Do performance-based research funding systems affect research production and impact?

By Daniele Checchi, Marco Malgarini, Scipione Sarlo (Anvur) – present version 05/10/17

Abstract

1. Introduction

Since the 1980s, a widespread wave of reforms has led to the redesign of the main public administration mechanisms at all levels and in all sectors (Kettl, 2005). Drivers of the reforms process have been the increase perception of government inefficiency (Olson, 1973), a more market oriented view of public management, a growing orientation to services, the progressive decentralisation of power, the separation of the role of government in purchasing and providing services and the increased importance attributed to government accountability. The research and education sector has also been heavily involved in the reform process; in fact, research and innovation are growingly considered as the main driver to produce knowledge useful to achieve economic and social goals (Whitley, 2007). According to the new paradigm, the adoption of a Performance based Research Funding Systems (PRFS) is one of the main instruments to steer and manage the higher education system (Herbst, 2007); while not immune from some risks (see for instance Hicks, 2012), PRFS are considered to stimulate research organizations improving their production in terms of quality or quantity, as well as to steer certain fields or topics of research (Geuna, A., Martin, B.R., 2003, Hicks, D., 2012).

However, even if PRFS are often advocated as an efficient mechanism to govern the higher education system, existing literature has not reached a consensus on its effects on performances (see for instance Adams and Gurney, 2014, or Wang and Hicks, 2003, among others). The aim of our paper is to contribute to this kind of literature providing an analysis of the effects of PRFS on research production and impact, considering a fairly large number of countries that have adopted such a system at some point over the last twenty years; other countries that are relevant in the international research landscape are also considered in our analysis as controls. Our idea is to evaluate whether the introduction of PRFS has a statistically significant impact on scientific productivity and impact, once controlling for other possible variables influencing the outcome. The rest of the paper is structured as follows: section 2 provides a definition of PRFS in the context of the more general wave of reforms that has radically changed public administrations since the 1980s'; section 3 presents the dataset that will be used in the econometric analysis performed in section 4. Section 5 concludes.

2. Performance based funding and the wave of public administration reforms

Reforms started in the 1980s' have led to the redesign of the main public administration mechanisms at all levels and in all sectors (Kettl, 2005). Six main drivers are considered to have inspired reforms: first, the perception of a growing inefficiency of governments (Olson, 1973), with a corresponding desire for more effective and cheaper services; secondly, a more market-oriented view of public management, stimulating the replacement of traditional command-and-control systems with market incentives; thirdly, growing orientation to service, implying putting citizens first instead of service providers, as it was common in the previous bureaucratic structure (Kettl, 2007, Herbst, 2007, Hicks, 2010). Moreover, in order for programmes and policies to become more responsive, the fourth strategy consists in moving toward the decentralization of power, with the gradual disaggregation of public administration into a multitude of semi-autonomous or para-governmental organizations (Hood, C., Schuppert, O., 1988), known as Agencies (so called *agentification*, see Verhoest, K., Thiel, S. van, Bouckaert, G. and Laegreid, P., 2012; Capano, G., Turri, M., 2017). The fifth driver was the idea of promoting a growing separation of the governments' role in purchasing and providing services. Finally, the emergence of a growing demand for governments' accountability, and a shift of focus from processes and structures to outputs and outcomes (Kettl, 2007, Herbst, 2007, Hicks, 2010): in this context, accountability may provide incentives for improving public organizations through the reallocation of public funds based on output and outcome measures; at the same time, it may become a powerful mean to encourage public organizations to act responsibly (Perrin, B., 2007, Perrin, B., 2015, Stame, N., 2016).

Understanding the challenges and changes characterizing public government scenario is crucial also for realising what has happened in the Research and Education sector. In fact, reforms have profoundly marked and revolutionized Research and Education environment and its key actors: Higher Education Institutions (mainly Universities), and Public Research Organizations (PRO) (Cruz-Castro, L., Bleda, M., Derrick, G.E., Jonkers, K., Martinez, C., and Sanz-Menendez, L., 2011). In the information society, the prevailing political view perceives research as the main driver to produce knowledge useful to achieve economic and social goals (Whitley, R., 2007). The connection between scientific research and public purposes becomes even more complex and implies composite and active tools in steering research activities and reallocate public funds, without violating the independence of scientists and their organisations (Braun, D., 2003). For these reasons, correctly designing the funding mechanism used to finance Universities and PRO is considered crucial in order to enhance research performances. In this context, starting from 1980s', a considerable number of countries have adopted funding mechanisms based on performance, the so-called Performance Based research Funding Systems (PRFS). According to the definition formulated by Hicks (2012), and used in Jonkers, K., Zacharewicz, T (2016), the main characteristics of a PRFS are the following:

- Research must be assessed.
- Research evaluation must be *ex post*.
- Research output and/or impact must be evaluated.
- Part of the governmental allocation of university research funding must depend on the outcome of the evaluation.
- The PRFS must be a national or regional system (Hicks, D. 2012, Jonkers, K., Zacharewicz, T., 2016).

Assessment exercises concerning degree programmes and teaching are excluded. Also evaluations of projects' proposals are not included, because they are mainly performed *ex ante*, and not directly involved with the outcomes of research activity derived from those proposals. For the same rationale, funding systems that reallocate funds only on the number of PhD students are not considered. Moreover, a PRFS must contribute directly or indirectly to assign research funds: any kind of evaluation exercises that provide only recommendation or feedback to HEIs and PROs is not considered, as well as evaluation or assessment exercises that are performed at local or institutional level (Hicks, D. 2012, Jonkers, K., Zacharewicz, T., 2016).

The greater leverage that PRFS offer to governments and management of HEIs and PROs is to encourage such organizations to improve their production in terms of quality or quantity, as well as highlighting and steering certain fields or topics of research (Geuna, A., Martin, B.R., 2003, Hicks, D., 2012; Tapper, T., Salter, B., 2003). Further expected benefits include providing incentives for best performing Institutions, in a competitive game which reward outputs (Herbst, 2007), while also increasing accountability of HEIs and other PROs (Hicks, D., 2012). Possible drawbacks in the use of such systems include the risk of creating perverse incentives for bad research practices (ranging from the multiplication of irrelevant publication to plagiarism, self-plagiarism and scientific fraud, see on this Hazelkon, 2010), that of stimulating the so-called Matthew' effect and that of discouraging interdisciplinary and innovative research (see on this Hicks 2012; Rafols et al, 2012; Wilsdon et al, 2016). Costs of implementation of PRFS are also non negligible (e.g. Boer et al, 2015; Hicks, 2012; Martin 2011; OECD 2010;), though Geuna and Piolatto (2015) indicate that at least in the short term the benefits outweight the costs in the case of the UK and Italy (Hicks, D. 2012, Jonkers, K., Zacharewicz, T., 2016).

