Employment Effect of Innovation§

d'Artis Kancs
European Commission
DG Joint Research Centre
41092 Seville, Spain
d'artis.kancs@ec.europa.eu

Boriss Siliverstovs
KOF Swiss Economic Institute
ETH Zurich
8092 Zurich, Switzerland
boriss.siliverstovs@kof.ethz.ch

July 28, 2015

Abstract

The present paper estimates and decomposes the employment effect of innovation by R&D intensity levels. Our micro-econometric analysis is based on a large international panel data set from the EU Industrial R&D Investment Scoreboard. Employing flexible semi-parametric methods – the generalised propensity score – allows us to recover the full functional relationship between R&D investment and firm employment, and to address important econometric issues, which is not possible in the standard estimation approach used in the previous literature. Our results suggest that modest innovators do not create and may even destruct jobs by raising their R&D expenditures. Most of the jobs in the economy are created by innovation followers: increasing innovation by 1% may increase employment up to 0.7%. The job creation effect of innovation reaches its peak when R&D intensity is around 100% of the total capital expenditure, after which the positive employment effect declines and becomes statistically insignificant. Innovation leaders do not create jobs by further increasing their R&D expenditures, which are already very high.

Keywords: Innovation, R&D investment, causal inference, semi-parametric, employment, job creation, GPS.
JEL code: C14, C21, F23, J20, J23, O30, O32, O33.

§The authors acknowledge helpful comments from Francesco Bogliacino and Daria Ciriaci as well as participants of the conference Counterfactual Methods for Policy Impact Evaluation (COMPIE) in Rome. We are grateful to Alexander Tübke for granting access to the EU Industrial R&D Investment Scoreboard data. The authors are solely responsible for the content of the paper. The views expressed are purely those of the authors and may not under any circumstances be regarded as stating an official position of the European Commission. Computations were performed in Ox 6.30 and PcGive 13.0 (Doornik, 2007; Doornik and Hendry, 2009).

1 Introduction

In setting the Europe 2020 Strategy, the European Union (EU) has defined five ambitious objectives – on employment, innovation, education, social inclusion and climate/energy – to be reached by 2020 (European Commission, 2013, 2015). In particular, concerning the first two key targets the Strategy is aimed at: (i) increasing employment by raising the employment rate of population to at least 75%; and (ii) promoting innovation by increasing research and innovation expenditures to at least 3% of GDP. For example, only from the EU Cohesion Policy 41.0 billions are allocated to research and innovation, and 71.7 billions to labour markets between 2014 and 2020. In light of the high policy priority, the objective of the present study is to assess to what extent and under which circumstances both innovation and employment can be increased simultaneously.

At the first glance, simultaneous boosting of both employment and innovation activity by increasing public investment may seem an easy and most natural task to achieve as, at least in the short-run, public investment expenditures tend to create jobs. However, the econometric results reported in the literature on employment effects of innovation are rather contradictory both with respect to its sign and magnitude, suggesting that increasing innovation intensity can have not only complementary, but also substitutionary effects on firm employment (Young, 1993; Piva and Vivarelli, 2005; Antonucci and Pianta, 2002; Van Reenen, 1997). More generally, the previous results imply that the relationship between innovation and employment may be far more complicated that one can naively assume initially.

The complexity arises due to both conceptual issues and empirical evidence. Conceptually, the challenges in understanding the relationship between the variables of interest arise, for example, due to the coexistence of many different transmission mechanisms and general equilibrium feedback loops, as the employment effect of innovation depends, among others, on the nature of innovation (product or process innovation); the purpose of innovation (to save labour or capital, neutral, or biased towards skills) and other factors (Pianta, 2004; Kancs and Ciaian, 2011). Empirically, the employment effect of innovation depends on the firm's sector of activity; formal and informal institutions; the time frame of the analysis; specifics of the existing production technology; the dimension of innovation (radical or incremental); consumer preferences; the fierceness of competition in intermediate input and labour markets; and the structure of workforce skills (Bogliacino and Vivarelli, 2012; Bogliacino *et al.*, 2012; Vivarelli, 2007; Lachenmaier and Rottmann, 2007).

The diversity in the channels of adjustment and reverse causality of interdependencies between innovation and employment suggest a non-linear functional relationship between these two variables. Hence, an accurate estimation of the functional form depends crucially on the ability to account for non-linearities in the innovation-employment nexus. In order to allow for differentiated impact of innovation on employment while accounting for differences among firms at different R&D intensity levels, an estimation approach is required which does not average across all firms, but instead allows for differentiated employment effect at different R&D intensity levels. Due to complexities in the challenges to the estimation approach, there are no studies available in the literature, that would attempt to identify the non-linearities in the R&D and firm employment relationship in such a continuous non-linear setting.

In the present study we estimate the full functional relationship between firm's innovation and employment growth by relying on flexible semi-parametric methods – the generalised propensity score (GPS) method – suggested by Hirano and Imbens (2004). The following two main features of the GPS methodology make it particularly attractive for our purpose: (i) the estimation is based on a flexible semi-parametric regression allowing for a non-linear dependence between the variables of interest without imposing any a priori restrictions; and (ii) the elimination of the selection bias arising from a non-random assignment of treatment (R&D expenditure) intensity across firms by conditioning on the observed firm characteristics. In applying the GPS methodology we are interested in identifying the R&D intensity levels under which innovation can be complementary with respect to firm employment and under which it may have an adverse impact on employment.

The assessment of the employment effect of innovation for different R&D intensity levels is our main contribution to the literature and policy debate – these insights can help to design policies, which contribute to achieving both the innovation and employment targets of the Europe 2020 Strategy simultaneously. To the best of our knowledge, the application of the GPS methodology to the employment – innovation nexus is the first of this sort in the literature.

We base our empirical micro-econometric analysis on a large international firm-level panel dataset, and our proxy for technology is a measurable and continuous variable, while most of the previous studies have relied on either indirect proxies of technological change or dummy variables (such as the occurrence of product and process innovation). In particular, we employ the EU Industrial R&D Investment Scoreboard data set, which comprises data on R&D investment, as well as other financial and economic variables for the top 1173 R&D global performers, 483 of which are active in high-tech sectors, which we analyse in detail, as high-tech companies create the most jobs both in absolute and relative terms.[1] In addition to firm-level innovation expenditures, we make use also of economic and financial variables, which allow us to control for important firm-specific effects. Moreover, the R&D Scoreboard also identifies the industrial sector (of the parent subsidiary) as well as the geographical region of R&D investment (according to the location of firm's headquarter), which allows us to control for fixed sector-specific and location-specific effects.

[1]As shown in Table 1, all top 20 global innovation leaders are active in either in high-tech and/or medium high-tech sectors.

Our results confirm previous findings that innovation can both create and destruct jobs (which, as we show, depends strongly on the innovation intensity). Second, the relationship between innovation and job creation is highly non-linear. At low innovation intensity levels (the share of R&D investment in the total capital expenditure between zero and 35-40%) an additional investment into R&D may even destruct jobs. At medium to medium-high innovation intensity levels (R&D intensity around 100%) the innovation impact on employment is positive and statistically significantly different from zero. The employment elasticity with respect to innovation is 0.7%, which implies that increasing innovation by 1% raises employment by 0.7%. The job creation effect of innovation reaches its peak when the R&D intensity is around 100% of the total capital expenditure, after which the positive employment effect declines and becomes statistically indifferent from zero. At high and very high innovation intensity levels (the share of R&D investment in the total capital expenditure above 150%) the innovation impact on employment becomes negative again, implying that, on average, additional R&D investment in innovation leaders destructs jobs. These results of decomposing the employment effect by innovation intensity are new and have not been reported in the literature before.

Our results have important messages for policy makers. First, our findings confirm the important role that innovation followers can play in creating jobs and in ensuring the sustainability of high employment in the medium- to long-run. In light of the results of Crepon *et al.* (1998),[2] two alternative policy strategies can be identified how policy makers can target this objective: policy instruments aiming at the growth of innovation followers creating jobs, and policy instruments aiming at increasing the number of innovation followers, as they both undertake innovative activities and create employment in the EU. Second, our results point to potential complementarities between the two Europe 2020 policy targets aiming to increase the R&D/GDP ratio and the employment rate, particularly by supporting innovation followers. Indeed, the empirical evidence, which we provide in this study, supports the view that R&D expenditures can be beneficial to job creation capacity. These findings imply that both the innovation and employment targets of the Europe 2020 Strategy can be reached simultaneously, by designing tailored policies for innovation followers as they create most of the employment in the economy. On the other hand, our results suggest that innovation leaders and modest innovators tend to destruct jobs through additional investment into R&D, implying that these companies should not by targeted by the policy to achieve both the innovation and employment targets of the Europe 2020 Strategy. According to Kancs and Siliverstovs (2012), innovation leaders are key for achieving the innovation the innovation target of the Europe 2020 Strategy by boosting firm productivity and competitiveness. In summary, the findings of the present

[2]The model of Crepon *et al.* (1998) distinguishes between four stages of innovation process: the decision to innovate, the decision on how much to spend on innovation activities, the relation between expenditure on innovation and innovation output, and the relation between innovation output and performance.

study and Kancs and Siliverstovs (2012) suggest that innovation leaders should be targeted, if policy objective is to boost productivity and competitiveness, whereas innovation followers should be targeted, if policy objective is to achieve both the innovation and employment targets of the Europe 2020 Strategy.

The rest of the paper proceeds as follows. The next section provides a brief review of relevant literature. The econometric methodology is outlined in Section 3. Data used in our study are described in Section 4. In Section 5 the estimation results are reported, our results are compared with those of previous studies, as well as the robustness checks' results with respect to changing the information set are reported. The final section summarises our findings and draws policy conclusions.

2 Previous literature

The question of whether technological change creates or destroys jobs has been posed since the beginning of the classical economics of Karl Marx:

> "Suppose that the making of the new machinery affords employment to a greater number of mechanics, can that be called compensation to the carpet makers, thrown on the streets?" (Marx, 1867: 479).[3]

Despite the high policy relevance of the issue, the existing evidence available in the literature does not allow for connecting all the dots in the innovation-employment relationship and often is even contradictory. Heterogeneous results, reported in the literature reflect, among others, the existence of complex adjustment and interdependency mechanisms at play. On the one hand, labour-biased technological change and labour-saving innovation can result in *technological unemployment*. For example, if there is a potential for creating a more efficient workforce by replacing workers through the acquisition of capital goods, innovation may result in technological unemployment. On the other hand, different market *compensation mechanisms*, which are triggered by technological change, can compensate for the initial labour-saving impact (Bogliacino *et al.*, 2012; Lucchese and Pianta, 2012). As noted by Bogliacino and Vivarelli (2012), innovation may reduce unit costs of production, which in a competitive market would translate into lower output prices. Lower prices, in turn, would stimulate additional demand for products, additional production and hence higher employment (*price effect*). Given that the price effect is not instantaneous, in the period between the decrease in production costs and the subsequent fall in prices, excess profits and excess income may be accumulated by the innovative firms and their employees. Whereas excess profits may be directly invested, creating in such a way new jobs, excess income may result into higher demand for goods,

[3]Das Kapital (1867), Volume I, Chapter 15, Section 6.

and hence a higher employment (*income effect*) (Freeman *et al.*, 1982; Freeman and Soete, 1987; Katsoulacos, 1986).

The compensation and displacement mechanisms (price effect and income effect) outlined above may create ir destruct jobs, depending on the nature of innovation. Indeed, empirical studies confirm that the nature of innovation is an important determinant of the overall employment effect of innovation (Pianta, 2004). Product innovation induces two countervailing effects: a direct and an indirect effect. Whereas the direct effect of product innovation leads to higher turnover and hence may increase employment, the indirect effect may reduce employment, for example, if product innovation creates monopoly power or displaces older, more labour intensive products. Similarly, process innovation triggers the same two countervailing effects: a direct and an indirect effect. Process innovation will likely have a negative direct effect on employment, as improved production processes reduce the need for labour. The indirect effect of process innovation may lead to an increase in employment, for example, if lower production costs are passed through to consumers, which, in turn, increase the demand for the product (Pianta, 2004).[4]

Empirical studies have found that the sectoral dimension of innovation is a particularly important determinant of the overall employment effect of innovation. On the one hand, the above discussed compensation mechanism in form of new products or new services may accelerate the secular shift from manufacturing to services (Vivarelli, 1995, 2013). On the other hand, new technologies in manufacturing seem to be characterised mainly by labour-saving embodied technological change, which are only partially compensated by the market mechanisms discussed above (Vivarelli, 2014). Inter-sectoral differences in the employment-innovation relationship have been confirmed also in other studies, e.g. Bogliacino *et al.* (2012).

