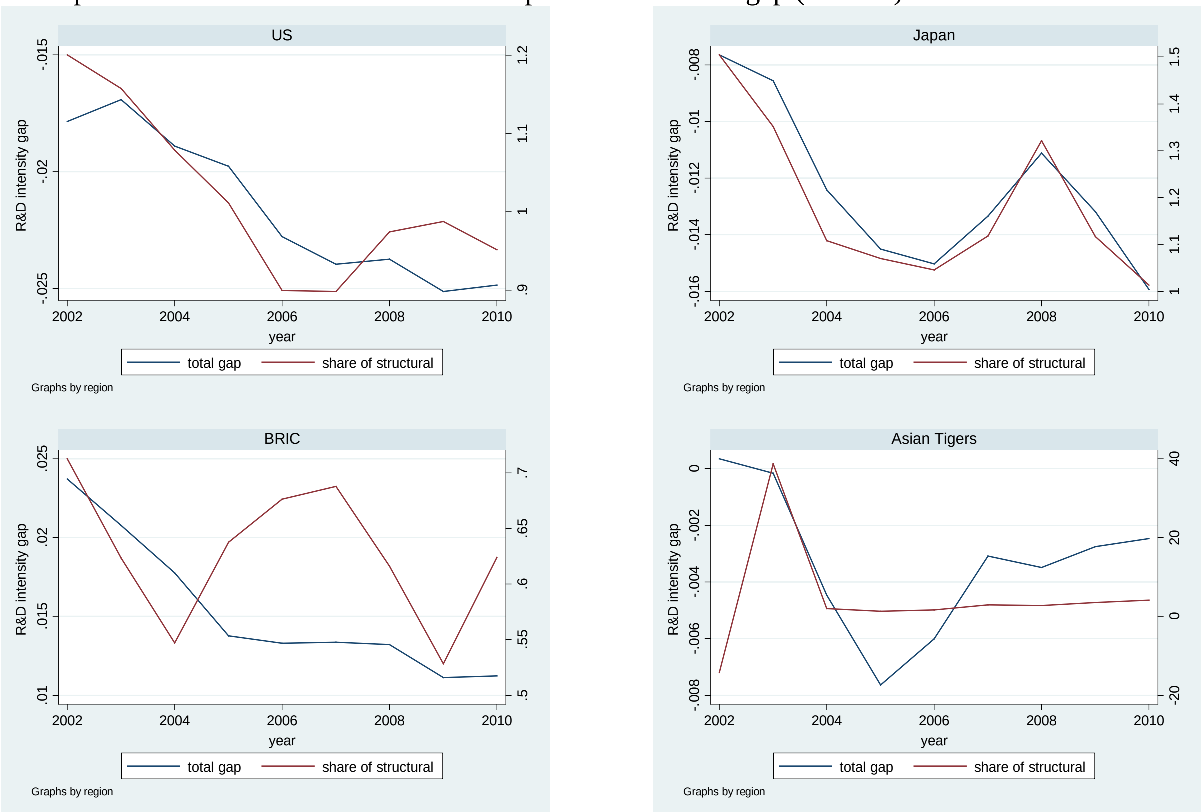

Figure 3: R&D intensity gap across regions and sectors.

This figure presents the R&D intensity gap for four major world competitors of the EU and for four sectoral categories defined in Data section. The reference region is always the EU. A negative value means that a region is more R&D intensive than the EU, while a positive value would indicate a higher R&D intensity of the EU.