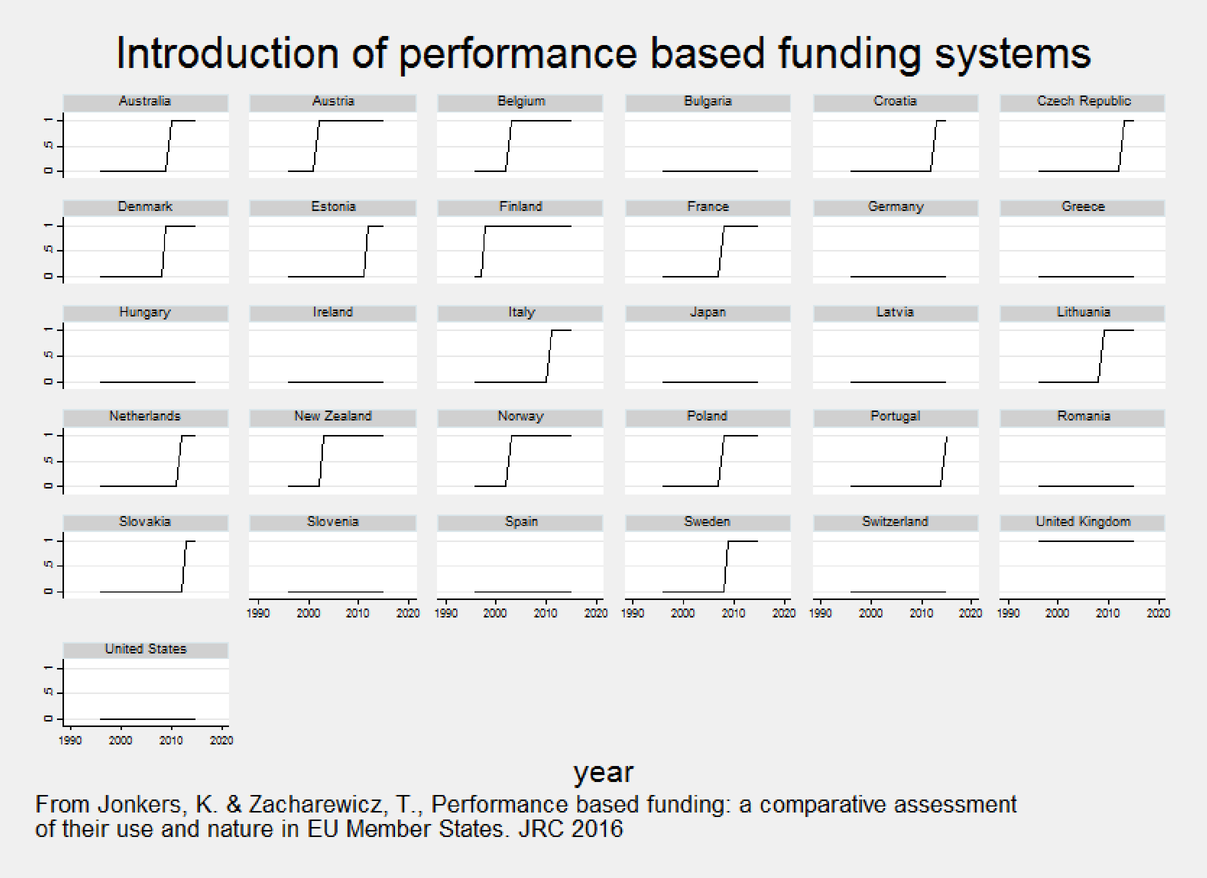
In the following, we will concentrate on the effect of PRFS on the system performance in terms of research productivity and impact; in fact, existing literature has not reached a consensus on the effects of PRFS on the system' performance. In the next section, we will present a first overview concerning the timing of the adoption of a PRFS in a fairly large number of European and non-European countries; we will then introduce the variable that we will use as a proxy for evaluating the size and impact of scientific production and the control variables to be used in order to properly assess the effects of PRFS on production and performance.

3. Data

3.1 Performance Based Funding Systems dataset

Jonkers, and Zacharewicz (2016) provide a general overview of the different assessment and allocation systems in use among the EU Member States and in some of the main extra-European countries. Among the latter, in the United States the system is mainly funded at the institutional level and is not usually considered as a PRFS according to the Hicks definition, while Australia and New Zealand have experienced several evaluation exercises in the last twenty years and are hence considered as proper PRFS systems. (Butler, L., 2003a, 2003b, Jonkers, K., Zacharewicz, T., 2016). The classification takes into account two kind of information: the presence of some type of PRFS, according to the Hicks' definition previously illustrated, and the year in which this kind of system has been introduced in each country. Figure 1 shows the situation for the 31 countries considered in our analysis[1] for the period 1996-2015. The figure shows that according to the available information 17 out of 31 countries have introduced some form of PRFS over time; some of the major player in the research scenario (including the US, Japan, Spain and Germany) are not considered to use a proper form of PRFS, together with some eastern European countries. On the other hand, the UK is the only country in which such a system was in place already at the beginning of the period considered in our analysis.

Figure 1: Presence or absence of a PBFS across countries and years considered in the analysis



3.2 Indicators of scientific production and impact

Our indicators of scientific production and impact are extracted from the Scival database and are referred to the period 1996-2016. We consider 31 countries worldwide, and distinguish

[1] The countries considered in the analysis are: Australia; Austria; Belgium; Bulgaria; Croatia; Czech Republic; Denmark; Estonia; Finland; France; Germany; Greece; Hungary; Ireland; Italy; Japan; Latvia; Lithuania; Netherlands; New Zealand; Norway; Poland; Portugal; Romania; Slovakia; Slovenia; Spain; Sweden; Switzerland; United Kingdom; United States.

among six main scientific categories, following the Field of Science and Technology (FOS) classification (see table 1)[2].

[2] See http://www.oecd.org/sti/inno/38235147.pdf and OECD (2015)

Table 1: List of scientific areas.

Scientific Category	Sub-categories included
Agricultural sciences	Agriculture, forestry and fisheries
	Animal and dairy science
	Veterinary science
	Agricultural biotechnology
	Other agricultural sciences
	Agriculture, forestry and fisheries
	Animal and dairy science
	Veterinary science
Engineering and Technologies	Civil engineering Electrical engineering, electronic engineering, information engineering Mechanical engineering Chemical engineering Materials engineering Medical engineering Environmental engineering Environmental biotechnology Industrial biotechnology Nano-technology Other engineering and technologies
Humanities	History and archaeology Languages and literature Philosophy, ethics and religion Art (arts, history of arts, performing arts, music) Other humanities
Medical sciences	Basic medical research Clinical medicine Health sciences Health biotechnology Other medical science
Natural sciences	Mathematics Computer and information sciences Physical sciences and astronomy Chemical sciences Earth and related environmental sciences Biological sciences Other natural sciences
Social sciences	Psychology Economics and business Educational sciences Sociology Law Political science Social and economic geography Media and communication Other social sciences

Source: OECD (2015)

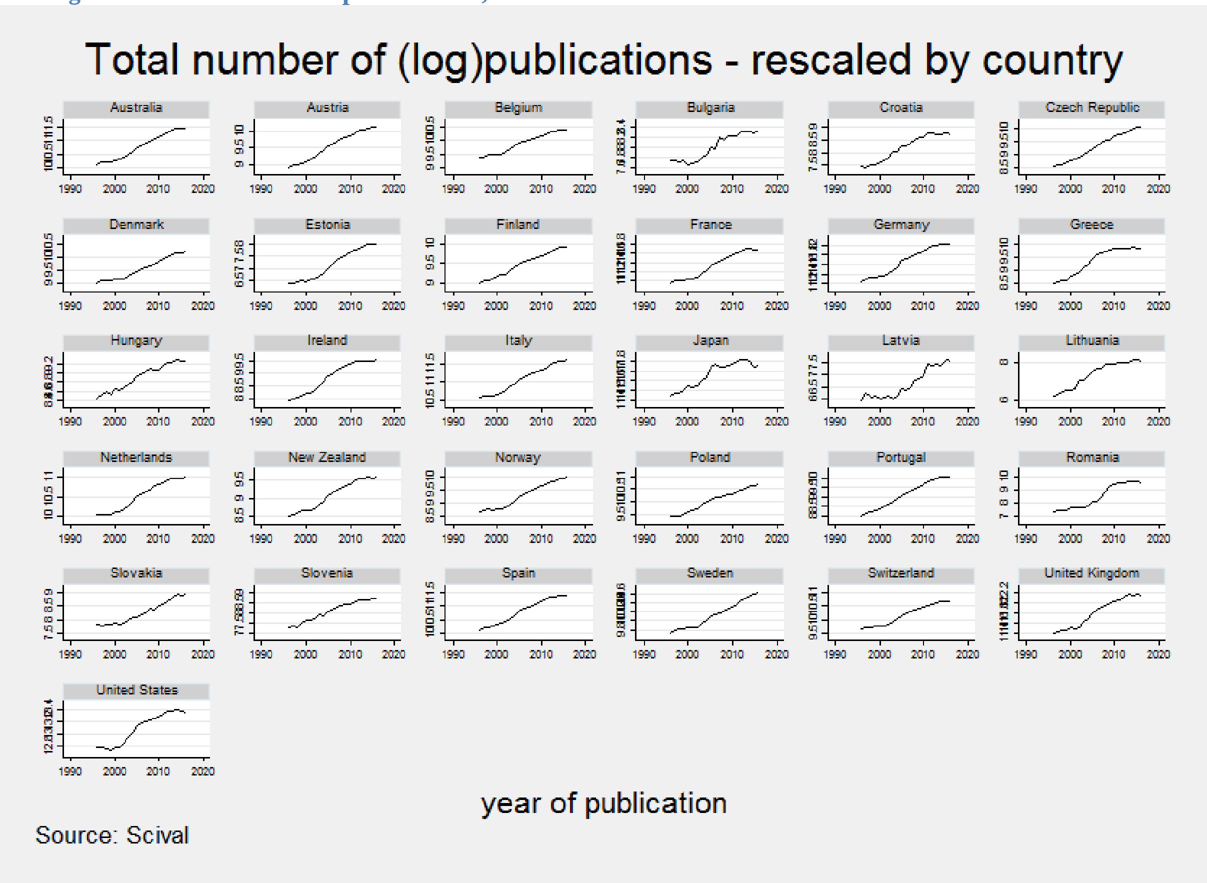
For each country, scientific category and year we consider an indicator of scientific production (the number of publications), one indicators of scientific impact (the Field Weighted Citation Impact, FWCI [3]) and one indicator of excellence of the scientific production (the share of paper in top 10% journals in terms of citations[4]). Figure 2 presents the trend of total publications in the period; the data show a marked increase in the number of scientific output indexed in Scopus, a common finding in this kind of literature, mostly attributable to the enlargement of the number of journals covered by this database. In the

[3] Defined as the ratio of citations received relative to the expected world average for the subject field, publication type and publication year. We also analysed the number of citations per paper, but the variable has a clear hump shaped dynamics over the sample period, which complicates the modelling of the dynamic path.
[4] Defined as the number of publications of a selected entity that have been published in the world top 10% journals, ranked according to the CiteScore percentiles.

same period, European countries also increased their scientific impact as measured by the FWCI; notably, Italy and Spain, that in 1996 were lagging behind in terms of scientific impact with respect to the other main industrial countries, are now well above the OECD averages for this indicator (see Figure 3).