The contradicting evidence coming from different studies suggesting that technological development can both create jobs as well as destruct them (a fact confirmed also in the present study) does not allow for understanding the underlying functional relationship between innovation and employment, which is required to be helpful for policy makers in identifying the 'right' types of firms at the 'right' stage of innovation process to ensure the desirable synergies between innovation and employment and to achieve both targets of the Europe 2020 Strategy. In order to increase innovation without reducing employment, policy makers require well-targeted policy initiatives at the 'right' stage of innovation process to a well-identified subset of firms. In the context of the Europe 2020 Strategy's objectives, particularly relevant questions are: (i) at which R&D intensity levels innovation and employment are complementary, and when innovation may have an adverse impact on firm employment: low, intermediate or high R&D intensity? (ii) what type of firms create more jobs (and hence provide the highest potential for policy synergies): innovation leaders,

[4]The impact of organisational and management innovation on firm employment is less clear-cut.

innovation followers or modest innovators?

The present study attempts to fill this research gap by identifying the R&D intensity levels under which firm innovation is likely to be complementary with respect to firm employment and under which conditions it may have an adverse impact on firm employment. Identifying and estimating the employment effect of innovation for the full range of all possible R&D intensity levels is our main contribution to the literature and policy debate; and to the best of our knowledge no comparable studies are available in the literature, which would decompose the gross employment effect by innovation intensity in a continuous setup.[5]

3 Econometric strategy

We estimate the functional relationship between innovation and employment by relying on the generalised propensity score (GPS) approach introduced in Hirano and Imbens (2004).[6] The GPS approach is a further elaboration on the popular binary treatment propensity score estimator of Rosenbaum and Rubin (1983) widely used for impact evaluation of various programs.[7] In the context of the present study the relevant features of the GPS methodology are as follows. First, it allows for continuous rather than binary treatment levels. Second, it allows to estimate the treatment effect also without a 'zero' control group. Third, the GPS procedure eliminates selection bias arising due to a non-random assignment (choice) of treatment (R&D) intensity across firms by conditioning on observed firm characteristics. Finally, it captures potential non-linearities in the functional relationship between R&D investment and firm employment, as it relies on a flexible semi-parametric regression. As a result, the estimated dose-response functions reveal the entire interval of the average and marginal treatment effects over all possible treatment levels (R&D intensity).

Following Hirano and Imbens (2004), we implement the GPS estimator in three steps. However, before describing these steps we stipulate the temporal framework of our analysis. The values of the response variable (employment) correspond to the year 2007, i.e. the last year before the collapse of Lehman Brothers in September 2008 that triggered the outbreak of the Great Recession. In order to avoid the simultaneity bias, the values of the treatment variable (R&D intensity) correspond to the year 2006. We derive the values of the generalised propensity score conditional on the observed

[5]The closest approach to ours is that of Ciriaci *et al.* (2013), who use a quantile regression methodology to decompose the gross employment effect according to quantiles of firm innovation intensity. Our study builds on the work of Bogliacino (2014), who points that R&D has a non-linear effect on employment. However, our results are more disaggregated, as they allow for a continuous impact of innovation on employment, which is not the case in Ciriaci *et al.* (2013) and Bogliacino (2014).

[6]This approach was already applied to the following pairs of variables: R&D intensity and productivity in Kancs and Siliverstovs (2012), migration and trade in Egger *et al.* (2012), and growth effects of the regional policy in the European Union in Becker *et al.* (2012), *inter alia*.

[7]For an accessible presentation of the logic underlying the propensity-score matching see Heinrich *et al.* (2010).

firms' characteristics for this year.

The first step is based on the assumption that the conditional distribution of treatment variable, r, or, as most often in the literature, its logarithmic transformation, $\ln r$, is normal given the covariates, X:

$$\ln r_{it}|X_{it} \sim N(X_i^{2006\prime}\gamma; \sigma^2), \tag{1}$$

where X_i^{2006} is a $z \times 1$ vector of both contemporaneous values of discrete and continuous covariates. The parameters of the conditional distribution (γ, σ^2) are evaluated using a standard OLS regression. The estimated GPS is defined as follows:

$$\widehat{s}_i^{2006} = \frac{1}{\sqrt{2\pi\widehat{\sigma}^2}} \exp\left[-\frac{1}{2\widehat{\sigma}^2}(\ln r_i^{2006} - X_i^{2006\prime}\widehat{\gamma})^2\right]. \tag{2}$$

The propensity score in Equation (2) fulfils its purpose of measuring the degree of similarity across heterogeneous firms when the so-called *balancing* property is satisfied, i.e. for those firms with assigned equal propensity scores (conditional on the firm-specific covariates) the associated treatment level is independent from firm characteristics. In this step, we follow the procedure specified in Hirano and Imbens (2004) in order to verify whether the balancing property is not violated in our data sample.

In the second step, the expected value of response variable, $\ln\omega_i^{2007}$, is modelled as a flexible semi-parametric function of treatment variable and the estimated generalised propensity score, $\ln r_i^{2006}$ and s_i^{2006}, respectively:

$$\begin{aligned} E[\ln\omega_{it}^{2007}|\ln r_i^{2006}, s_i^{2006}] = Incpt &+ \alpha_{11} * \ln r_i^{2006} + \alpha_{12} * \left[\ln r_i^{2006}\right]^2 + \alpha_{13} * \left[\ln r_i^{2006}\right]^3 \\ &+ \alpha_{21} * s_i^{2006} + \alpha_{22} * \left[s_i^{2006}\right]^2 + \alpha_{23} * \left[s_i^{2006}\right]^3 \\ &+ \alpha_3 * (\ln r_i^{2006} * s_i^{2006}), \end{aligned} \tag{3}$$

where the latter is substituted with its estimates, $\widehat{s}_i^{2006}$, from the first step. The flexibility of the functional form can be controlled for by varying the power of variables $\ln r_i^{2006}$ and s_i^{2006} and their cross-products.

The average expected response of target variable, ω, for a given treatment dose, ρ, is estimated in the third step:

$$\begin{aligned} E[\ln\widehat{\omega}^{2007}(\rho)] = \frac{1}{N}\sum_{i=1}^{N} \Big[&\widehat{Incpt} + \widehat{\alpha}_{11} * \rho + \widehat{\alpha}_{12} * [\ln\rho]^2 + \widehat{\alpha}_{13} * [\ln\rho]^3 \\ &+ \widehat{\alpha}_{31} * \widehat{s}(\rho, X_i^{2006}) + \widehat{\alpha}_{32} * \left[\widehat{s}(\rho, X_i^{2006})\right]^2 + \widehat{\alpha}_{33} * \left[\widehat{s}(\rho, X_i^{2006})\right]^3 \\ &+ \widehat{\alpha}_3 * (\ln\rho * \widehat{s}(\rho, X_i^{2006}))\Big], \end{aligned} \tag{4}$$

where the coefficient estimates from Equation (3) are used. The whole dose-response function is obtained by computing Equation (4) for each treatment level by using a grid of values in the corresponding range of treatment variable.

In a final step, we derive the treatment effect and elasticity functions. The former is defined as a first derivative of $E[\ln \widehat{\omega}^{2007}(\rho)]$ with respect to the argument ρ. The latter function is computed in usual way $\partial E[\ln \widehat{\omega}^{2007}(\rho)]/(\partial\rho/\rho)$. The estimated employment elasticity with respect to R&D are of a particular interest for us, as it allows to directly compare our results with those reported in the previous literature. Following Hirano and Imbens (2004), confidence intervals around the estimated dose-response, treatment effect and elasticity functions are obtained by means of a bootstrap procedure.

4 Data sources and variable construction

4.1 Data sources

The principal data source is the EU Industrial R&D Investment Scoreboard. The R&D Scoreboard is an annual data set compiled and provided by the European Commission. Firstly released in 2004, it comprises data on R&D investment, as well as other financial and economic variables (e.g. net sales, operating profits, employees) for the top 1173 R&D global performers,[8] around half of which are based in the EU and another half are based outside the EU, and 483 of which are active in the high-tech sectors (see Table 2). In the present study we focus on the high-tech firms as, according to our data, innovation creates most jobs in the high-tech sectors. During the 2004-2012 period the overall employment growth was 22.6% in the EU companies. In the high-tech sectors employment increased by 49.2%, whereas only by 24.2% in the medium-high-tech and by 18.5% in the low-tech sectors (see section 4.2). We verify robustness of our results by extending the information set to include both high- and medium-tech firms as well as all firms in the sample.

In addition to economic and financial variables, the R&D Scoreboard also identifies the industrial sector (of the parent company) as well as the geographical region of R&D investment (according to the location of company's headquarter). The R&D Scoreboard data are reported in two ways. On the one hand, the R&D Scoreboard data are reported as national aggregates broken down by NACE Rev.1.1 in the Eurostat dissemination database. On the other hand, given that the presentation of the aggregated statistics per economic activity and per country has no data for certain economic activities and certain countries, the full set of data is also reported as broken down by individual enterprise group.

The R&D Scoreboard data set is compiled from companies' annual reports and accounts with

[8]In total in the Scoreboard data there are 1372 companies, from which 1173 are without missing observations.

reference date of the first of August of each year. For those companies, whose accounts are expected close to the cut-off date, preliminary information is used. In order to maximise the completeness and to avoid double counting, the consolidated group accounts of the ultimate parent company are used. Companies which are subsidiaries of another company are not considered separately. Where consolidated group accounts of the ultimate parent company are not available, subsidiaries are however included. In case of a demerger, the full history of the continuing entity is included, whereas the history of the demerged company goes only back as far as the date of the demerger to avoid double counting. In case of an acquisition or merger, the estimated figures for the year of acquisition are used along with the estimated comparative figures if available.[9]

An important caveat of the R&D Scoreboard data concerns sample selection, putting under question the general validity of our results. Given the underlying sampling and selection rules of the R&D Scoreboard data set – ranking and selecting companies according to the total amount of their R&D expenditures – the R&D Scoreboard is not a random sample. Hence the R&D Scoreboard data set may be criticised that it has a sample bias affecting the results, as it only represents the top R&D investors. However, given our interest in the employment effect of innovation, this issue is of second order magnitude, because we are covering almost the entire population of world-wide R&D expenditure (Moncada-Paterno-Castello *et al.*, 2010). The 1173 companies in our sample altogether represent around 80% of the world-wide business R&D expenditure. While small R&D investors and non-R&D-performers are excluded from the sample, the objective of the present study is to focus on the impact of R&D-driven innovation on employment, but not to examine the determinants of labour demand in the whole economy. Finally, the particular estimation approach we adopt in the present study allows us to estimate the counterfactual treatment effects also without a zero control group.