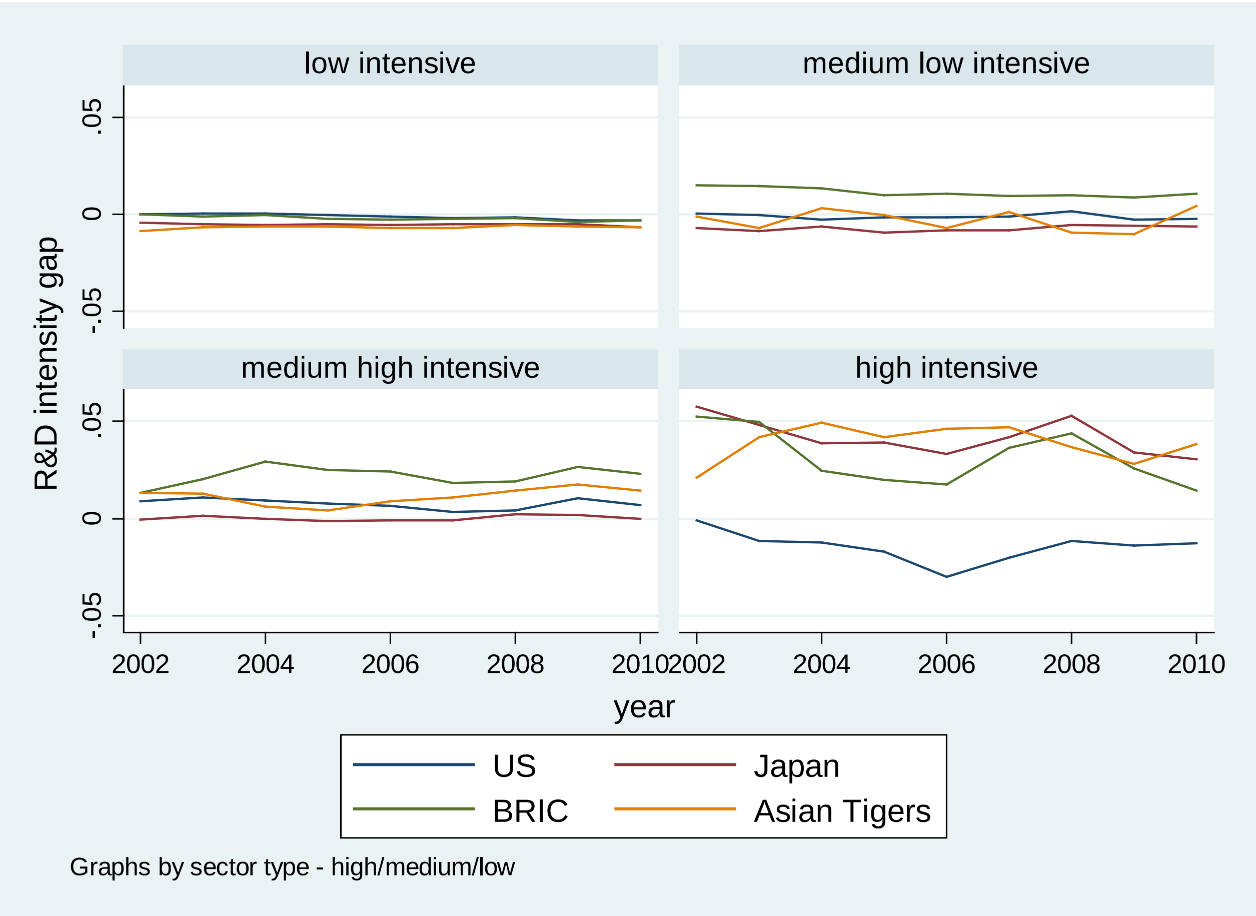

Figure 4: Sector weights across regions.

This figure presents the difference in macro-sector weights for the four sectoral categories defined in Data section and for four major world competitors of the EU. The reference region is always the EU. A negative value means that a sector in a region has bigger weight in its economy than the same sector within the EU economy.