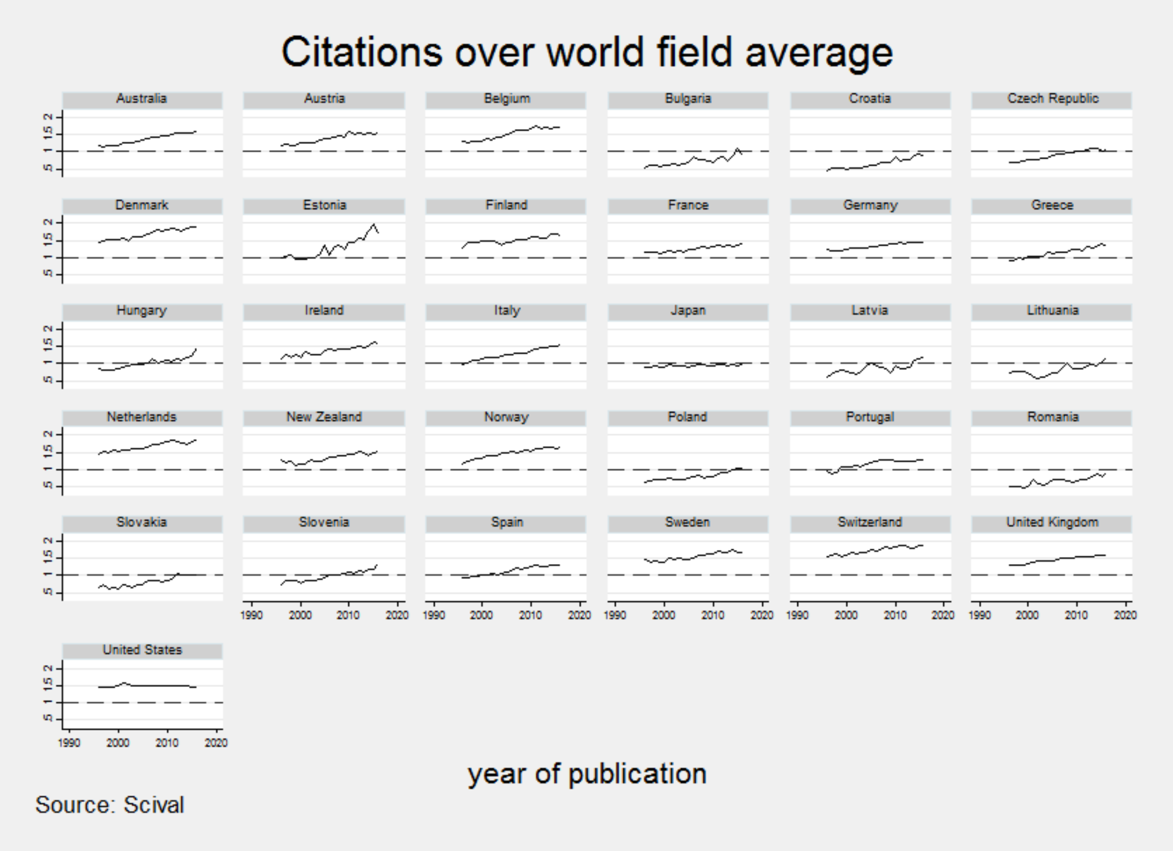
Finally, figure 4 reports, for the same countries considered above, the share of output comprised in the top 10% of the world publication in terms of number of citations. All European countries outperform with respect to the world average: the share of papers comprised in the top 10% worldwide was equal in 2016 to 23.4% in the UK, to over 20% in France, Germany and UK, being equal however to around 18% also in Spain and Italy; those values are close to the ones reached for the average of the OECD countries in the same period.

Figure 2: Total number of publications, 1996-2015



Source: Scival

Figure 3: Field Weighted Citation Impact, 1996-2015



Source: Scival

Figure 4: Share of papers in the top 10% in the world in terms of number of citations, 1996-2015

Source: Scival

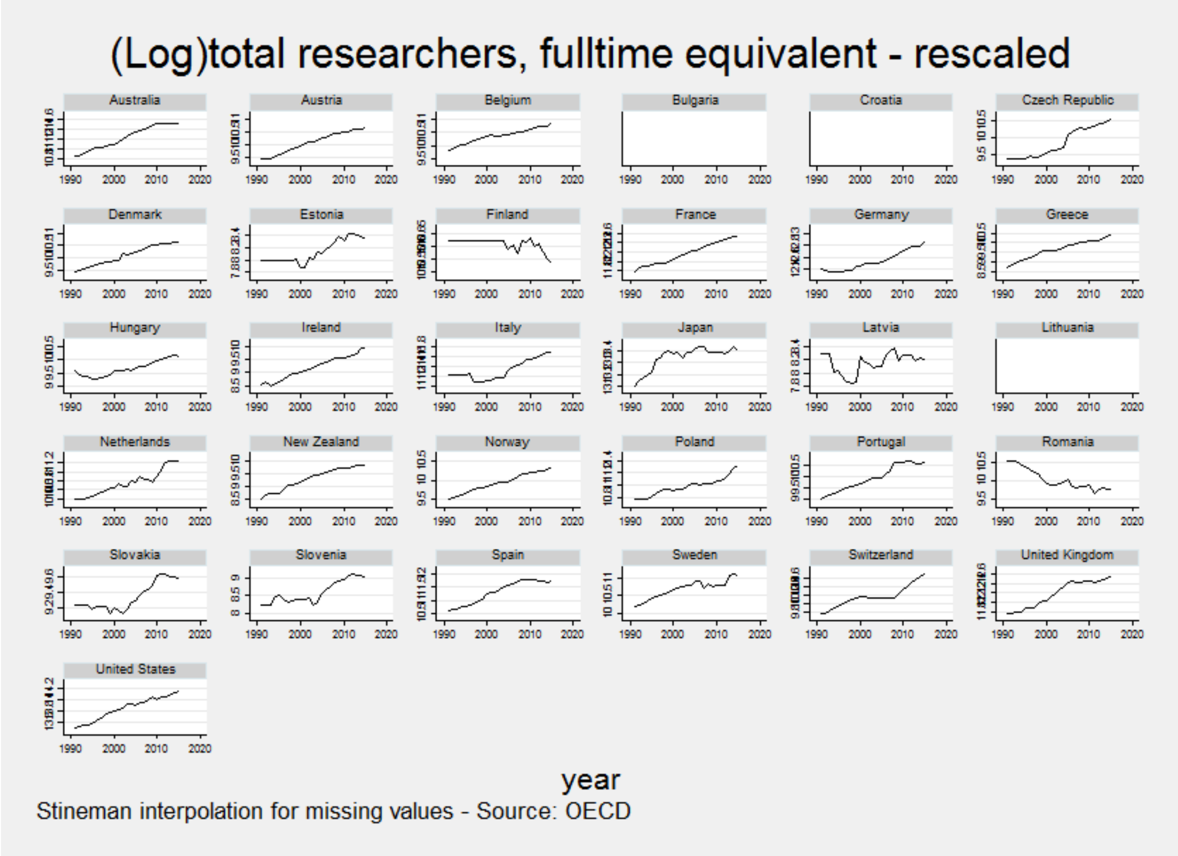
3.3 Science and technology indicators

In order to better assess the impact of the introduction of PRFS on scientific performance, in the following we will also control for the possible effects on it of the level and structure of the efforts undertaken by the countries in the field of science and technology. Related indicators are extracted from the OECD Main Science and Technology Indicators (MSTI) database[5]; more specifically, we choose to concentrate our attention on the indicators concerning the number of total researchers (in terms of full time equivalent) and the share of total gross R&D expenditures with respect to GDP. Figure 5 shows the log of total researchers (information is missing for non OECD countries): in the period considered, France, Germany and the UK show a similar upward tendency, especially when expressed as share with respect to total labour force. The published data are rather incomplete, and we had to resort to data interpolation in order not to lose relevant information. Figure 6 shows the dynamic over time of the share of gross domestic expenditures on R&D with respect to total GDP: in this case, Germany stands out, with a higher share that continues to grow even in the more recent period, reaching almost 3% in terms of the country' GDP, while in Italy and Spain the share is just slightly over 1%.