4.2 Dependent (response) variable

The dependent (response) variable is firm-specific employment measured by the number of employees. In order to calculate firm-specific employment, we use the average number of employees or, if the

[9] It is important to note that the R&D Scoreboard data are different from the official R&D statistics provided by statistical offices. The R&D Scoreboard data refers to all R&D financed by a particular company from its own funds, regardless of where the R&D activity is performed. Hence, because companies are identified with country of their registered head office which, in some cases, may be different from the operational or R&D headquarters. In contrast, the R&D statistics usually refers to all R&D activities performed by businesses within a particular sector and country, regardless of the location of the business's headquarters and regardless of the origin of the sources of finance. Second, the R&D Scoreboard collects data from audited financial accounts and reports, whereas the R&D statistics are compiled on the basis of statistical surveys, in general covering the known R&D performer. Further differences concern sectoral classifications (R&D statistics follows the classification of economic activities in the European Community, NACE Rev.1.1, whereas the R&D Scoreboard allocates companies in accordance to the sectoral classification as defined by the Financial Times Stock Exchange Index (ICB classification) and then converts them into NACE Rev.1.1. These differences need to be kept in mind when comparing the results reported in this paper to studies employing statistical R&D data.

annual average is not available, the number of employees at the end of the reference period. In total the companies included in the R&D Scoreboard data set employed 48471 million workers in 2012, 1.5% more than the previous year. The distribution of employees by region was 18357 million in the companies based in the EU, 11138 million in the US companies, 8206 million in the Japanese companies and 10770 million in the companies from other countries.[10]

The development of employment in companies contained in the R&D Scoreboard data over the 2004-2012 period can be summarised as follows. Overall, the worldwide employment increased by 27.9% from 2004 to 2012 led by increases in high-tech sectors (42.0%) and medium-high-tech sectors (29.9%). The overall employment growth was 22.6% in the EU companies, increasing by 49.2% in high-tech sectors, by 24.2% in medium-high-tech and by 18.5% in low-tech sectors. The overall employment growth (25.1%) in the US companies greatly varies by sector group: a strong increase for high-tech sectors (43.7%) and a sharp decrease in low-tech sectors (-23.2%). The overall employment increase of 24.0% in the Japanese companies corresponded to an increase of 31.4% in low-tech sectors and of 28.5% in medium-high-tech sectors. The ratio of employment in high-tech to medium-high-tech sectors for companies based in Japan fell from 38% to 32%, rose slightly for EU companies, from 29% to 35%, and went up a lot for US companies from 80% to 98%.

4.3 Explanatory (treatment) variable

We define the explanatory (treatment) variable, r_{it}, as the share of R&D investment in the total capital expenditure. The constructed measure of R&D intensity includes all cash investment in R&D funded by the companies themselves, but excludes any R&D undertaken under contract for customers, such as governments or other companies, and the companies' share of any associated company or joint venture R&D investment. R&D expenditures are calculated based on the R&D accounting definition set out in the International Accounting Standard (IAS) 38 "Intangible assets", which is based on the OECD "Frascati" manual. Research is defined as original and planned investigation undertaken with the prospect of gaining new scientific or technical knowledge and understanding. Expenditure on research is recognised as an expense when it incurred. Development is the application of research findings or other knowledge to a plan or design for the production of new or substantially improved materials, devices, products, processes, systems or services before the start of commercial production or use. Development costs are capitalised when they meet certain criteria and when it can be demonstrated that the asset will generate probable future economic benefits. Where part or all of R&D costs have been capitalised, the additions to the appropriate

[10]Note, however, that data reported by the Scoreboard companies do not inform about the actual geographic distribution of the number of employees. A detailed geographic analysis should take into account the location of subsidiaries of the parent Scoreboard companies as well as the location of other production activities involved in the value-chains.

intangible assets are included to calculate the cash investment and any amortisation eliminated.

In order to account for sectoral heterogeneity with respect to R&D intensity, we regroup all firms into four sub-samples according to the level of technological sophistication. Following the OECD classification, all firms in our sample are regrouped into four 3-digit Industry Classification Benchmark (ICB) groups: high-, medium-high-, medium-low-, and low-tech companies:

- *High-tech*: Technology hardware & equipment (THE), Software & computer services (SCS), Pharmaceuticals & biotechnology (PBT), Health care equipment & services (HCE), and Leisure goods (LGO);
- *Medium-high-tech*: Industrial engineering, Electronic & electrical equipment, General industrials, Automobiles & parts, Personal goods, Other financials, Chemicals, Aerospace & defence, Travel & leisure, Support services, and Household goods & home construction;
- *Medium-low-tech*: Food producers, Fixed line telecommunications, Beverages, General retailers, Alternative energy, Media, Oil equipment, services & distribution, and Tobacco;
- *Low-tech*: Gas, water & multi-utilities, Oil & gas producers, Nonlife insurance, Industrial metals & mining, Construction & materials, Food & drug retailers, Banks, Electricity, Industrial transportation, Mobile telecommunications, Forestry & paper, Mining, Life insurance.

The descriptive statistics of R&D activity for each group of companies is reported in Table 2. According to Table 2, the R&D activity of high-tech firms, measured both in absolute and relative terms, substantially exceeds that of medium-tech and low-tech companies. In the present study we focus on the high-tech sub-sample which, as reported in Table 2, contains 483 firms. We use data on firm-specific employment and R&D intensity for the years 2007 and 2006, respectively, in order to avoid devastating effects of the global financial crisis on the world economy.

4.4 Covariates

The set of covariates used in our analysis is selected based on previous studies (e.g. see Hall *et al.*, 2008, 2010), subject to their availability in our data set. It includes:

- *Net sales*, SALE: In line with the accounting definition of sales, sales taxes and shares of sales of joint ventures & associates are excluded. For banks, sales are defined as the "Total (operating) income" plus any insurance income. For insurance companies, sales are defined as "Gross premiums written" plus any banking income.
- *Operating profit*, OP: Profit (or loss) before taxation, plus net interest cost (or minus net interest income) and government grants, less gains (or plus losses) arising from the sale/disposal

of businesses or fixed assets. Due to the fact that companies report both positive and negative operating profit, we cannot take a logarithmic transformation of this variable. In order to do so, we created the following two variables $\ln \mathrm{OP}^{+}_{2006}$ and $\ln \mathrm{OP}^{-}_{2006}$. The former variable is equal to the log of actual values whenever a firm reports positive profit and zero otherwise. The latter variable is equal to the log of absolute actual values multiplied by minus one whenever a firm reports negative profit and zero otherwise.

- *Capital expenditure*, CAPEXP: The expenditure used by a company to acquire or upgrade physical assets such as equipment, property, industrial buildings. In company accounts capital expenditure is added to the asset account (i.e. capitalised), thus increasing the amount of assets. It is disclosed in accounts as additions to tangible fixed assets.

- *Market capitalisation*, MCAP: The share price multiplied by the number of shares issued at a given date. Market capitalisation data have been extracted from both the Financial Times London Share Service and Reuters. These reflect the market capitalisation of each company at the close of trading on 4 August 2006. The gross market capitalisation amount is used to take into account those companies for which not all the equity is available on the market.

- *Industry sectors*: The industry sectors are based on the ICB classification. The level of disaggregation is generally the three-digit level of the ICB classification, which is then converted to NACE Rev.1.1.

- *Sectoral dummies*: Sectors are classified into high-tech, medium-high-tech, medium-low-tech, and low-tech, according to 3-digit ICB groups.

- *Regional dummies*: "Asian Tigers" (AT), "BRIC", "EU", "Japan", "RoW", "Switzerland" (CH), and "USA".

5 Results

5.1 Main results

The results of the first step GPS estimation procedure (see Equation (1)) are reported in Table 3, which suggest that the variation in the R&D intensity is best captured by variables such as operating profits, market capitalisation and its square, as well as sales. Also the industry- and region-specific dummy variables contribute substantially to the explanatory power of the first step of the GPS regression. The goodness-of-fit of the regression is quite high, yielding the R^2 of 42%, creating a powerful GPS, see Equation (2). [11]

[11] The assumption of normally distributed GPS, see Equation (2), was verified by means of the Kolmogorov-Smirnov test. The associated p-value is 0.26.

Next, we verify whether the GPS is appropriately specified by testing the so-called balancing property, following the procedure suggested by Hirano and Imbens (2004). Each covariate is subdivided into three groups of approximately similar sizes using distribution of the treatment intensity variable. The initial testing of the balancing property amounts to testing whether the average value of a particular variable in every group is equal to the average value in the remaining groups. The results of these tests are reported in Table 4. Only for a handful of groups we cannot reject the tested null hypothesis at the usual significance levels, indicating that there is very strong heterogeneity among the covariates belonging to these three groups pertinent to different values of treatment intensity. A well specified GPS should be able to successfully account for these differences.

In order to check whether this is the case, we subdivided each group into blocs of approximately the same sizes corresponding to the quintiles of the respective GPS. The resulting cell sizes are reported in Table 5. Observe that the total number of firms, reported in Table 5, is less than reported above in Table 4, i.e. 442 vs. 483. This is because we imposed the so-called common support condition, ensuring that we deal with observations with similar GPS values but different treatment intensities. As argued by Becker *et al.* (2012), it is advisable to impose the common-support condition in order to substantially improve the balancing properties of the GPS and hence achieve more reliable estimation results.

The balancing properties of covariates adjusted for the GPS are reported in Table 6. Compared to the results for the unadjusted covariates reported in Table 4, there is a substantial improvement, as only three test statistics exceed the nominal 5% significance level.[12] The mean absolute value of all t-statistics reported in Table 6 drops to 0.90 from the corresponding value of 3.41 computed across all groups and covariates in Table 4. Hence we conclude that the generalised propensity scores are appropriately defined. Next, we proceed to the estimation of the dose-response relationship between the innovation and employment variables.

The estimation results for the second-step regression corresponding to Equation (3) are reported in Table 7. The estimated R^2 is 0.40, which is quite remarkable given the parsimonious specification. The second step GPS regression results reported in Table 7 clearly show that employment response to firm innovation (proxied by R&D expenditures) is highly non-linear with all included polynomial terms of the latter variable reporting highly significant coefficients. It is also worthwhile noticing that the GPS variable enters as a significant variable both in levels and via the interactive term with the (log) of our treatment variable, confirming its relevance in eliminating the sample selection bias.[13]

[12]Observe that given a total number of reported t-test statistics this empirical rejection rate approximately corresponds to the nominal test level of 5%.

[13]The higher order power transformations of the GPS variable turned out to be insignificant and therefore were omitted from the model specification for the sake of parsimony.

In order to facilitate the interpretation of the estimation results, we plot the estimated dose-response, treatment effect and elasticity functions in Figures 1, 2 and 3, respectively. In order to provide an idea about firm distribution for different R&D intensity levels, vertical lines are added to distinguish between the four quartiles. For example, in the high-tech sectors the bottom quartile contains firms with R&D intensity levels up to 160%. The cut-offs for the other three quartiles are at 350% and 690%, respectively.

At low innovation intensity levels (the share of R&D investment in the total capital expenditure between zero and 35-40%) an additional investment into R&D may even destruct jobs. This can be seen in the negative interval in Figure 3. The job destruction effect of moderate innovators can be explained by missing critical mass and insufficient absorptive capacity to benefit from intramural research in companies with insufficient innovative capacity. Our results are consistent with findings of Geroski (1998) as well as Kancs and Siliverstovs (2012), who find that a certain critical mass of R&D capacity is required, before a significant firm growth can be achieved from investment in R&D. Our results are also consistent with the hypothesis of absorptive capacity, which has been found to be important particularly for moderate innovators. Firms must be capable of absorbing and using new knowledge effectively, if they are to benefit from intramural and extramural R&D investment (which apparently is not the case at very low R&D levels) (Fabrizio, 2009). Another reason why increasing R&D expenditures may destruct jobs in modestly innovating companies could be a larger room for technological efficiency improvements. Given that modest innovators have a higher potential for creating a more efficient workforce and replacing workers through the acquisition of capital goods, the compensation may be only partial for modest innovators.

At medium to medium-high innovation intensity levels (the share of R&D investment in the total capital expenditure is around 100%) the innovation impact on employment is positive and statistically significantly different from zero. The employment elasticity with respect to innovation goes up to 0.7%, which implies that increasing innovation by 1% raises employment by 0.7%. There may be several reasons, why innovation followers create more jobs than modest innovators. First, through new investments. Given that the convergence between falling costs and lower prices is not instantaneous, extra profits that are accumulated by innovative firms are often reinvested.[14] A larger production capacity in innovation followers requires more workers and hence creates more jobs. Second, by increasing income. Given that more improvements in productivity are transmitted to higher wages in innovation followers, likely they will induce higher consumption and hence higher employment. According to Leonardi (2003), more educated workers (which are employed in more innovative firms) consume more skill-intensive goods.[15] Hence, an increase in the income

[14]Note, however, that the new investments can be capital-intensive, which may partially mitigate the compensation effect.