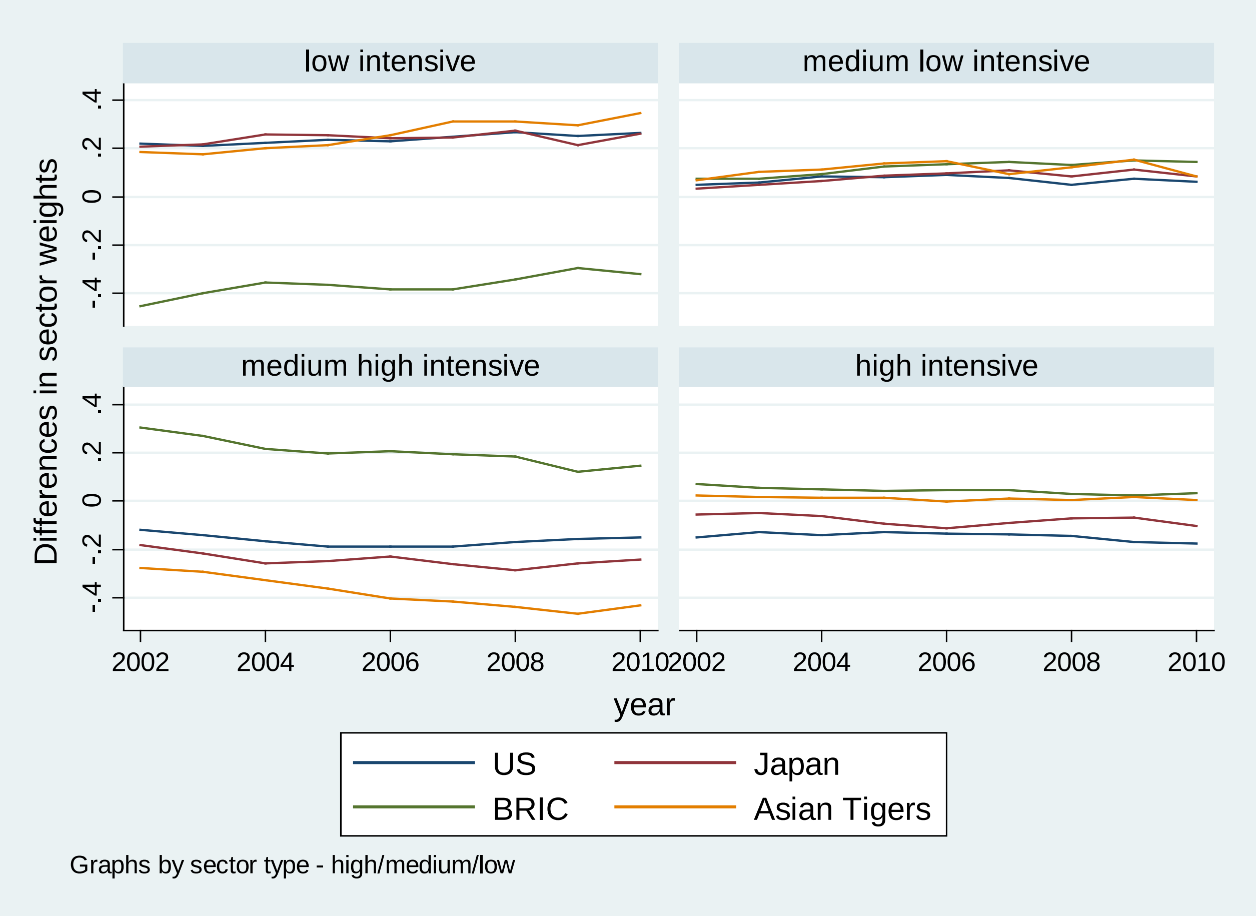

Table 5: Sector dummies and interaction with the US dummy from Model (1)

This table presents the results form firm level estimation for the period 2002-2010. Only the EU and US are considered.

Variables	Coef/Se	*Variables*	Coef/Se
Constant	-4.5574	*US dummy*	0,4268
	(0,2508)		*(0,4377)*
Basic Materials dummy	0,5789**	*US*Basic Materials dummy*	0,1133
	(0,2689)		*(0,4716)*
Industrials dummy (without 271 and 273)	0,8360***	*US*Industrials dummy (without 271 and 273)*	-0,1070
	(0,2493)		*(0,4532)*
Consumer Goods dummy	1,1089***	*US*Consumer Goods dummy*	-0,1392
	(0,2556)		*(0,4617)*
Consumer Services dummy	-0,0011	*US*Consumer Services dummy*	0,9981
	(0,3378)		*(0,6679)*
Telecommunications dummy	0,4168	*US*Telecommunications dummy*	-1,0008**
	(0,2884)		*(0,4678)*
Utilities dummy	-0,5702*	*US*Utilities dummy*	1,0846**
	(0,3238)		*(0,4906)*
Financials dummy	0,4179	*US*Financials dummy*	1,8131***
	(0,2998)		*(0,4813)*
Aerospace & Defense dummy	2,0047***	*US*Aerospace & Defense dummy*	-0,9221*
	(0,2612)		*(0,4775)*
Electronic & Electrical Equipment dummy	1,8943***	*US*Electronic & Electrical Equipment dummy*	-0,2285
	(0,2552)		*(0,4756)*
Health care equipment & services dummy	2,0966***	*US*Health care equipment & services dummy*	-0,2096
	(0,2688)		*(0,4616)*
Pharmaceuticals & Biotechnology dummy	2,7158***	*US*Pharmaceuticals & Biotechnology dummy*	-0,0682
	(0,2620)		*(0,4606)*
Software & Computer Services dummy	2,6094***	*US*Software & Computer Services dummy*	-0,0657
	(0,2536)		*(0,4512)*
Technology Hardware & Equipment dummy	2,8490***	*US*Technology Hardware & Equipment dummy*	-0,2959
	(0,2478)		*(0,4481)*
Year dummies	*yes*		
R^2	*0.515*		
N	*10,325*		

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 6: cycle, age, employment and their interactions with the US dummy for Models (1), (2) and (3).
Log of lagged profit-to-sales ratio for Models (4), (5) and (6)

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Variables	*Coef/Se*	*Coef/Se*	*Coef/Se*	*Coef/Se*	*Coef/Se*	*Coef/Se*
Constant	-4.5574***	-2.5665***	-2.5902***	-1.6721***	-1.8722***	-3.5581***
	(0.2508)	*(0.2436)*	*(0.2789)*	*(0.4874)*	*(0.4533)*	*(0.6194)*
GDP cycle	0.1941	-0.0193	0.0131	0.2345	-0.5690	0.2349
	(0.1938)	*(0.1740)*	*(0.1724)*	*(0.5801)*	*(0.3727)*	*(0.3225)*
ln(*age*)	-0.0929***	0.0340	0.0580*			
	(0.0253)	*(0.0243)*	*(0.0298)*			
*US** ln(*age*)			-0.0952*			
			(0.0494)			
ln(*employment*)		-0.2712***	-0.2781***	-0.2922***	-0.3508***	-0.1299***
		(0.0182)	*(0.0231)*	*(0.0442)*	*(0.0397)*	*(0.0391)*
*US** ln(*employment*)			0.0331	-0.0131	0.0606	0.0460
			(0.0352)	*(0.0663)*	*(0.0626)*	*(0.0596)*
L.ln(*profit/sales*)				0.2555***	0.1107**	0.0924
				(0.0771)	*(0.0508)*	*(0.0704)*
*US**L.ln(*profit/sales*)				-0.2209***	-0.0286	-0.1185
				(0.0830)	*(0.0657)*	*(0.0792)*
Sector dummies	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes
R^2	0.515	0.592	0.593	0.639	0.658	0.420
N	10,325	10,235	10,235	1,497	2,968	2,639

Note: *** p<0.01, ** p<0.05, * p<0.1