Eventually we searched for some control on the supply side of potential researchers, and we ended to the Unesco database (http://data.uis.unesco.org/), which however covers a shorter period of years (compared to the other series) and exhibits unusual fluctuations (see Figure 7). We will make use of this variable simply as additional control, without putting much emphasis on it.

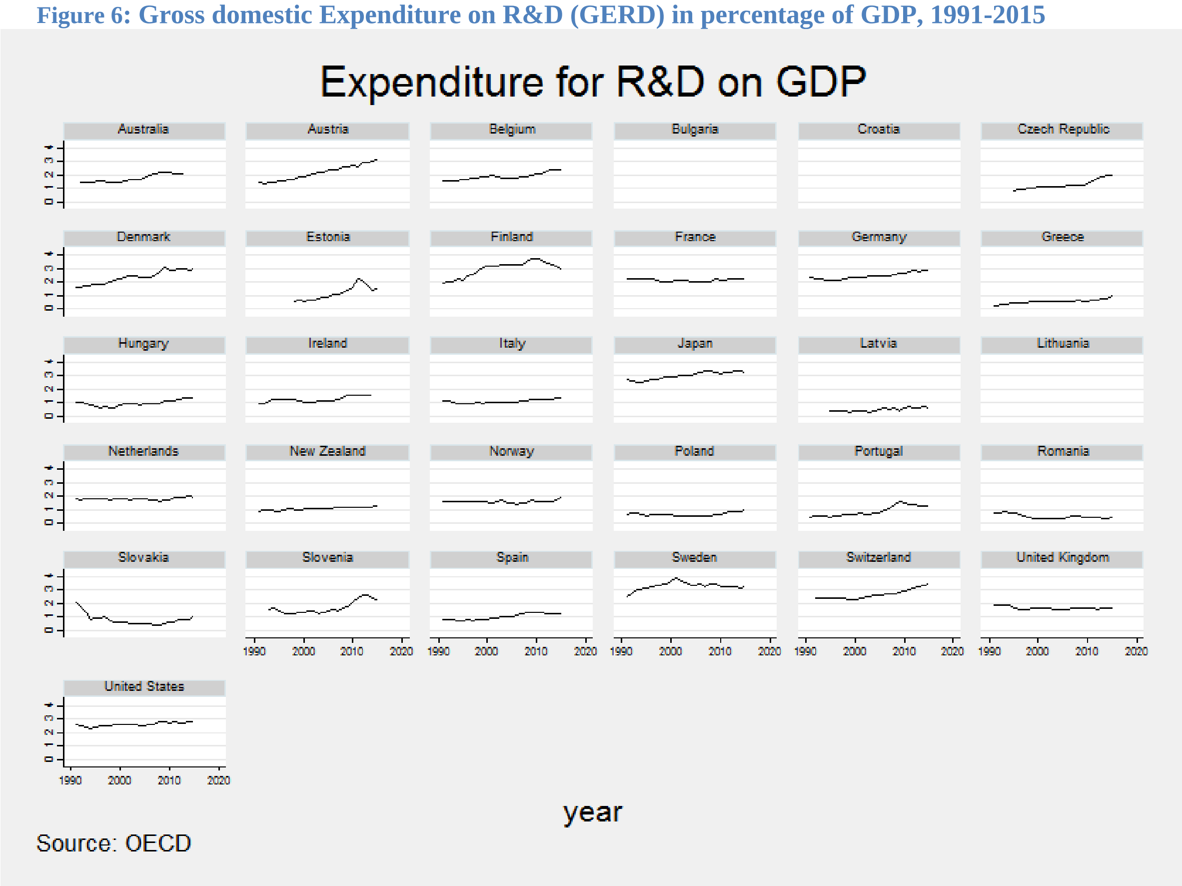
[5] http://stats.oecd.org/Index.aspx?DataSetCode=MSTI_PUB

Figure 5: Total researchers, 1991-2015



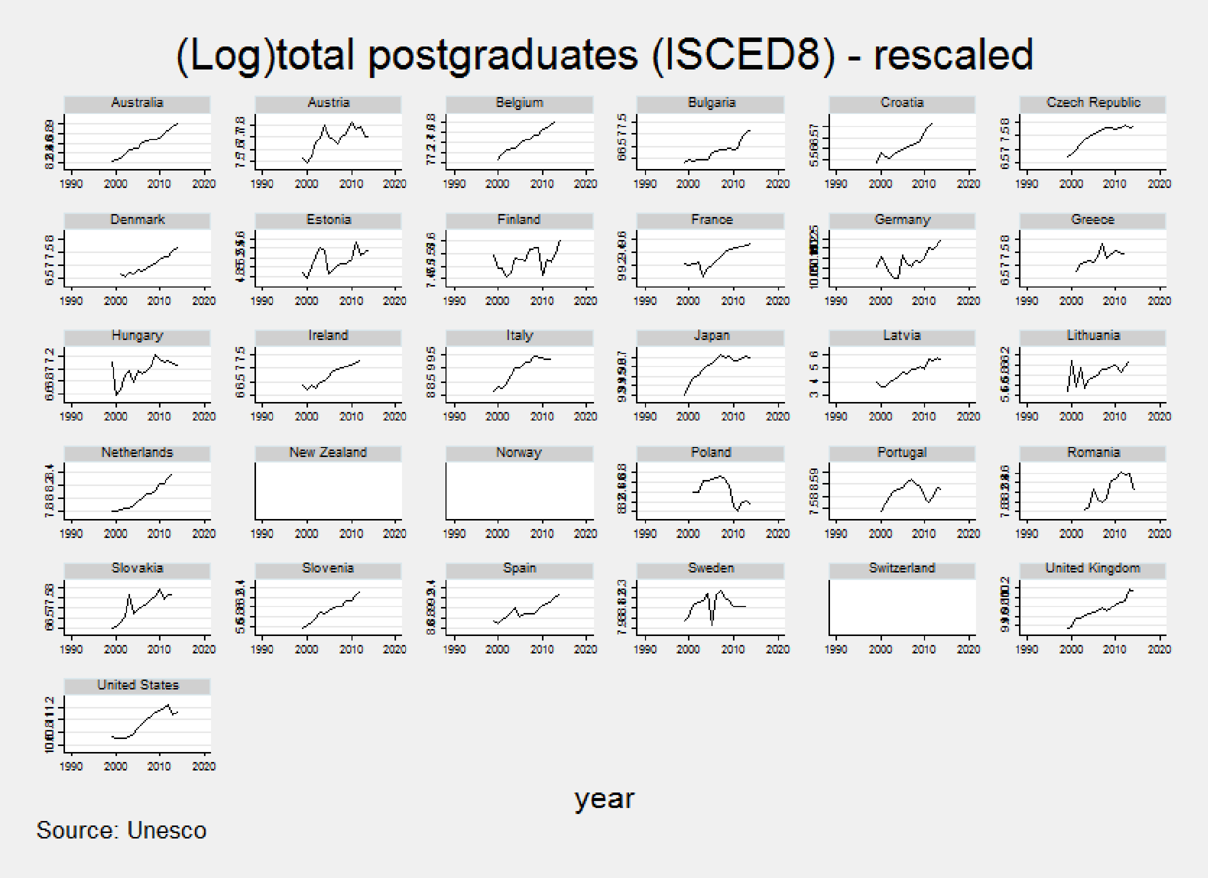
Source: OECD Main Science and Technology Indicators

Figure 6: Gross domestic Expenditure on R&D (GERD) in percentage of GDP, 1991-2015



Source: OECD Main Science and Technology Indicators

Figure 7: Log of total number of post-graduate degrees, 1999-2014



4. Results

We approach our question of interest (whether the introduction of PRFS has impact on the research activity of a country) following a diff-in-diff approach, since our sample includes 13 countries without any performance assessment systems, which will represent our control group. The remaining group of countries have introduced at different points of time various form of research assessment, either based on algorithms or on peer review. These countries receive their treatment at different point of the sample period, possibly with different intensity (since formula-based systems tend to bite more than peer-based ones). We implement our strategy by estimating the following model

$$y_{it} = \alpha_i + \beta x_{it} + \gamma \cdot \text{Reform}_{it} + \epsilon_{it} \qquad (1)$$

where y_{it} indicates the output variable of country i at time t, α_i are country fixed effects, x_{it} are additional controls (like number of researchers or PhDs, expenditure in R&D) and Reform_{it} is a step-dummy variable which takes unitary value when a PRFS is introduced.