[15]Leonardi (2003) derives theoretically in a general equilibrium model, and estimates empirically for the UK that

of high-skill workers's income increases the demand for skill-intensive goods, resulting, in such a way, in higher output of innovative firms in high-tech sectors employing high-skill workers. An increase in aggregate demand in turn increases production and employment. Third, through new products/varieties resulting from innovation. New products/varieties entail a creation of new jobs in innovation followers, but a destruction of jobs in modest innovators Bogliacino *et al.* (2012); Bogliacino and Vivarelli (2012); Bogliacino *et al.* (2013). Finally, through a decrease in output prices, resulting in lower production costs, which stimulates demand for innovative firms' products and, as a result, increases demand for workers.

The job creation effect of innovation reaches its peak when R&D intensity is around 100% of the total capital expenditure, after which the positive employment effect declines and becomes statistically indifferent from zero. At high and very high innovation intensity levels (the share of R&D investment in the total capital expenditure above 150%) the innovation impact on employment becomes negative again, implying that, on average, additional R&D investment in innovation leaders destructs jobs.[16] These results of decomposing the employment effect by innovation intensity, suggesting that the displacement effect seems to be larger than the compensation effect for innovation leaders, whereas the compensation effect seems to be greater than the displacement effect for innovation followers, are new and have not been reported in the literature before.

5.2 Comparison with previous studies

Our estimation results complement those of Ciriaci *et al.* (2013), Bogliacino *et al.* (2012), Bogliacino and Vivarelli (2012) and Bogliacino (2014) who provide the initial attempts to decompose the employment effect of innovation according to R&D intensity levels. Using the balanced panel comprising characteristics of 3300 Spanish firms observed for the period 2002—2009, Ciriaci *et al.* (2013) investigate the employment effects of innovation both for innovative and non-innovative firms. Ciriaci *et al.* (2013) find that those firms, which engage more intensively in innovation activities, create more jobs than less innovative firms. In particular, this effect is more pronounced for small and young innovative firms. At the same time they point out that for this group of firms, successful launch of new products in the market as a result of boosting innovation activity can lead to a higher growth in sales rather than in employment, which is consistent with the labour-saving effects of technological advances, discussed above.

Bogliacino *et al.* (2012) studies the employment effect of R&D expenditure using the sample of 677 EU firms observed during the period 1990—2008. The elasticities of interest are estimated

more educated workers demand more skill-intensive goods. According to Leonardi (2003), in the UK the induced demand shift can explain 3% of the total relative demand shift between 1981 and 1997.

[16] At extremely high R&D intensity levels (above 1300%) our results suggest positive employment effect of innovation again. However, the number of firms with such an extremely high R&D intensity is very small in our sample (and in the population). Therefore, these results for very high innovation intensity levels should be considered with caution.

using the dynamic panel model allowing for lagged employment by means of the Least Squares Dummy Variable Corrected (LSDVC) estimator (Bun and Kiviet, 2003; Bruno, 2005). The results are obtained for the sample of all firms as well as for the samples comprising service-sector firms, all manufacturing firms and samples comprising manufacturing firms further subdivided into high-tech and non-high-tech firms. The reported short-run elasticities are 0.023% for the whole sample, 0.056%—for service-sector firms, and 0.049% for high-tech manufacturing firms. It is interesting to observe that the corresponding elasticity estimate for non-high-tech manufacturing firms is not significant though positive (0.021%). Using the estimated coefficient on the lagged employment variable Bogliacino *et al.* (2012, Table 1), the long-run elasticities can be derived. The long-run elasticity of employment calculated for the whole sample is 0.075%, 0.097%—for service-sector firms and approximately of equal magnitude of 0.11% both for all manufacturing firms and high-tech manufacturing firms.

Bogliacino and Vivarelli (2012) conduct another study of employment effects of innovation activity using a sample of 2295 firms from 15 European countries available for the period 1996—2005. In the main part of the paper, the results are reported for a number of dynamic panel data estimators such as random-effects, fixed-effects as well as two versions of the Generalised Method of Moments [GMM-DIF, Arellano and Bond (1991)] and [GMM-SYS, Blundell and Bond (1998)], where the last estimator is referred to as the most reliable one (Bogliacino and Vivarelli, 2012, Section IV). These estimators are applied for the whole sample of firms. The short-run elasticity reported by the GMM-SYS estimator is 0.025%, which is very similar to that reported in Bogliacino *et al.* (2012). However, the long-run elasticity is about 0.31%, which is about four times larger than that reported in Bogliacino *et al.* (2012) for the whole sample (0.075%). In the section containing robustness results, the distinction is made between firms with different innovation intensity by allowing for differential employment effects of high-tech, medium-tech and low-tech firms. The elasticities of interest are obtained by means of the LSDVC rather than GMM estimator, while it is argued that the former one outperforms the latter one under given estimation conditions. The main result is that the job creation effect of R&D expenditure only is evident for high-tech sector; both for medium- and low-tech sectors the estimated short-run elasticities are not significantly different from zero. For the high-tech sector, the short- and long-run elasticities are 0.017% and 0.17%, respectively.

Our results, emphasising a complex non-linear interaction of employment and innovation, are also consistent with those of Bogliacino (2014), who equally finds that R&D investment expenditure has a non-linear effect on firm employment. Bogliacino (2014) also reports that productivity impact of R&D takes significant lags, whereas employment effect is effective already since the beginning of the R&D process. According to Bogliacino (2014), both the intensive and the extensive margins of R&D work in the same direction: for a given firm size, increasing the R&D intensity raises the employment elasticity, and for a given R&D intensity, increasing the firm size increases also the employment effect. These results confirm our policy conclusions that policy makers have two alternative policy strategies for targeting the innovation and employment objectives: policy instruments aiming at the growth of innovative companies, and policy instruments aiming at increasing the number of companies that undertake innovative activities in the EU.

It is instructive to compare our results with the traditional point estimates available in the literature, despite the fact that the studies cited above focus on the employment elasticity with respect to nominal measure of R&D expenditure, whereas we focus on the employment elasticity with respect to the relative measure of R&D expenditure. Even though the range of elasticity estimates for high-tech companies reported in our study in Figure 3 is quite large, as it varies with the level of R&D intensity [-0.80%, 0.70%], the positive values of the employment elasticity observed for the firms pertaining to the lower quartile of the R&D intensity distribution are in the comparable range with the estimates of the long-run employment elasticities of R&D, as discussed above.

Our results are also consistent with the evidence from general equilibrium macroeconomic models, which simulate the impact of R&D and innovation policies (Brandsma and Kancs, 2015; Di Comite and Kancs, 2015; Di Comite *et al.*, 2015). The simulated employment effects of innovation in macroeconomic models capture important general equilibrium effects and vertical and horizontal linkages between firms, which is not possible in micro-econometric studies, such as the one presented in this paper. Combining the micro and macro approaches for studying the employment effect of innovation is a promising area for future research.

5.3 Robustness checks

In order to verify the robustness of the results reported in the previous section, we perform several robustness checks. First, we re-estimate regressions in Equations (2) and (3) using enlarged data sets that include both high- and medium-tech firms as well as all firms in our sample. Secondly, we re-estimates the second-step regression using our preferred sample of high-tech firms but after taking the logarithmic transformation of the GPS variable obtained in the first step, similarly as it was done in the empirical example in Hirano and Imbens (2004). The estimation results are presented in Table 8.

In panel (A) of Table 8 we report the estimation results of Equation (3) keeping square and cubic transformations of the score variable. As seen, these are not significant at the usual levels, and therefore the more parsimonious form of the regression, reported in Table 7, is preferred. This choice of the model specification is also supported in terms of the Schwarz Information Criterion (SIC).

In panel (B) the estimation results are reported using enlarged information set including both high- and medium-high tech firms. This results in correspondingly increased sample size of 771 firms. However, such a model specification yields a much lower explanatory power with the reported goodness-of-fit measure decreasing from $R^2 = 0.40$ in Table 7 to $R^2 = 0.21$. In addition, the coefficient pertaining to the score variable in levels is no longer significantly different from zero. Only the interaction terms between the score and (log of) the treatment variables remains statistically significant. The further increase of the sample size incorporating all the firms in our sample yields similar results, see panel (C). We observe a further decrease in the regression explanatory power with the reported $R^2 = 0.15$ and the score variable s is not significant in this model. After comparing these estimation results with those reported in Table 7, we can conclude that focusing on a smaller data set involving less heterogeneous firms yields more clear-cut results that are statistically superior to those obtained using larger pool of more diverse firms. It seems that for the latter data set more explanatory variables are needed than we have at hand in order to account for the inherent firm heterogeneity.

In panel (D) of Table 7 we report the estimation results of the second-step regression, where we inserted the logarithmic transformation of the score variable. Also in this case we observe that the underlying model in Table 7 is statistically superior to that one both in terms of the regression explanatory power and the SIC values.

6 Conclusions, policy recommendations and limitations

The question of whether technological change creates or destroys jobs has been posed since the beginning of the Industrial Revolution in the 19th century. While, the theoretical models, the estimation strategy and the empirical evidence have improved significantly since then, the key questions and challenges surrounding the innovation-employment relationship remain. The present paper aims to contribute to this literature by attempting to identify the R&D intensity levels under which firm innovation is complementary with respect to firm employment and under which it may have an adverse impact on firm employment. The objective of the study is to reveal the entire innovation-employment relationship, which is done by estimating the employment effect of innovation for different R&D intensity levels in a continuous framework. This is our main contribution to the literature and policy debate; to the best of our knowledge no comparable studies analysing the employment effect of innovation in a continuous setting are available in the literature.

In order to answer these questions, we base our empirical micro-econometric analysis on a large international firm-level panel dataset, and our proxy for technology is a measurable and continuous variable, while most of the previous studies have relied on either indirect proxies of technological change or dummy variables (such as the occurrence of product and process innovation).

In particular, we employ the EU Industrial R&D Investment Scoreboard data set, which comprises data on R&D investment, as well as other financial and economic variables for the top 1173 R&D global performers, 483 of which are active in high-tech sectors, which we analyse in detail, as high-tech companies create the most jobs both in absolute and relative terms. In addition to firm-level innovation expenditures, we make use also of economic and financial variables, which allow us to control for important firm-specific effects. Moreover, the R&D Scoreboard also identifies the industrial sector (of the parent subsidiary) as well as the geographical region of R&D investment (according to the location of firm's headquarter), which allows us to control for fixed sector-specific and location-specific effects.

In order to decompose the employment effect by innovation intensity, we employ flexible semi-parametric methods, which allow us to recover the full functional relationship between R&D investment and firm employment. This is not possible in the standard estimation approach, which yields only point estimates and hence may hide important non-linearities in the innovation-employment relationship (Kancs and Siliverstovs, 2012). We use semi-parametric methods for causal inference in quasi-experimental settings with continuous treatments, by considering the innovation expenditure of firms as a continuous treatment and employment at the firm-level as an outcome. The functional form of the impact of R&D expenditure on firm employment is identified under the assumption of weak unconfoundedness, implying that the systematic information in innovation expenditure can be conditioned out by controlling for observable determinants of innovation expenditure, achieving quasi-randomisation. This allows us to address important estimation issues, such as the simultaneity bias, from which many previous studies suffer (Rosenbaum and Rubin, 1983; Hirano and Imbens, 2004).

Our results confirm previous findings that innovation can both create and destruct jobs (which, as we show, depends strongly on the innovation intensity). Second, the relationship between innovation and job creation is highly non-linear. At low innovation intensity levels (the share of R&D investment in the total capital expenditure between zero and 35-40%) an additional investment into R&D may even destruct jobs. At medium to medium-high innovation intensity levels (R&D intensity around 100%) the innovation impact on employment is positive and statistically significantly different from zero. The employment elasticity with respect to innovation is 0.7%, which implies that increasing innovation by 1% raises employment by 0.7%. The job creation effect of innovation reaches its peak when the R&D intensity is around 100% of the total capital expenditure, after which the positive employment effect declines and becomes statistically indifferent from zero. At high and very high innovation intensity levels (the share of R&D investment in the total capital expenditure above 150%) the innovation impact on employment becomes negative again, implying that, on average, additional R&D investment in highly innovative companies destructs jobs. These results of decomposing the employment effect by innovation intensity are new and have not been reported in the literature before.

Our results have important messages for policy makers. First, our findings confirm the important role that innovation followers can play in creating jobs and in ensuring the sustainability of high employment in the medium- to long-run. In light of the results of Crepon *et al.* (1998),[17] two alternative policy strategies can be identified how policy makers can target this objective: policy instruments aiming at the growth of innovation followers creating jobs, and policy instruments aiming at increasing the number of innovation followers, as they both undertake innovative activities and create employment in the EU. Second, our results point to potential complementarities between the two Europe 2020 policy targets aiming to increase the R&D/GDP ratio and the employment rate, particularly by supporting innovation followers. Indeed, the empirical evidence, which we provide in this study, supports the view that R&D expenditures can be beneficial to job creation capacity. These findings imply that both the innovation and employment targets of the Europe 2020 Strategy can be reached simultaneously, by designing tailored policies for innovation followers as they create most of the employment in the economy. On the other hand, our results suggest that innovation leaders and modest innovators tend to destruct jobs through additional investment into R&D, implying that these companies should not by targeted by the policy to achieve both the innovation and employment targets of the Europe 2020 Strategy. According to Kancs and Siliverstovs (2012), innovation leaders are key for achieving the innovation the innovation target of the Europe 2020 Strategy by boosting firm productivity and competitiveness. In summary, the findings of the present study and Kancs and Siliverstovs (2012) suggest that innovation leaders should be targeted, if policy objective is to boost productivity and competitiveness, whereas innovation followers should be targeted, if policy objective is to achieve both the innovation and employment targets of the Europe 2020 Strategy.

Turning to limitations, an important caveat of our empirical analysis concerns the nature of the Scoreboard sample. First, while other data sets, such as the OECD BERD data, can be considered as fully representative of the OECD economies, in the EU Industrial R&D Investment Scoreboard data used in the present study only the R&D "champions" are considered. This is a clear limitation of our data, the results of which cannot be straightforwardly extrapolated to e.g. SMEs. However – notwithstanding this source of sample selection – our analysis still provides interesting insights, and in addition has support from the empirical evidence on concentration of innovative activities. It is well documented in the previous literature that innovative activities are highly concentrated – only a small share of firms around the world innovate, the majority of firms in most regions around

[17] The model of Crepon *et al.* (1998) distinguishes between four stages of innovation process: the decision to innovate, the decision on how much to spend on innovation activities, the relation between expenditure on innovation and innovation output, and the relation between innovation output and performance.

the world do not engage in any significant R&D activities, they imitate (Slivko and Theilen, 2014). Hence, by considering the top 1173 innovators which account for almost 80% of the global R&D expenditure (top 2500 companies account for more than 90% of the global R&D expenditure), ensures also certain representativeness. A further limitation of the data used in our study is that the R&D Scoreboard data does not allow us to identify the effects of product and process innovations separately. However, as discussed in the introduction, the employment effect of innovation can be very different depending on the nature of innovation. In order to separately identify the employment effects of product and process innovation, other sources of data, such as Community Innovation Survey (CIS), need to be used, which is a promising area for future research.

References

ANTONUCCI, T. and PIANTA, M. (2002). Employment effects of product and process innovation in Europe. *International Review of Applied Economics*, **16** (3), 295–307.

ARELLANO, M. and BOND, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, **58** (2), 277–97.

BECKER, S. O., EGGER, P. H. and VON EHRLICH, M. (2012). Too much of a good thing? On the growth effects of the EU's regional policy. *European Economic Review*, **56** (4), 648–668.

BLUNDELL, R. and BOND, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, **87** (1), 115–143.

BOGLIACINO, F. (2014). Innovation and employment: A firm level analysis with European R&D Scoreboard data. *Economia*, **15** (2), 141–154.

—, LUCCHESE, M. and PIANTA, M. (2013). Job creation in business services: Innovation, demand, and polarisation. *Structural Change and Economic Dynamics*, **25**, 95–109.

—, PIVA, M. and VIVARELLI, M. (2012). R&D and employment: An application of the LSDVC estimator using European microdata. *Economics Letters*, **116** (1), 56–59.

— and VIVARELLI, M. (2012). The job creation effect of R&D expenditures. *Australian Economic Papers*, **51** (2), 96–113.

BRANDSMA, A. and KANCS, D. (2015). RHOMOLO: A Dynamic General Equilibrium Modelling Approach to the Evaluation of the EU's R&D Policies. *Regional Studies*, **49**, http://dx.doi.org/10.1080/00343404.2015.1034665.

BRUNO, G. S. F. (2005). Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals. *Stata Journal*, **5** (4), 473–500.

BUN, M. J. G. and KIVIET, J. F. (2003). On the diminishing returns of higher-order terms in asymptotic expansions of bias. *Economics Letters*, **79** (2), 145–152.

CIRIACI, D., MONCADA-PATERNO-CASTELLO, P. and VOIGT, P. (2013). *Innovation and job creation: A sustainable relation?* Working Papers on Corporate R&D and Innovation 2013-01, European Commission, DG Joint Research Centre.

Crepon, B., Duguet, E. and Mairesse, J. (1998). Research, Innovation And Productivity: An Econometric Analysis At The Firm Level. *Economics of Innovation and New Technology*, **7** (2), 115–158.

Di Comite, F. and Kancs, D. (2015). *Macro-Economic Models for R&D and Innovation Policies: A Comparison of QUEST, RHOMOLO, GEM-E3 and NEMESIS*. IPTS Working Papers JRC94323, European Commission, DG Joint Research Centre.

—, — and Torfs, W. (2015). *Macroeconomic Modelling of R&D and Innovation Policies: An application of RHOMOLO and QUEST*. IPTS Working Papers JRC89558, European Commission, DG Joint Research Centre.

Doornik, J. and Hendry, D. (2009). *Empirical Econometric Modelling Using PcGive 13*. London: Timberlake Consultants Press and Oxford: www.doornik.com.

Doornik, J. A. (2007). *Object-Oriented Matrix Programming Using Ox*. London: Timberlake Consultants Press and Oxford: www.doornik.com.

Egger, P. H., von Ehrlich, M. and Nelson, D. R. (2012). Migration and Trade. *The World Economy*, **35** (2), 216–241.

European Commission (2013). *EUROPE 2020: A strategy for smart, sustainable and inclusive growth*. Brussels: European Commission.

European Commission (2015). *Mid-term review of the EUROPE 2020 strategy for smart, sustainable and inclusive growth*. Brussels: European Commission.

Fabrizio, K. R. (2009). Absorptive capacity and the search for innovation. *Research Policy*, **38** (2), 255–267.

Freeman, C., Clark, J. and Soete, L. (1982). *Unemployment and Technical Innovation*. London: Pinter.

— and Soete, L. (1987). *Technical Change and Full Employment*. Oxford: Basil Blackwell.

Geroski, P. A. (1998). An applied econometrician's view of large company performance. *Review of Industrial Organization*, **13** (3), 271–294.

Hall, B. H., Lotti, F. and Mairesse, J. (2008). Employment, innovation, and productivity: evidence from Italian microdata. *Industrial and Corporate Change*, **17** (4), 813–839.

—, Mairesse, J. and Mohnen, P. (2010). Measuring the returns to R&D. In B. H. Hall and N. Rosenberg (eds.), *Handbook of the Economics of Innovation*, vol. 2, *22*, Elsevier, pp. 1033–1082.

HEINRICH, C., MAFFIOLI, A. and VÁZQUEZ, G. (2010). *Impact-evaluation guideline*. Technical Notes No. IDB-TN-161, Inter-American Development Bank.

HIRANO, K. and IMBENS, G. W. (2004). The propensity score with continuous treatments. In A. Gelman and X.-L. Meng (eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, Chichester: Wiley, pp. 73–84.

KANCS, D. and CIAIAN, P. (2011). Modelling the flow of knowledge and human capital: A framework of innovative capital. *International Journal of Public Policy*, **7** (1), 134–160.

— and SILIVERSTOVS, B. (2012). *R&D and non-linear productivity growth of heterogeneous firms*. Working Papers on Corporate R&D and Innovation 2012/06, European Commission, DG Joint Research Centre.

KATSOULACOS, Y. (1986). *The Employment Effect of Technical Change*. Brighton: Wheatsheaf.

LACHENMAIER, S. and ROTTMANN, H. (2007). Employment effects of innovation at the firm level. *Journal of Economics and Statistics*, **227** (3), 254–272.

LEONARDI, M. (2003). *Product Demand Shifts and Wage Inequality*. Discussion Papers 908, Institute for the Study of Labor.

LUCCHESE, M. and PIANTA, M. (2012). Innovation and employment in economic cycles. *Comparative Economic Studies*, **54** (2), 341–359.

MONCADA-PATERNO-CASTELLO, P., CIUPAGEA, C., SMITH, K., TUBKE, A. and 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** (4), 523–536.

PIANTA, M. (2004). Innovation and employment. In J. Fagerberg, D. C. Mowery and R. R. Nelson (eds.), *Handbook of Innovation*, *22*, Oxford University Press.

PIVA, M. and VIVARELLI, M. (2005). Innovation and employment: Evidence from Italian microdata. *Journal of Economics*, **86** (1), 65–83.

ROSENBAUM, P. R. and RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70** (1), 41–50.

SLIVKO, O. and THEILEN, B. (2014). Innovation or imitation? The effect of spillovers and competitive pressure on firms R&D strategy choice. *Journal of Economics*, **112** (3), 253–282.

VAN REENEN, J. (1997). Employment and Technological Innovation: Evidence from U.K. Manufacturing Firms. *Journal of Labor Economics*, **15** (2), 255–84.

Vivarelli, M. (1995). *The economics of technology and employment: theory and empirical evidence.* Cheltenham: Edward Elgar.

— (2007). *Innovation and Employment: A Survey.* Discussion Papers 2621, Institute for the Study of Labor.

— (2013). Technology, Employment and Skills: An Interpretative Framework. *Eurasian Business Review*, **3** (1), 66–89.

— (2014). Innovation, employment and skills in advanced and developing countries: A survey of economic literature. *Journal of Economic Issues*, **48** (1), 123–154.

Young, A. (1993). Substitution and complementarity in endogenous innovation. *Quarterly Journal of Economics*, **108** (3), 775–807.