However this model is inadequate for our purposes since our dependent variable is trended (like in the case of the total number of publications) or is anyhow affected by the repeated enlargement of the dataset due to the inclusion of new journals. We deal with this problem either introducing a time trend, which can either be general

$$y_{it} = \alpha_i + \delta \cdot t + \beta \cdot x_{it} + \gamma_0 \cdot \text{Reform}_{it} + \gamma_1 \text{Reform}_{it} \cdot t + \epsilon_{it} \quad \text{(2a)}$$

or country specific as in

$$y_{it} = \alpha_i + \delta_i \cdot t + \beta \cdot x_{it} + \gamma_0 \cdot \text{Reform}_{it} + \gamma_1 \text{Reform}_{it} \cdot t + \epsilon_{it} \quad \text{(2b)}$$

Alternatively we follow a non-parametric approach and we use year dummies as in

$$y_{it} = \alpha_i + \delta_i + \beta \cdot x_{it} + \gamma_0 \cdot \text{Reform}_{it} + \gamma_1 \text{Reform}_{it} \cdot t + \epsilon_{it} \quad \text{(2c)}$$

Using various specifications of the model represented by equation (2) we estimate a possible average impact of the reform on the intercept (coefficient γ_0) and on the slope (coefficient γ_1), but we ignore the time profile of this impact. In order to investigate this aspect, we have constructed dummies variables corresponding to the year distance from the reform. In such a case we estimate

$$y_{it} = \alpha_i + \delta \cdot t + \beta \cdot x_{it} + \sum_{\tau=0}^{10} \gamma_\tau \cdot \text{Ref.year}_{i\tau} + \epsilon_{it} \quad \text{(3)}$$

where $\text{Ref.year}_{i\tau}$ is a dummy variable taking a unitary value after τ years from the reform. All models have been estimating taking into account the potential clustering of the errors by countries.

Going to comment our results, in table 2 we present the estimates of models (2.a)- (2.b)- (2.c) in column 1, 2 and 3 respectively, while columns 4, 5 and 6 presents the estimation of model (3) under different assumption regarding the time trend. Looking at first column, the estimate suggests an average positive impact on the intercept and a negative impact on the slope: the number of publications would increase by 17% on average, but this increase fades away in 13 years. This order of magnitude is confirmed in column 3, though at a lower level of statistical significance. When we look at the time profile of this impact in column 4, we notice that most of the effect occurs in the initial 3 years, when publications grow by almost 15% in comparison to the control cases, but this effect tends to vanish later on. While the magnitude of the coefficients is consistent across columns of table 2, their statistical significance relies on the assumption we make on the error distribution. Should we remove the clustering assumption, the impact of the introduction of performance base funding system would be statistically significant in all specifications for the initial 3-4 years, and then disappears.

If quantity seems to react, let’s now consider what happens to the quality of scientific production, which is captured by two outcomes, the share of papers published in top journals

(10%)[6] and the average citations obtained by each paper (net of differences attributable to field or year of publication). In table 3 we show the results for the share of papers in top journals: the average effect is nil, though one can identify some positive effect after 2-4 years from the introduction of the reform. The magnitude is not impressive, in the order of 1-2 percentage points, which reduce with the time passing. The limited impact on a share is not surprising, since at world level the sum of the changes must be zero by construction. On the contrary, the second measure of quality seems more reactive to the introduction of performance based system. Looking at table 4 we observe a positive and statistically significant effect on the number of citations obtained by each paper, in the order of 5 percentage points, which are accumulated by the third year of the reform, reaching a peak during the fourth year and then almost disappearing by the eighth year. All these effects are statistically significant at usual confidence intervals.

Once we disaggregate the outcome variables by the six research fields indicated in table 1, we find that most of the previously described dynamics are driven by specific research area. Looking at table 5, we notice that most of the impact on the total number of publications derives from Agricultural Sciences, Humanities and Medical Sciences, while Engineering and Natural Sciences seem unaffected by the introduction of research assessment. On the contrary, these two research areas seem more responsive to quality, for the limited significant correlations between reform and share of papers in top 10% journals and/or citations emerges for these two areas (see tables 6 and 7). The research area in Social Sciences seems rather unaffected by research assessments.

Overall we may summarise the results obtained so far by saying that RPFSs have a short run impact on quantities (number of papers) that reveals however temporary, since it vanishes in 6-7 years. Vice versa, the impact on quality of scientific research is more permanent, since it peaks after 3-4 years and retains its effects for at least a decade. This dynamics is not equally distributed among research areas: those areas that are less more exposed to an international audience (like Engineering or Natural Sciences) drive the "quality", while remaining ones are mostly responsible for a "quantity" one.

As robustness checks we have also introduced some measures of resources invested in research activity, like expenditure in Research and Development and the total number of researchers or the number of PhDs. However this check reduces the sample by one fifth, since

[6] We have experimented with alternative outcomes (like share of papers in top 1% or in top 5% journals), but we were unable to identify robust effects of the reform.

information on these variable is absent in 6 countries.[7] In addition, in order not to reduce sample size even further, we have interpolated missing data[8] and used initial and terminal values to extend the sample size. As a consequence, more than robustness checks the results are to be taken as investigation on the correlation between research outcomes and research inputs. In table 8 in the Appendix we find that the total number of publications is positively correlated with R&D expenditure and number of PhDs, while the number of researchers does not exhibit any correlation. On the contrary the number of researchers exhibits some positive correlation with the share of papers in top journals in table 9 in the Appendix, while the other two input measures do not show a consistent pattern of association. Finally, in table 10 in the Appendix the three input measures exhibit positive correlation with number of citations, especially in the case of PhDs.

A final aspect that we have investigated is whether some results are due to the specific form of the research assessment. We have therefore classified the countries introducing RPFS into two groups: one group uses bibliometric performance indicators, which allows a continuous monitoring of scientific productivity, both in terms of quantity and quality, thanks to the existence of international datasets (like Scopus and WoS). On the contrary, the second group mostly relies on periodical assessment, based on peer reviews. Not surprisingly, the former system puts a stronger hold on researchers than the latter, and this is reflected in the estimation. Looking at table 11 in the Appendix we observe that most of the effect is driven by RPFS of the first type, especially regarding the number of publications.

[7] Bulgaria, Croatia and Lithuania do not possess information on expenditure and researchers, while New Zealand, Norway and Switzerland are missing in the number of postgraduate degrees.

[8] We have followed the Stineman procedure in R.

Table 2 - Total publications (logs) - all fields - 1996-2015

	(1) com.trend	(2) dif.trend	(3) yr.dummy	(4) com.trend	(5) dif.trend	(6) yr.dummy
reform	0.175*	0.057	0.170+	-0.140	-0.131	-0.131
	(0.074)	(0.037)	(0.096)	(0.123)	(0.142)	(0.142)
reform×time	-0.013+	-0.005	-0.013			
	(0.007)	(0.004)	(0.010)			
refyear1				0.165	0.152	0.152
				(0.098)	(0.109)	(0.109)
refyear2				0.148+	0.138	0.138
				(0.085)	(0.098)	(0.098)
refyear3				0.144+	0.137	0.137
				(0.079)	(0.094)	(0.094)
refyear4				0.127+	0.116	0.116
				(0.074)	(0.094)	(0.094)
refyear5				0.108	0.094	0.094
				(0.064)	(0.084)	(0.084)
refyear6				0.115+	0.106	0.106
				(0.061)	(0.080)	(0.080)
refyear7				0.090	0.096	0.096
				(0.059)	(0.075)	(0.075)
refyear8				0.070	0.057	0.057
				(0.062)	(0.068)	(0.068)
refyear9				0.119**	0.069	0.069
				(0.038)	(0.055)	(0.055)
refyear10				0.111***	0.060	0.060
				(0.029)	(0.044)	(0.044)
N	620	620	620	620	620	620
N.countries	31	31	31	31	31	31
R^2 adj	0.847	0.956	0.858	0.845	0.855	0.855

Standard errors in brackets - country fixed effects included - common time trend, country-specific time trend or year fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001

Table 3 - Share of articles in top10% journals - all fields - 1996-2015

	(1) com.trend	(2) dif.trend	(3) yr.dummy	(4) no trend	(5) com.trend	(6) yr.dummy
reform	0.009	0.005	0.006			
	(0.006)	(0.005)	(0.006)			
refyear1				0.008	-0.001	-0.000
				(0.005)	(0.005)	(0.004)
refyear2				0.014*	0.004	0.003
				(0.006)	(0.006)	(0.005)
refyear3				0.013*	0.001	-0.000
				(0.006)	(0.006)	(0.005)
refyear4				0.019**	0.007	0.006
				(0.007)	(0.007)	(0.004)
refyear5				0.018*	0.005	0.003
				(0.008)	(0.007)	(0.005)
refyear6				0.023*	0.009	0.006
				(0.009)	(0.009)	(0.006)
refyear7				0.019*	0.003	0.002
				(0.009)	(0.008)	(0.005)
refyear8				0.008	-0.007	-0.002
				(0.009)	(0.008)	(0.004)
refyear9				0.006	-0.008	0.004
				(0.006)	(0.005)	(0.005)
refyear10				0.016*	0.000	0.002
				(0.007)	(0.006)	(0.005)
N	620	620	620	620	620	620
N.countries	31	31	31	31	31	31
R^2 adj	0.285	0.527	0.495	0.060	0.272	0.486

Standard errors in brackets - country fixed effects included - common time trend, country-specific time trend or year fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001.

Table 4 - Citations per paper relative to field world average - all fields - 1996-2015

	(1) com.trend	(2) dif.trend	(3) yr.dummy	(4) no trend	(5) com.trend	(6) yr.dummy
reform	0.051*	0.031+	0.054*			
	(0.021)	(0.017)	(0.022)			
refyear1				0.129***	0.032+	0.040+
				(0.031)	(0.018)	(0.020)
refyear2				0.146***	0.035*	0.044*
				(0.033)	(0.017)	(0.017)
refyear3				0.183***	0.053+	0.052+
				(0.039)	(0.027)	(0.027)
refyear4				0.210***	0.075+	0.072+
				(0.052)	(0.038)	(0.040)
refyear5				0.173***	0.030*	0.028+
				(0.027)	(0.015)	(0.015)
refyear6				0.190***	0.035*	0.032*
				(0.030)	(0.015)	(0.014)
refyear7				0.187***	0.019	0.019
				(0.040)	(0.023)	(0.023)
refyear8				0.181***	0.020	0.014
				(0.044)	(0.025)	(0.026)
refyear9				0.233***	0.081*	0.079+
				(0.058)	(0.037)	(0.040)
refyear10				0.225***	0.054**	0.045+
				(0.042)	(0.019)	(0.024)
N	620	620	620	620	620	620
N.countries	31	31	31	31	31	31
R^2 adj	0.765	0.852	0.774	0.199	0.764	0.772

Standard errors in brackets - country fixed effects included - common time trend, country-specific time trend or year fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001.

Table 5 - Total publications (logs) - by fields - 1996-2015

	(1) agricultural	(2) engineering	(3) humanities	(4) medical	(5) natural	(6) social
reform	-0.392+	0.007	-0.219	-0.281	-0.077	-0.210
	(0.224)	(0.147)	(0.235)	(0.193)	(0.134)	(0.206)
refyear1	0.325*	0.092	0.347+	0.254	0.081	0.235
	(0.159)	(0.122)	(0.176)	(0.157)	(0.103)	(0.153)
refyear2	0.295+	0.102	0.267	0.233	0.071	0.199
	(0.145)	(0.116)	(0.160)	(0.142)	(0.090)	(0.140)
refyear3	0.302*	0.099	0.282+	0.226	0.079	0.226+
	(0.140)	(0.109)	(0.153)	(0.135)	(0.088)	(0.133)
refyear4	0.287*	0.084	0.218	0.243+	0.048	0.192
	(0.137)	(0.107)	(0.162)	(0.135)	(0.085)	(0.126)
refyear5	0.269+	0.077	0.184	0.214+	0.042	0.185
	(0.134)	(0.089)	(0.125)	(0.125)	(0.078)	(0.111)
refyear6	0.223+	0.079	0.230+	0.210+	0.056	0.186
	(0.130)	(0.087)	(0.132)	(0.123)	(0.073)	(0.115)
refyear7	0.181	0.080	0.242+	0.193	0.042	0.182+
	(0.123)	(0.083)	(0.142)	(0.118)	(0.063)	(0.104)
refyear8	0.152	0.035	0.133	0.130	0.026	0.091
	(0.106)	(0.079)	(0.146)	(0.101)	(0.059)	(0.095)
refyear9	0.121	0.083	0.054	0.110	0.041	0.061
	(0.094)	(0.059)	(0.076)	(0.076)	(0.050)	(0.065)
refyear10	0.081	0.063	0.090	0.112+	0.036	0.045
	(0.077)	(0.044)	(0.056)	(0.064)	(0.040)	(0.048)
N	620	620	619	620	620	620
N_countries	31	31	31	31	31	31
R^2 adj	0.730	0.836	0.880	0.769	0.853	0.873

Standard errors in brackets - country and time fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001

Table 6 - share of articles in top10% journals - by fields - 1996-2015

	(1) agricultural	(2) engineering	(3) humanities	(4) medical	(5) natural	(6) social
refyear1	-0.018	-0.017*	-0.023	0.004	0.001	0.004
	(0.014)	(0.008)	(0.014)	(0.005)	(0.005)	(0.008)
refyear2	-0.009	-0.014+	-0.001	0.008	0.005	0.003
	(0.016)	(0.008)	(0.022)	(0.006)	(0.004)	(0.008)
refyear3	-0.011	-0.019*	-0.012	0.007	0.002	-0.003
	(0.018)	(0.008)	(0.011)	(0.007)	(0.005)	(0.007)
refyear4	0.005	-0.007	0.003	0.011	0.005	-0.005
	(0.019)	(0.008)	(0.014)	(0.008)	(0.004)	(0.011)
refyear5	0.003	-0.009	-0.012	0.002	0.006	-0.005
	(0.024)	(0.011)	(0.018)	(0.006)	(0.005)	(0.012)
refyear6	-0.004	-0.001	-0.008	0.005	0.009	-0.007
	(0.022)	(0.012)	(0.016)	(0.006)	(0.005)	(0.012)
refyear7	-0.000	-0.009	-0.008	0.007	0.005	-0.008
	(0.018)	(0.009)	(0.017)	(0.005)	(0.004)	(0.012)
refyear8	-0.008	-0.011	-0.009	0.001	0.002	-0.001
	(0.013)	(0.009)	(0.020)	(0.006)	(0.005)	(0.009)
refyear9	0.010	0.005	0.013	-0.000	0.009+	-0.003
	(0.019)	(0.009)	(0.018)	(0.008)	(0.004)	(0.010)
refyear10	0.011	-0.003	-0.012	0.002	0.004	-0.008
	(0.018)	(0.009)	(0.015)	(0.010)	(0.004)	(0.009)
N	620	620	619	620	620	620
N_countries	31	31	31	31	31	31
R^2 adj	0.723	0.341	0.439	0.532	0.499	0.086

Standard errors in brackets - country and time fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001

Table 7 - Citations per paper relative to field world average - all fields - 1996-2015