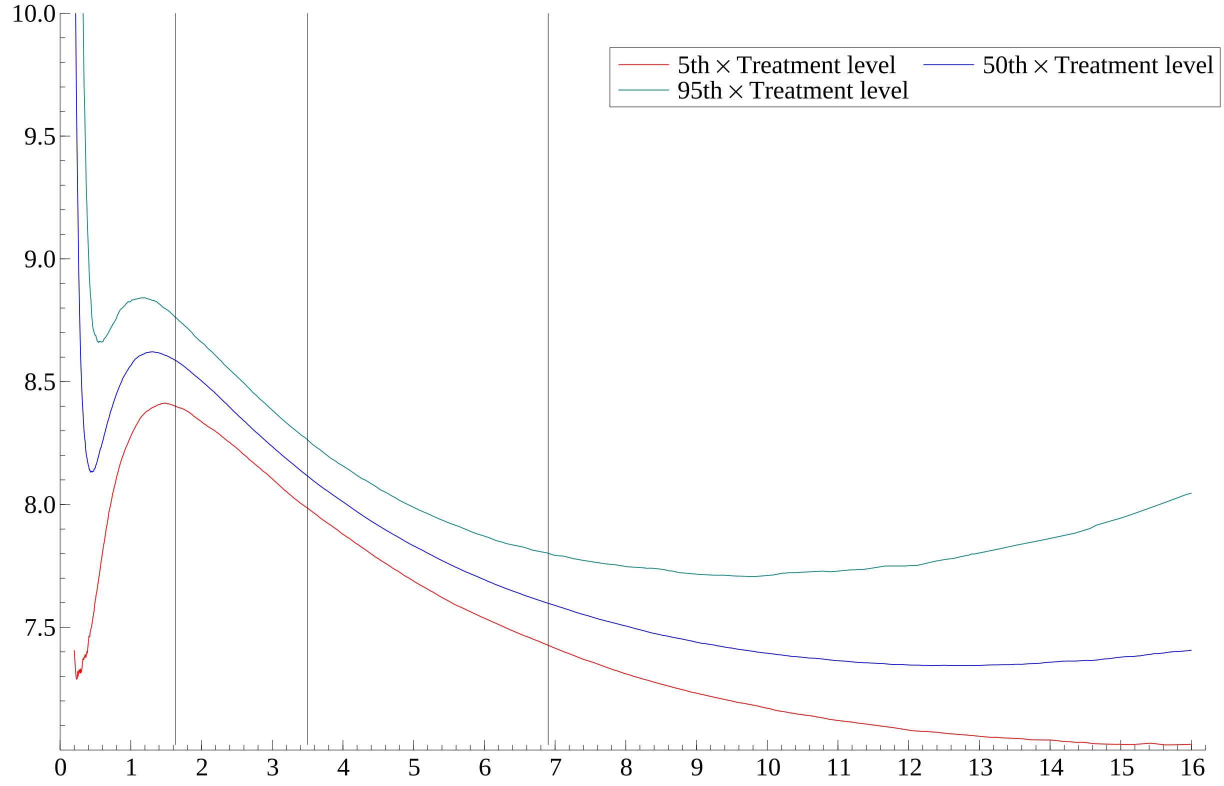

Figure 1: Dose-Response function of **high-tech** companies: Average expected response of employment (2007) [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

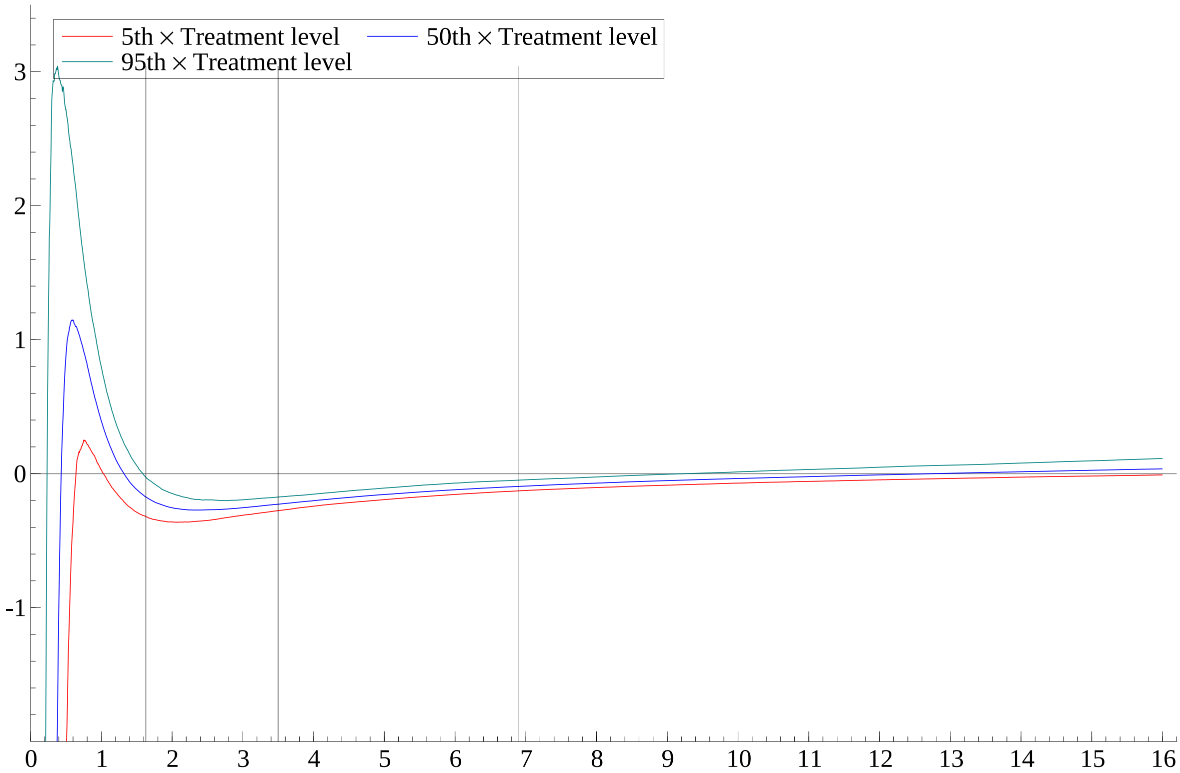

Figure 2: Treatment Effect function of **high-tech** companies: Derivative of the average expected response of employment (2007) [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

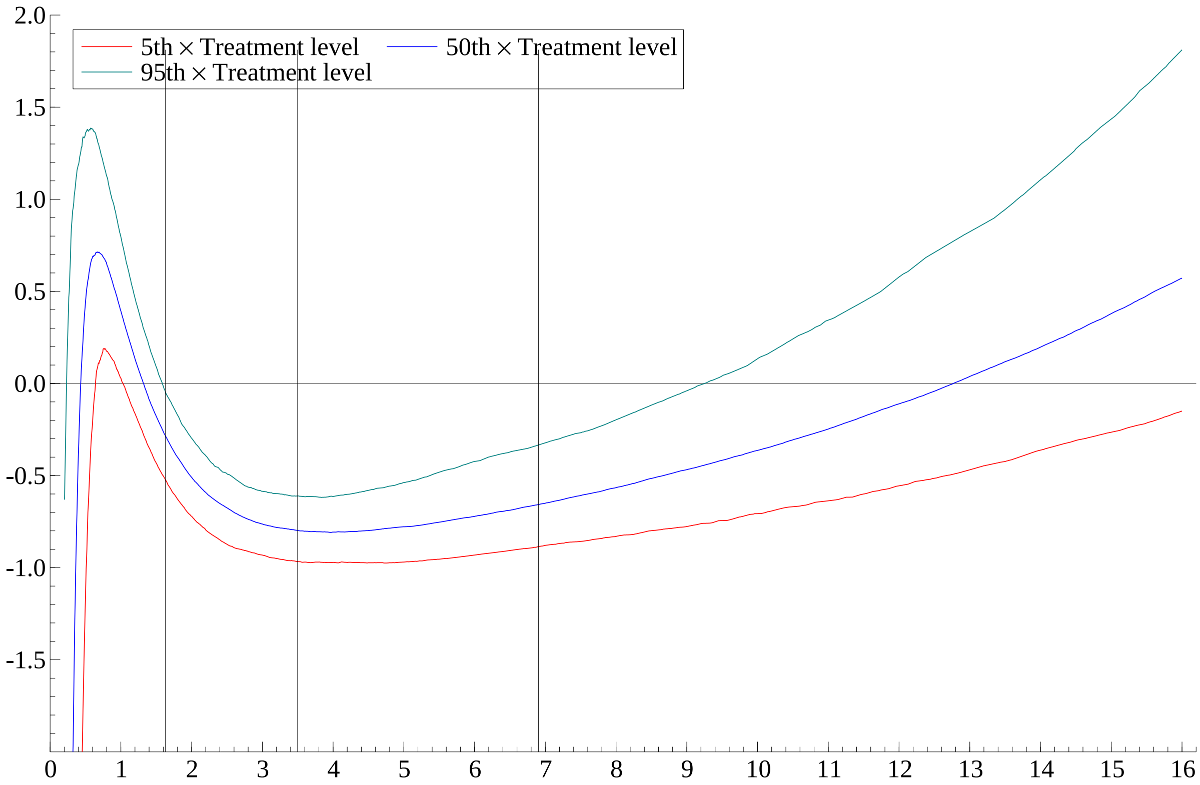

Figure 3: Elasticity of **high-tech** companies: Average expected response of employment in 2007 [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

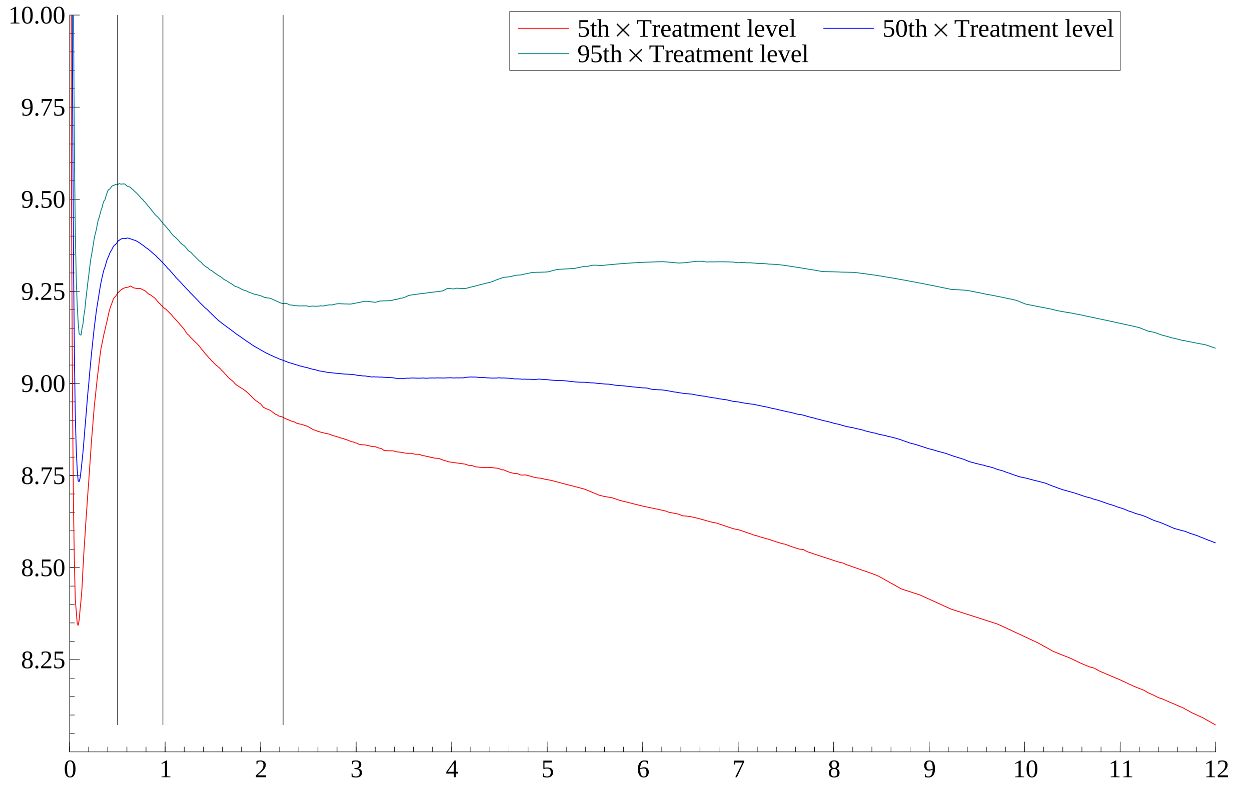

Figure 4: Dose-Response function of **all** companies: Average expected response of employment (2007) [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