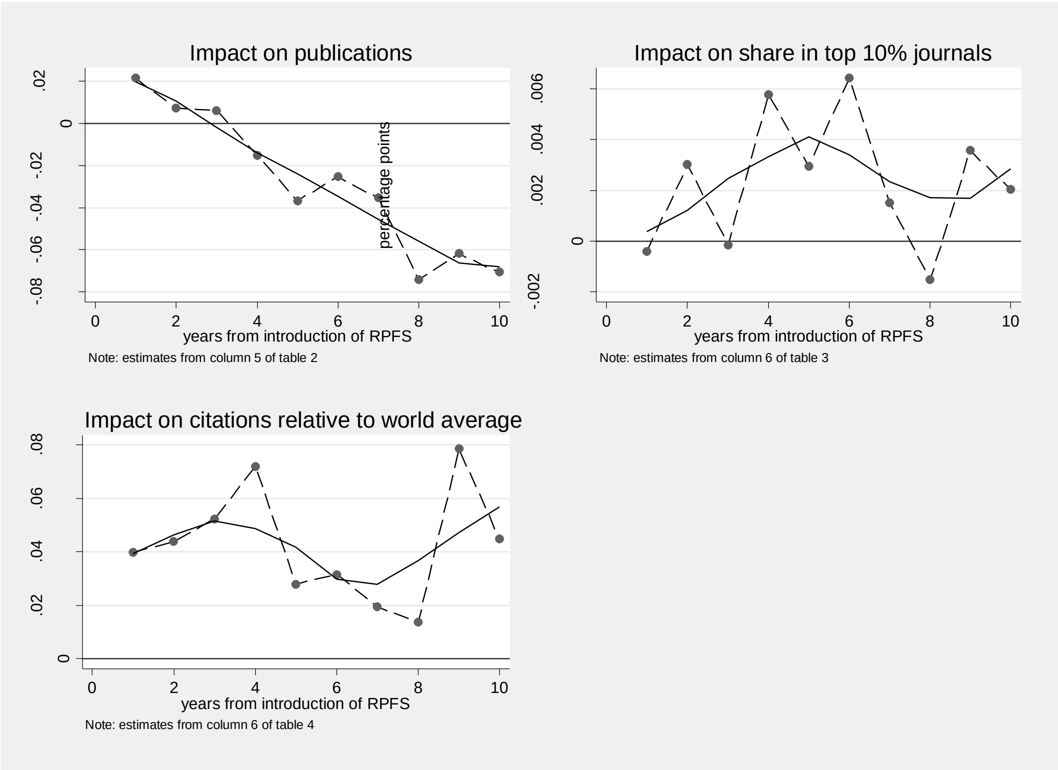
	(1) agricultural	(2) engineering	(3) humanities	(4) medical	(5) natural	(6) social
refyear1	0.010	0.072*	-0.046	0.042	0.029	0.055
	(0.023)	(0.031)	(0.057)	(0.028)	(0.027)	(0.038)
refyear2	0.020	0.060*	-0.028	0.070	0.027	0.010
	(0.025)	(0.028)	(0.055)	(0.044)	(0.020)	(0.033)
refyear3	0.028	0.050	-0.043	0.120	0.026	0.032
	(0.028)	(0.034)	(0.072)	(0.092)	(0.020)	(0.028)
refyear4	0.060+	0.019	0.073	0.168	0.045+	-0.008
	(0.031)	(0.036)	(0.070)	(0.119)	(0.026)	(0.042)
refyear5	0.018	0.020	-0.030	0.007	0.027	-0.001
	(0.033)	(0.025)	(0.046)	(0.030)	(0.020)	(0.030)
refyear6	0.044+	0.024	0.025	0.016	0.026	-0.030
	(0.025)	(0.038)	(0.060)	(0.040)	(0.018)	(0.040)
refyear7	0.030	0.004	0.058	-0.021	0.035	-0.011
	(0.046)	(0.038)	(0.069)	(0.044)	(0.031)	(0.034)
refyear8	0.002	0.022	0.105	0.003	-0.001	0.041
	(0.039)	(0.040)	(0.077)	(0.046)	(0.033)	(0.040)
refyear9	0.064	0.120***	0.095	0.096	0.042	0.040
	(0.042)	(0.028)	(0.056)	(0.069)	(0.035)	(0.060)
refyear10	0.042	-0.011	0.096*	0.111*	-0.007	-0.028
	(0.044)	(0.043)	(0.045)	(0.049)	(0.021)	(0.025)
N	620	620	619	620	620	620
N_countries	31	31	31	31	31	31
R^2 adj	0.436	0.277	0.045	0.566	0.657	0.348

Standard errors in brackets - country and time fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001

5. Conclusions

Our main results can be visualised by the following graph, which plots the time pattern of reaction of national research systems to the introduction of performance based funding systems.

Figure 8: Time profile of the impact of RPFS



References

Adams, J., & Gurney, K. (2014). *Funding selectivity, concentration and excellence - how good is the UK's research?* . HEPI.

Braun, D. (2003). Lasting tensions in research policy-making—a delegation problem. *Science and Public Policy*, 30(5), 309-321.

Butler, L. (2003a). Modifying publication practices in response to funding formulas. *Research Evaluation*, 12(1), 39–46.

Butler, L. (2003b). Explaining Australia's increased share of ISI publications—The effects of a funding formula based on publication counts. *Research Policy*, 32, 143–155.

Capano, G., Turri, M., 2017. Same Governance Template but Different Agencies, *Higher Education Policy*, Volume 30, Issue 2, pp. 225–243.

Cruz-Castro, L., Bleda, M., Derrick, GE., Jonkers, K., Martinez, C., and Sanz-Menendez, L., 2011. Issue brief.: Public sector research funding.

De Boer, HF., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., Lemmens-Krug, K., Vossensteyn, H., 2015. Performance-based funding and performance agreements in fourteen higher education systems - Report for the Ministry of Education, Culture and Science, Enschede: CHEPS, University of Twente

Geuna, A., Martin, B.R., 2003. University research evaluation and funding: an international comparison. *Minerva* 41, 277–304.

Geuna, A., Piolatto, M., 2016, Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while), *Research Policy*, Volume 45, Issue 1, Pages 260–271

Hazelkorn, E., 2010. Pros and cons of research assessment, in ISSC/UNESCO *World Social Science Report 2010*, Paris: UNESCO Press

Herbst, M. (2007). Financing public universities. *Higher education dynamics*, 18, 90.

Hicks, D, 2009. "Evolving Systems of Multi-University Research Evaluation", *Higher Education* 57(4): 393-404.

Hicks, D., 2010. Overview of models of performance-based research funding systems, in OECD, 2010. Performance-Based Funding for Public Research in Tertiary Education Institutions, Workshop Proceedings, OECD: Paris, pp. 23-52.

Hicks, D., 2012. Performance-based university research funding systems, *Research Policy*, Volume 41, Issue 2, 2012, Pages 251–261

Hood, C., Schuppert, O., 1988. *Delivering Public Services in Western Europe: Sharing Western European experience of para-government organisation*, London: Sage.

Jonkers, K., Zacharewicz, T., 2016. *Research Performance Based Funding Systems: a Comparative Assessment*, Publications Office of the European Union, Luxembourg,

Kalpazidou Schmidt, E., Langberg, K., Aagraard, K., 2006. Funding Systems and Their Effects on Higher Education Systems. Country study – Denmark, Paris OECD.

Kettl D.F., 2005. *The Global Public Management Revolution*, Brookings Institution Press, Washington, D.C.

Martin B., 2011. The Research Excellence Framework and the impact agenda, are we creating a Frankenstein Monster?, *Research Evaluation*, 20(3) 247-254

Moed, H. F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity?. *Scientometrics*, 74(1), 153-161

OECD, 2010. Performance-Based Funding for Public Research in Tertiary Education Institutions, Workshop Proceedings, OECD: Paris.

Olson M., 1973. Evaluating performance in Public Sector. In Moss M. (eds) The Measurement of Economic and Social Performance, National Bureau of Economic Research, Columbia University Press, pp. 355-409.

Osuna, C, Cruz-Castro, L, and Sanz-Menéndez, L 2011, Overturning some assumptions about the effects of evaluation systems on publication performance, *Scientometrics* 86:3, 575-592

Radosevic, S., Lepori, B., 2009. Public research funding systems in central and eastern Europe: between excellence and relevance: introduction to special section, *Science and Public Policy*, 36 (9): 659-666.

Perrin, B. (2007). Towards a new view of accountability, in Bemelmans-Videc, M. L., Lonsdale, J., & Perrin, B. (Eds). *Making accountability work: Dilemmas for evaluation and for audit* (Vol. 14). Transaction Publishers, pp. 41-59.

Perrin, B. (2015). Bringing accountability up to date with the realities of public sector management in the 21st century. *Canadian Public Administration*, 58(1), 183-203.

Pollitt, C., Bouckaert, G. (2011). *Public Management Reform: A comparative analysis-new public management, governance, and the Neo-Weberian state*. Oxford University Press.

Rafols I., Leydersdorff L., O'Hare A., Nightingale P., Stirling A. (2012), How Journal rankings can suppress interdisciplinary research: a comparison between innovation studies and business and management, *Research Policy*, vol. 41, pp. 1262-1282

Rebora G, and Turri M. (2013) 'The UK and Italian Research Assessment Exercises Face to Face', *Research Policy* 42(9): 1657–1666.

Sörlin, S. (2007), "Funding Diversity: Performance-based Funding Regimes as Drivers of Differentiation in Higher Education Systems", *Higher Education Policy* 20, 413-440

Stame, N. (2016). *Valutazione pluralista*. FrancoAngeli, Milano.

Tapper, T., Salter, B., 2003. Interpreting the process of change in higher education: the case of the Research Assessment Exercises. *Higher Education Quarterly* 57, 4–23.

Vanecek, J., 2013, The Effect of Performance-based Research Funding on Output of R&D Results in the Czech Republic'. *Scientometrics* 10, 1–25.

Verhoest, K., Thiel, S. van, Bouckaert, G. and Laegreid, P. (2012) *Government Agencies: Practices and Lessons from 30 Countries*, London: Palgrave.

Wang, J., & Hicks, D. (2013). Detecting structural change in university research systems: A case study of British research policy. *Research Evaluation, 22*, 258-268.