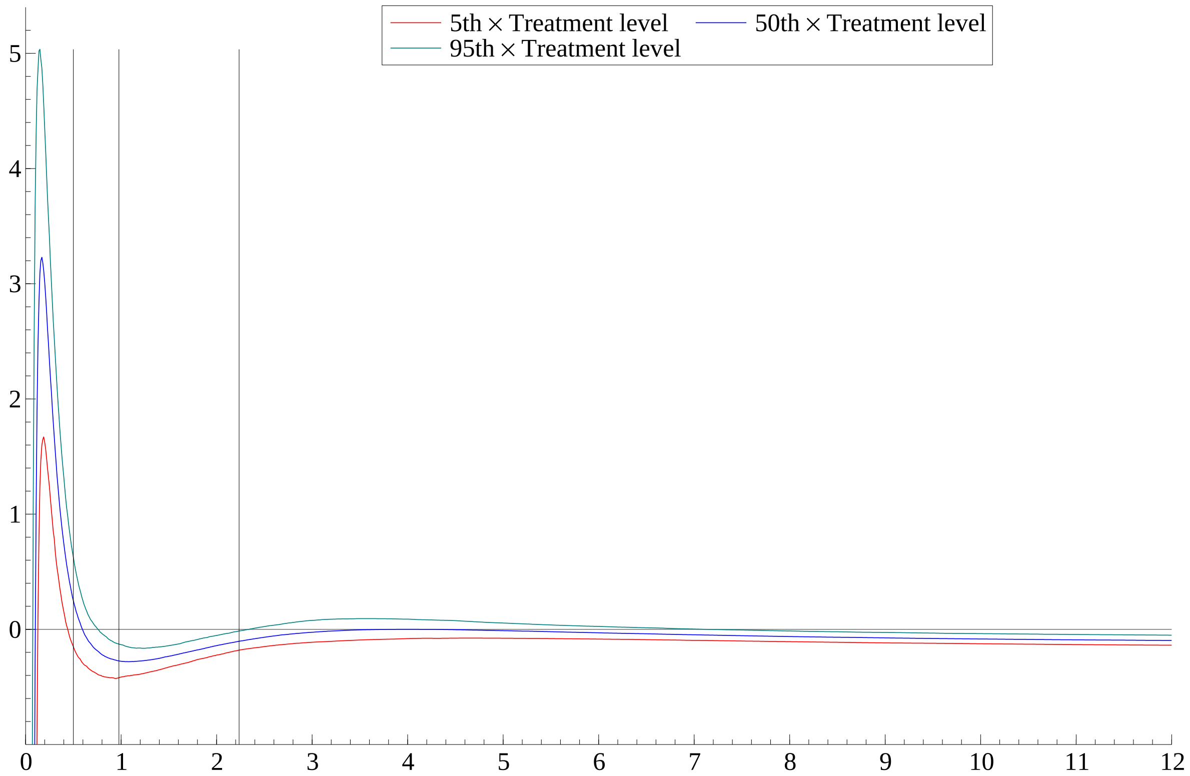

Figure 5: Treatment Effect function of **all** companies: Derivative of the average expected response of employment (2007) [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

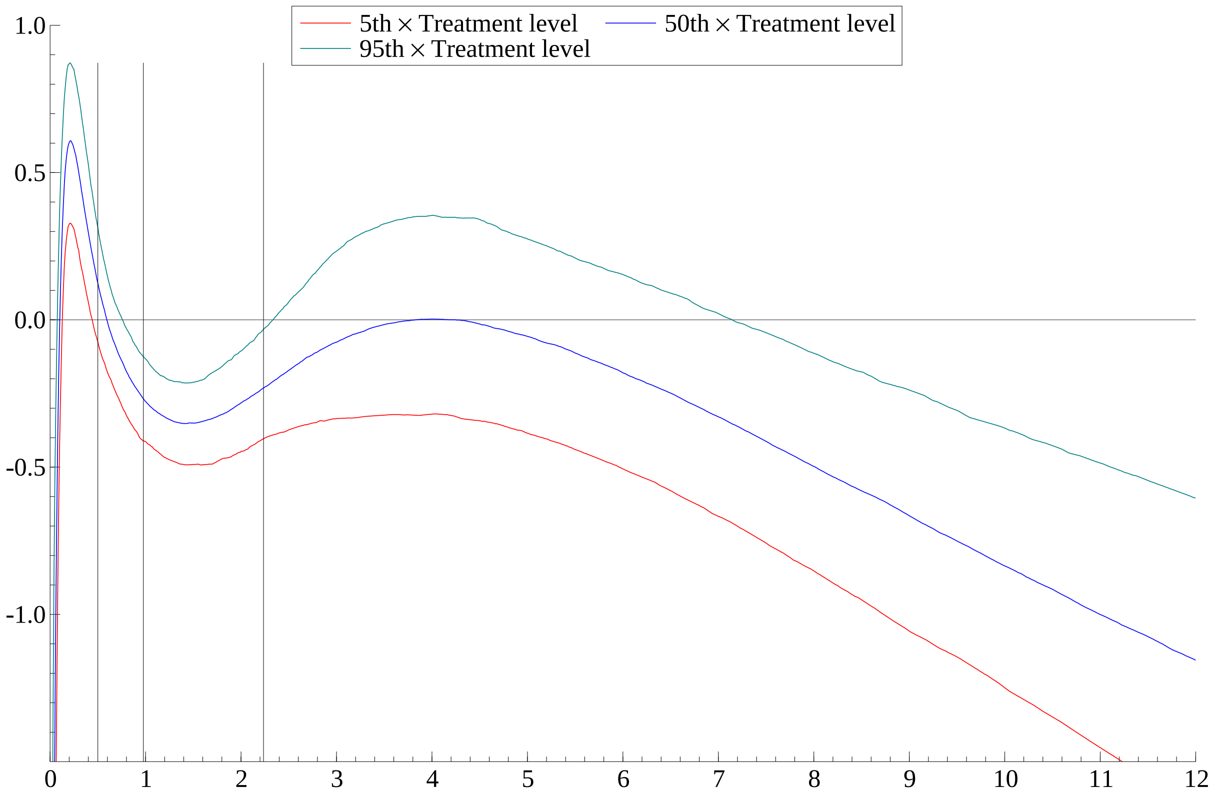

Figure 6: Elasticity of **all** companies: Average expected response of employment in 2007 [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

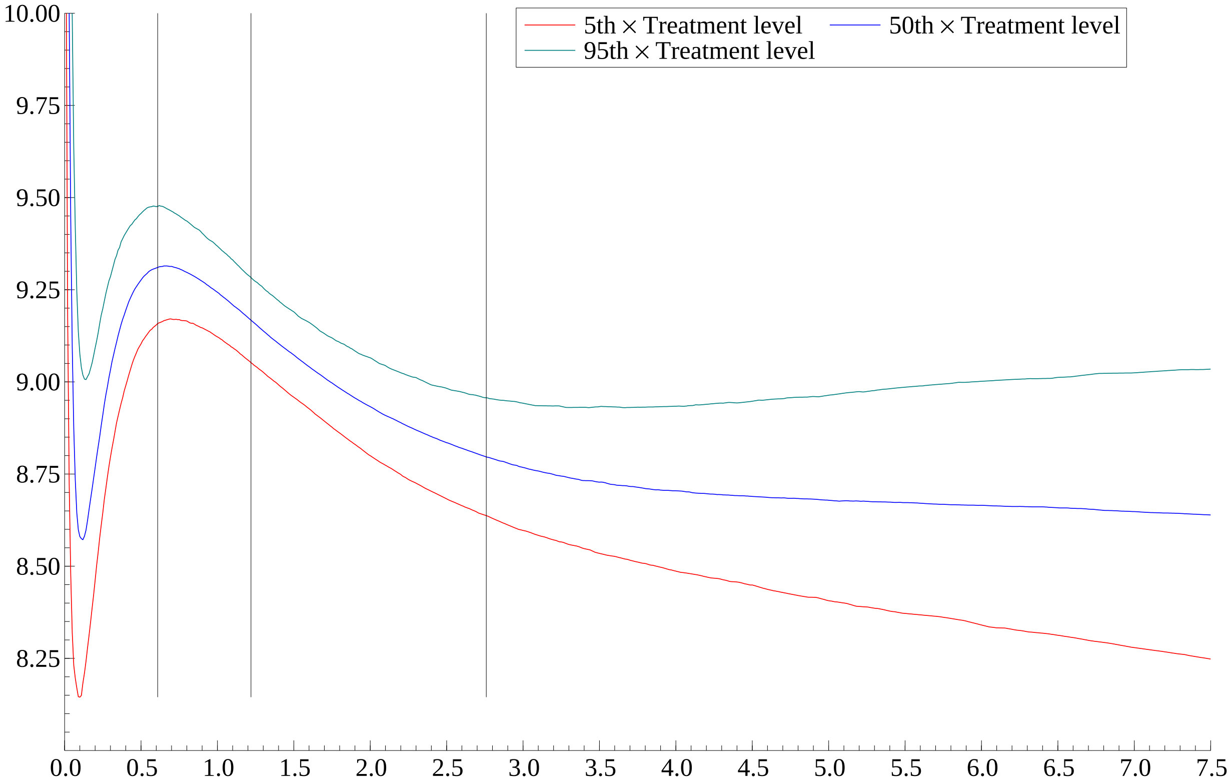

Figure 7: Dose-Response function of **high- and medium-high-tech** companies: Average expected response of employment (2007) [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

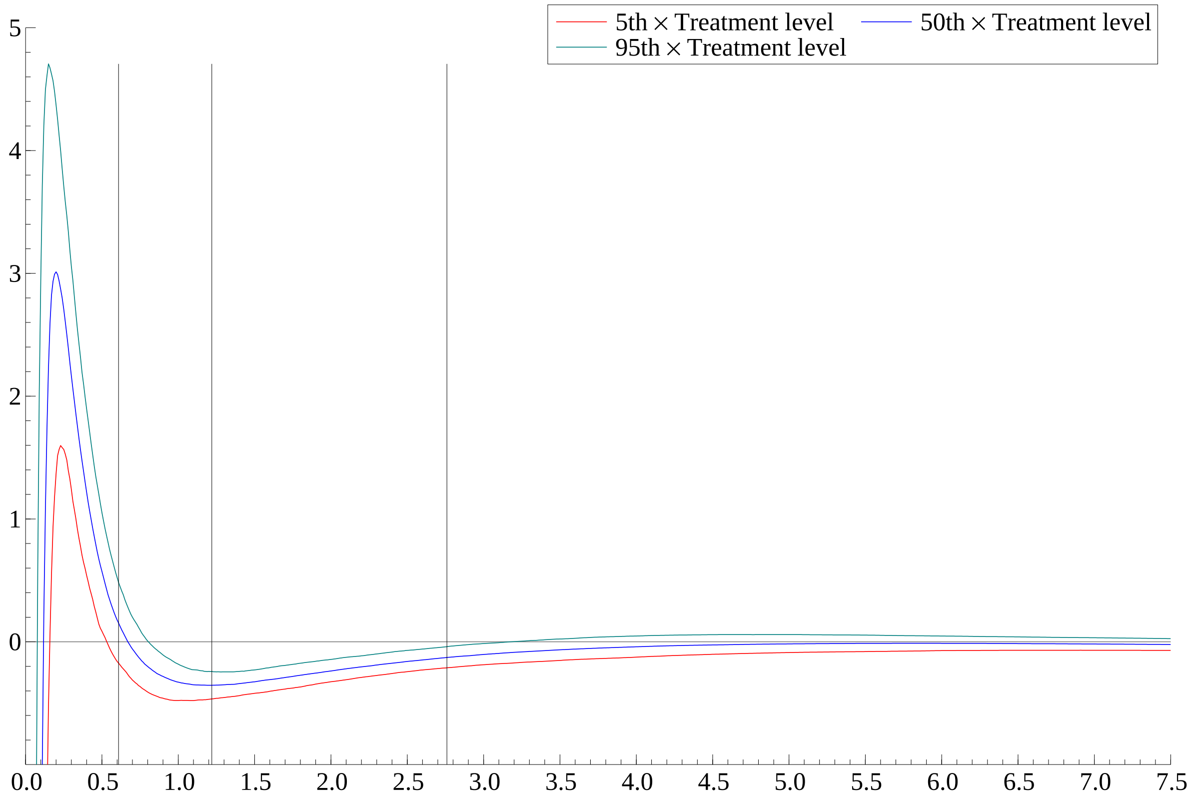

Figure 8: Treatment Effect function of **high- and medium-high-tech** companies: Derivative of the average expected response of employment (2007) [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

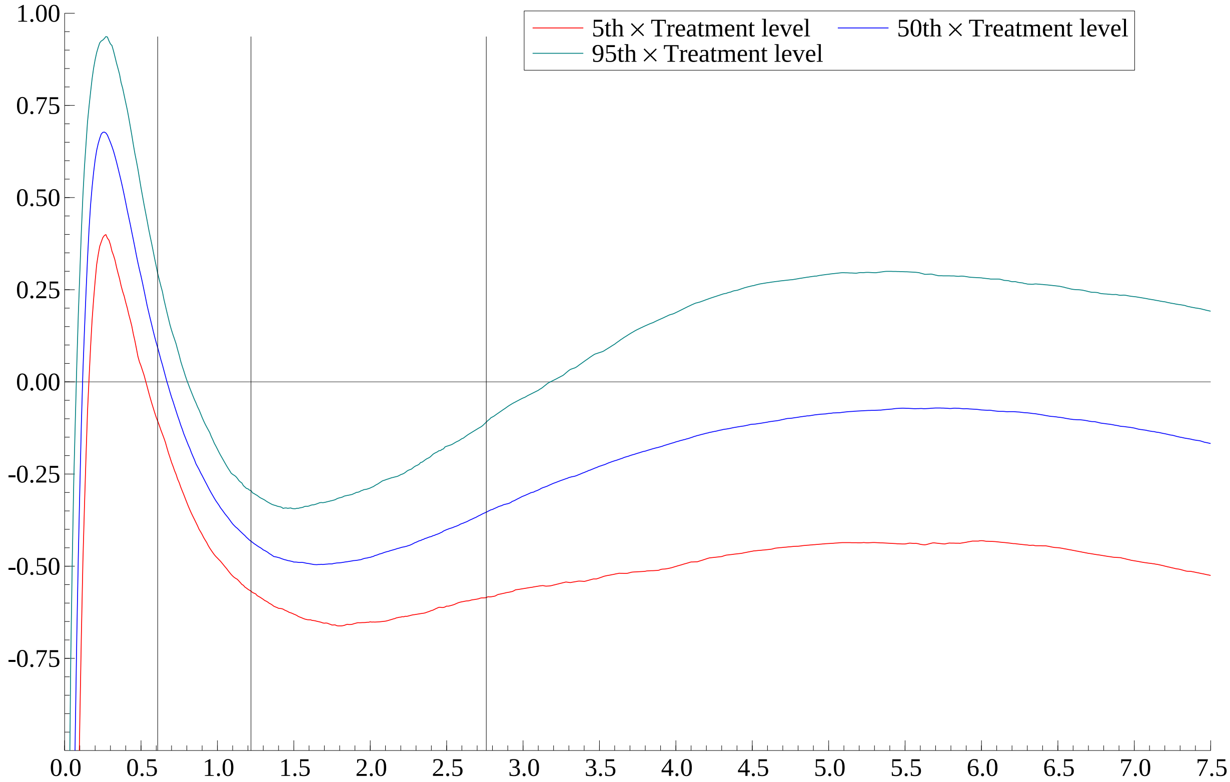

Figure 9: Elasticity of **high- and medium-high-tech** companies: Average expected response of employment in 2007 [Y-axis] to R&D intensity in 2006 [X-axis], GPS-adjusted. Dashed lines: bootstrapped 90 % confidence interval based on 1000 replications. Vertical lines denote quartiles of the R&D intensity distribution.

Table 1: Top 20 global innovation leaders in 2014

Rank	Company	Industry	R&D*	ICB classification
1	Volkswagen	Automotive	13.5	medium high-tech
2	Samsung	Computing and electronics	13.4	high & medium high-tech
3	Intel	Computing and electronics	10.6	high & medium high-tech
4	Microsoft	Software and internet	10.4	high-tech
5	Roche	Health care	10	high & medium high-tech
6	Novartis	Health care	9.9	high & medium high-tech
7	Toyota	Automotive	9.1	medium high-tech
8	Johnson & Johnson	Health care	8.2	high & medium high-tech
9	Google	Software and internet	8	high-tech
10	Merck	Health care	7.5	high & medium high-tech
11	GM	Automotive	7.2	medium high-tech
12	Daimler	Automotive	7	medium high-tech
13	Pfizer	Health care	6.7	high & medium high-tech
14	Amazon	Software and internet	6.6	high-tech
15	Ford	Automotive	6.4	medium high-tech
16	Sanofi-Aventis	Health care	6.3	high & medium high-tech
17	Honda	Automotive	6.3	medium high-tech
18	IBM	Computing and electronics	6.2	high & medium high-tech
19	GlaxoSmithKline	Health care	6.1	high & medium high-tech
20	Cisco	Computing and electronics	5.9	high & medium high-tech

Source: EU Industrial R&D Investment Scoreboard (2015). Notes: *Billion USD.

Table 2: Distribution characteristics of R&D intensity

	Quantiles					
	min	25%	50%	75%	max	Obs.
low-tech	0.003	0.041	0.102	0.280	4.306	133
medium-low-tech	0.012	0.093	0.255	0.462	6.804	79
medium-high-tech	0.004	0.428	0.773	1.412	9.457	478
high-tech	0.045	1.679	3.707	7.448	126.380	483

Notes:

RDCAPEX: R&D intensity is defined as a share of R&D expenditure in capital expenditure in 2006.

Table 3: Regression results

	GPS: first-step regression, Equation (2)			
	Coef.	S.E.	T-stat.	P-val.
Incpt	4.305	0.74	5.79	0.000
$\ln \mathrm{OP}^{+}_{2006}$	-0.081	0.05	-1.75	0.081
$\ln \mathrm{OP}^{-}_{2006}$	-0.036	0.05	-0.74	0.462
$\ln \mathrm{MCAP}_{2006}$	-0.433	0.18	-2.42	0.016
$[\ln \mathrm{MCAP}_{2006}]^2$	0.040	0.01	3.20	0.001
$\ln \mathrm{SALE}_{2006}$	-0.314	0.08	-3.81	0.000
$[\ln \mathrm{SALE}_{2006}]^2$	-0.001	0.01	-0.12	0.905
AT	-1.713	0.58	-2.97	0.003
BRIC	-1.471	0.64	-2.31	0.021
EU	-0.162	0.44	-0.37	0.711
Japan	0.014	0.45	0.03	0.975
RoW	-0.207	0.50	-0.41	0.680
USA	0.044	0.43	0.10	0.919
THE	0.184	0.21	0.86	0.391
SCS	0.525	0.23	2.32	0.021
PBT	0.199	0.23	0.87	0.383
HCE	-0.411	0.25	-1.66	0.098
R^2	0.42			
Obs.	483			

Notes:

The dependent variable r_{it} in the first-step regression is the log of R&D intensity in 2006, defined as the share of R&D expenditure in capital expenditure in the same year. The regression contains regional (AT, BRIC, EU, Japan, RoW, USA) and industry (THE, SCS, PBT, HCE) dummies, see Section 4.

Table 4: Initial balancing properties of covariates

	Group 1	Group 2	Group 3
$\ln \mathrm{OP}^{+}_{2006}$	9.43	0.91	-10.26
$\ln \mathrm{OP}^{-}_{2006}$	5.16	0.31	-4.73
$\ln \mathrm{MCAP}_{2006}$	7.58	1.07	-8.63
$[\ln \mathrm{MCAP}_{2006}]^2$	6.80	1.10	-8.55
$\ln \mathrm{SALE}_{2006}$	9.94	1.23	-10.03
$[\ln \mathrm{SALE}_{2006}]^2$	8.77	0.77	-10.53
AT	1.99	-2.47	-1.00
BRIC	1.39	-0.38	-2.01
EU	-2.16	-0.93	3.00
Japan	3.06	-0.23	-3.95
RoW	0.91	0.91	-2.53
USA	-1.23	0.90	0.32
THE	0.67	1.26	-1.99
SCS	-4.77	-0.84	4.76
PBT	-1.82	1.08	0.65
HCE	4.49	-1.81	-4.59
Obs.	161	161	161

Notes:

Groups of equal size were created using distribution of the continuous treatment variable, R&D intensity. Table entries are t-values of the test for the equal means between observations belonging to a particular group and those observations that do not belong to this group.

Table 5: Cell size for testing the balancing property of GPS

	Group 1	Control 1	Group 2	Control 2	Group 3	Control 3
Block 1	31	185	31	93	28	212
Block 2	30	50	31	52	27	39
Block 3	30	24	30	60	27	18
Block 4	30	19	31	36	27	25
Block 5	31	12	31	47	27	12
Total	152	290	154	288	136	306

Notes:

The block size of each treatment group is held approximately the same. For each group it is determined by quintiles of the estimated GPS.

Table 6: GPS-adjusted balancing properties of covariates

	Group 1	Group 2	Group 3
$\ln \mathrm{OP}^{+}_{2006}$	1.28	0.47	-1.08
$\ln \mathrm{OP}^{-}_{2006}$	1.05	-0.59	0.60
$\ln \mathrm{MCAP}_{2006}$	1.28	0.48	-1.06
$[\ln \mathrm{MCAP}_{2006}]^2$	1.01	0.72	-1.16
$\ln \mathrm{SALE}_{2006}$	1.36	0.43	-1.62
$[\ln \mathrm{SALE}_{2006}]^2$	0.98	0.44	-1.60
AT	-1.00	-1.00	1.00
BRIC	1.41	-1.42	-1.42
EU	-0.74	-0.04	0.06
Japan	0.28	-0.24	-0.10
RoW	0.39	0.73	-2.84
USA	0.51	-0.20	0.52
THE	-0.15	0.31	1.52
SCS	-0.69	-1.33	0.55
PBT	-0.12	1.97	-0.66
HCE	1.16	-0.68	-2.84
Obs.	152	154	136

Notes:

Table entries are t-values of the test for the equal means between observations belonging to a particular group and those observations that do not belong to this group, accounting for GPS.

Table 7: Regression results

	GPS: second-step regression, Equation (3)			
	Coef.	S.E.	T-stat.	P-val.
Incpt	7.242	0.289	25.088	0.000
$\ln r$	0.488	0.171	2.862	0.004
$[\ln r]^2$	0.430	0.097	4.425	0.000
$[\ln r]^3$	-0.145	0.025	-5.908	0.000
s	6.708	0.946	7.095	0.000
$\ln r * s$	-5.485	0.593	-9.246	0.000
σ^2	1.678			
R^2	0.40			
Obs.	442			
SIC	259.36			

Table 8: GPS: Second-stage regression — Robustness check

PANEL	(A)			(B)			(C)				(D)		
ICB3	High-tech			(High+Med-High)-tech			All				High-tech		
	Coef.	S.E.	P-val.	Coef.	S.E.	P-val.	Coef.	S.E.	P-val.		Coef.	S.E.	P-val.
Incpt	6.675	0.506	0.000	8.805	0.376	0.000	9.286	0.354	0.000	Incpt	9.502	0.546	0.000
$\ln r$	0.430	0.182	0.019	0.544	0.181	0.003	0.634	0.170	0.000	$\ln r$	-1.943	0.253	0.000
$[\ln r]^2$	0.467	0.100	0.000	-0.031	0.039	0.432	-0.097	0.032	0.002	$[\ln r]^2$	0.479	0.136	0.000
$[\ln r]^3$	-0.149	0.025	0.000	-0.081	0.022	0.000	-0.092	0.021	0.000	$[\ln r]^3$	-0.132	0.031	0.000
s	17.562	7.142	0.014	2.497	5.282	0.637	-1.439	5.231	0.783	$\ln s$	-0.418	0.767	0.586
s^2	-50.613	32.563	0.121	-5.032	23.116	0.828	10.484	23.951	0.662	$[\ln s]^2$	-0.554	0.301	0.066
s^3	67.447	44.330	0.129	4.802	29.761	0.872	-15.943	32.191	0.621	$[\ln s]^3$	-0.068	0.033	0.038
$\ln r * s$	-5.510	0.595	0.000	-2.858	0.446	0.000	-2.822	0.429	0.000	$\ln r * \ln s$	-0.517	0.079	0.000
σ^2	1.677			1.738			1.689				1.829		
R^2	0.402			0.205			0.153				0.348		
Obs.	442			771			818				442		
SIC	269.080			471.390			474.400				307.450		