Whitley, R., 2007, Changing governance of the Public Sciences, The Consequences of establishing research evaluation systems for knowledge production in different countries and different fields in Whitley, R., Glaser J (Eds) The Changing Governance of Science, Dordrecht: Springer

Appendix

Table 8 - Total publications (logs) - all fields - 1996-2015

	(1) com.trend	(2) dif.trend	(3) yr.dummy	(4) com.trend	(5) dif.trend	(6) yr.dummy
reform	0.141	0.014	0.169+	-0.188	-0.210	-0.210
	(0.089)	(0.038)	(0.093)	(0.184)	(0.202)	(0.202)
reform×time	-0.012	0.002	-0.014			
	(0.009)	(0.003)	(0.011)			
refyear1				0.186	0.203	0.203
				(0.151)	(0.162)	(0.162)
refyear2				0.170	0.187	0.187
				(0.138)	(0.148)	(0.148)
refyear3				0.168	0.184	0.184
				(0.128)	(0.139)	(0.139)
refyear4				0.164	0.179	0.179
				(0.117)	(0.132)	(0.132)
refyear5				0.134	0.152	0.152
				(0.112)	(0.122)	(0.122)
refyear6				0.148	0.169	0.169
				(0.105)	(0.114)	(0.114)
refyear7				0.122	0.148	0.148
				(0.105)	(0.112)	(0.112)
refyear8				0.110	0.125	0.125
				(0.095)	(0.102)	(0.102)
refyear9				0.095	0.081	0.081
				(0.077)	(0.089)	(0.089)
refyear10				0.084	0.065	0.065
				(0.060)	(0.071)	(0.071)
exp.R&D/GDP	0.158+	0.044	0.141	0.194+	0.188+	0.188+
	(0.085)	(0.060)	(0.085)	(0.098)	(0.097)	(0.097)
log(researchers)	-0.044	0.118	-0.055	-0.071	-0.092	-0.092
	(0.350)	(0.123)	(0.351)	(0.369)	(0.375)	(0.375)
log(PhDs)	0.244**	0.204*	0.222**	0.237**	0.215*	0.215*
	(0.079)	(0.086)	(0.075)	(0.082)	(0.079)	(0.079)
N	500	500	500	500	500	500
N.countries	25	25	25	25	25	25
R^2 adj	0.859	0.962	0.865	0.857	0.862	0.862

Standard errors in brackets - country fixed effects included - common time trend, country-specific time trend or year fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001

Appendix

Table 9 - Share of articles in top10% journals - all fields - 1996-2015

	(1) com.trend	(2) dif.trend	(3) yr.dummy	(4) no trend	(5) com.trend	(6) yr.dummy
reform	0.010+	0.004	0.006			
	(0.005)	(0.007)	(0.005)			
refyear1				0.001	-0.000	0.000
				(0.004)	(0.004)	(0.004)
refyear2				0.008+	0.006	0.005
				(0.005)	(0.006)	(0.005)
refyear3				0.008	0.004	0.001
				(0.006)	(0.006)	(0.005)
refyear4				0.016*	0.011	0.008
				(0.007)	(0.007)	(0.005)
refyear5				0.016+	0.011	0.006
				(0.008)	(0.008)	(0.006)
refyear6				0.021+	0.015	0.009
				(0.011)	(0.011)	(0.007)
refyear7				0.015	0.009	0.003
				(0.011)	(0.009)	(0.007)
refyear8				0.001	-0.005	-0.004
				(0.011)	(0.009)	(0.006)
refyear9				-0.003	-0.006	0.005
				(0.007)	(0.006)	(0.006)
refyear10				0.001	-0.002	0.003
				(0.008)	(0.007)	(0.005)
exp.R&D/GDP	0.004	-0.003	0.004	0.015*	0.007	0.005
	(0.007)	(0.009)	(0.006)	(0.007)	(0.007)	(0.006)
log(researchers)	0.013	-0.022	0.020+	0.029*	0.012	0.019+
	(0.010)	(0.014)	(0.010)	(0.011)	(0.011)	(0.011)
log(PhDs)	-0.017+	-0.009	-0.012	-0.006	-0.017+	-0.012
	(0.009)	(0.007)	(0.010)	(0.009)	(0.010)	(0.010)
N	500	500	500	500	500	500
N.countries	25	25	25	25	25	25
R^2 adj	0.376	0.542	0.575	0.299	0.370	0.568

Standard errors in brackets - country fixed effects included - common time trend, country-specific time trend or year fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001.

Appendix

Table 10 - Citations per paper relative to field world average - all fields - 1996-2015

	(1) com.trend	(2) dif.trend	(3) yr.dummy	(4) no trend	(5) com.trend	(6) yr.dummy
reform	0.058*	0.030	0.062*			
	(0.025)	(0.024)	(0.026)			
refyear1				0.044*	0.036+	0.043*
				(0.020)	(0.018)	(0.020)
refyear2				0.050+	0.034+	0.043*
				(0.026)	(0.018)	(0.020)
refyear3				0.083+	0.057	0.058
				(0.044)	(0.036)	(0.037)
refyear4				0.129*	0.091+	0.089
				(0.059)	(0.051)	(0.052)
refyear5				0.072*	0.036	0.038+
				(0.032)	(0.022)	(0.019)
refyear6				0.088*	0.045+	0.046*
				(0.038)	(0.024)	(0.022)
refyear7				0.062	0.016	0.016
				(0.049)	(0.034)	(0.037)
refyear8				0.073	0.024	0.022
				(0.052)	(0.037)	(0.040)
refyear9				0.106	0.087	0.083
				(0.066)	(0.059)	(0.062)
refyear10				0.061*	0.035	0.025
				(0.027)	(0.023)	(0.033)
exp.R&D/GDP	0.056	-0.010	0.061	0.124*	0.065	0.070
	(0.041)	(0.027)	(0.045)	(0.048)	(0.042)	(0.046)
log(researchers)	0.071	0.056	0.064	0.191**	0.064	0.057
	(0.050)	(0.054)	(0.054)	(0.055)	(0.050)	(0.054)
log(PhDs)	0.033*	-0.003	0.033*	0.107***	0.027	0.027
	(0.015)	(0.027)	(0.015)	(0.025)	(0.016)	(0.017)
N	500	500	500	500	500	500
N.countries	25	25	25	25	25	25
R^2 adj	0.780	0.865	0.786	0.695	0.779	0.783

Standard errors in brackets - country fixed effects included - common time trend, country-specific time trend or year fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001.

Appendix

Table 11 - Outcomes by performance assessment type - 1996-2015

	(1) bibliometric log(publications)	(2) bibliometric top10	(3) bibliometric field w.citat.	(4) peer rev. log(publications)	(5) peer rev. top10	(6) peer rev. field w.citat.
reform	-0.318+	0.007	-0.078	-0.234	-0.005	-0.002
	(0.150)	(0.015)	(0.076)	(0.401)	(0.007)	(0.036)
refyear1	0.257+	-0.005	0.104	0.251	0.001	0.001
	(0.133)	(0.016)	(0.075)	(0.297)	(0.007)	(0.042)
refyear2	0.254+	-0.004	0.089	0.182	0.003	0.015
	(0.118)	(0.017)	(0.057)	(0.240)	(0.005)	(0.035)
refyear3	0.262+	-0.007	0.109	0.174	-0.001	0.019
	(0.116)	(0.015)	(0.062)	(0.210)	(0.007)	(0.034)
refyear4	0.262+	0.000	0.158	0.140	0.006	0.048
	(0.125)	(0.013)	(0.118)	(0.178)	(0.008)	(0.027)
refyear5	0.194+	0.002	0.064	0.123	0.003	0.032
	(0.097)	(0.014)	(0.055)	(0.159)	(0.009)	(0.026)
refyear6	0.164	0.001	0.056	0.140	0.009	0.011
	(0.091)	(0.014)	(0.037)	(0.149)	(0.008)	(0.019)
refyear7	0.172+	-0.004	0.041	0.100	-0.003	0.011
	(0.082)	(0.009)	(0.030)	(0.164)	(0.006)	(0.036)
refyear8	0.198**	-0.012	0.038	-0.042	0.007*	0.001
	(0.059)	(0.007)	(0.028)	(0.140)	(0.003)	(0.041)
refyear9	0.193**	-0.002	0.097	0.115	-0.002	0.019
	(0.054)	(0.006)	(0.057)	(0.082)	(0.006)	(0.020)
refyear10	0.161**	-0.002	0.043	0.071	-0.006	0.081**
	(0.040)	(0.007)	(0.035)	(0.057)	(0.006)	(0.014)
N	200	200	200	140	140	140
N_countries	10	10	10	7	7	7
R^2 adj	0.943	0.654	0.799	0.842	0.473	0.808

Standard errors in brackets - country and time fixed effects included - clustered errors by country - statistical significance + 0.10 * 0.05 ** 0.01 *** 0.001