Workshop on "Assessing the Impact of Science Funding in Europe"

Bruegel, Rue de la Charité 33

Brussels

Tuesday January 27, 2015

Power Point presentations are available on SIMPATIC website (www.simpatic.eu).

9:00-9:10 Welcoming remarks

Domenico Rossetti, DG RTD, EC, SIMPATIC project officer

In his opening speech, Domenico Rossetti reminded that the budget to R&I has never been so large: €80 bn in H2020 and €220 bn in ESIF. However, due to the crisis, budget competition for different purposes has also became fiercer. Therefore, he mentioned how critical it is right now to assess the impact of science funding from a socio-economic perspective.

The main aspects of interest for the EC to be analysed were:

- Leverage effects of science funding: how much more private investment can 1 euro of public money generate;
- Sectorial spillovers: are there sectors for which public investment can indirectly generate more investment in other sectors;
- EU vs national funding: are there significant differences between these types of funding.

For the future, the next key aspects of attention for H2020 projects are likely to be the measurement of progress beyond GDP and the economic assessment of innovation policies.

9:10-09:30 Introduction to the workshop

Reinhilde Veugelers, Scientific Coordinator of SIMPATIC, KU Leuven & Bruegel

The availability of big data and methodological improvements have been contributing to develop the science of impact assesment. As explained by Reinhilde Veugelers, with better measures of quality and new quantitative methodologies (e.g. RCTs, Differences in Differences or Matching tecnhiques) an increasing number of challenges are being addressed, as the establishment of causal relationships, selection effects and heterogeneity problems. Given the multiplicity of methodologies, data and results, it is important to find common grounds for comparability purposes. As a result, SIMPATIC contributes to coordinate different projects in this field and thus build a better pool of knowledge by allowing for comparability of its findings.

9.30 - 10.00 Presentation of the KU Leuven pilot

Reinhilde Veugelers, KU Leuven & Bruegel; **Otto Toivanen**, KU Leuven; **Cindy Lopes Bento** KU Leuven; **Stijn Kelchtermans**, KU Leuven

In this project the authors used a panel dataset of professors in the biomedical and exact sciences from KU Leuven to study the effects of research funding from the regional government. The 3 main objectives are 1) to understand who is more likely to get grants and what is the impact of that funding; 2) whether research funding is supporting young researchers and contributes to their career development; and 3) are these internal grants helping researchers to apply for and get European funding.

As the best research projects might be the ones that attract more funding, the authors need to address the problem of reverse causality when studying the impact of funding on the quality of projects. Thus, lagged and pre-sample independent variables were used, along with an IV and 2 step estimation procedure. Moreover, fixed effects were used to take into account unobserved heterogeneity correlated with funding and productivity, as for example talent.

Some first preliminary results indicate that young researchers are more likely to be funded by smaller grants, once corrected for other effects (like past funding record, past publication track record, talent, etc...). Therefore, these smaller grants are acting as an "entry" program for young researchers. However, being funded by smaller grants does not seem to be helping young researchers to progress to the bigger grants. In fact, young researchers have shown a stronger persistency into being funded by smaller grants.

The next steps of this research will be to analyse the impact of funding in the progression to top performance, on the research networks and on the likelihood of attracting other sources of funding.

10.00 - 10.30 Presentation of the Swiss National Science Foundation pilot

Fabiana Visentin, EPFL ; Other team members: **Dominique Foray**, EPFL; **Jacques Mairesse**, ENSAE; **Michele Pezzoni**, Bocconi University

The aim of this project is to evaluate the success of the project Sinergia. This programme is one of funding schemes of the Swiss National Science Foundation, which aims at promoting multidisciplinary research. In order to receive this funding, applicants need to prove to have a multidisciplinary, collaboration-based project.

The key research question of this study are to assess the impact of a Sinergia grant on:

- The creation of scientific collaborations;

- The entry of scientists into new scientific fields;
- The creation of employment opportunities for junior researchers.

By using a dataset containing information on 232 grants awarded between 2008 and 2014, the authors will answer the above research questions by computing the effect of receiving a grant on the:

- Probability of publishing a paper in co-authorship (as a measure of scientific collaboration);
- Probability of a researcher to publish a paper in a new subject category (as a measure of entry into new scientific fields);
- Probability of a young researcher to be employed in a highly-ranked university after graduation.

This will allow for the evaluation of the grant efficacy along different social dimensions: collaborations, entry into new scientific fields and job opportunities.

11.00 - 11.30 Presentation of the University of Strasbourg pilot

Patrick Llerena, University of Strasbourg; Other team members: **Stefano Bianchini**, University of Strasbourg; **Julia Lane**, American Institutes for Research and University of Strasbourg;

In order to measure the impacts of science funding, there are 3 key questions one must pose:

- What are the returns to investment in research?
- How are the returns created?
- How can national governments foster activities that increase those returns?

These questions can only be answered with a comprehensive measurement of research inputs, outputs and processes, as it is performed in the US StarMetrics program.

In line with the StarMetrics, the University of Strasbourg pilot aims at building a longitudinal microdataset to describe the composition of the workforce, the purchases of inputs and the geographical dimension of research expenditures in France. This information can be matched with output statistics.

The next steps of this project will be to extend the current data to more research units active in different scientific fields and perform a cross-country comparison and harmonization. The authors are particularly interested in studying the effect of team composition on scientific productivity and the short-term effects of science spending on the real economy. They also aim at examining the performance and innovativeness of firms that constitute part of the scientific value chain.

11.30 - 12.00 Presentation of the ANR pilot

Nicolas Carayol, University of Bordeaux

The ANR pilot is an on-going work in France, concerning one specific source of funding – the ANR (created in 2005 to implement project-based funding). The objective of this study is to evaluate the impact of ANR's funding on a researcher's performance, against the counterfactual where he or she does not receive it.

While the data on funded projects is available at the ANR records, information on the unfunded projects is not. Therefore, the counterfactual data on potential grantees was compiled on the basis of the full population of potential users. This allowed the authors to apply a propensity score matching estimation method based on two candidate sets of variables:

- Age, volume of publications, citations and top citations;
- Volume, maximum impact and size of the laboratory.

The results suggest that the research performance of the funded researchers are significantly higher than the reference group.

For the next steps, this project will apply the same methodology to different programs and make a connection with the ERC.

12.00 - 12.30 Other potential pilots

UK's Engineering and Physical Sciences Research Council:

Albert Banal, UPF; **Ines Macho-Stadler**, **David Perez-Castrillo**, UA Barcelona

This project tries to analyze the impact of the UK's Engineering and Physical Science Research Council (EPSRC), which provides funding to engineering and physical science research in UK. In partcular, it tries to analyze the causes and the consequences of (i) the cooperation among academic researchers, (ii) the collaboration between teams of academic researchers and firms to develop research projects and (iii) the key drivers for the success of grants applications, using data of grants awarded from 1975 to 2007. The ideal type of data (that they do not have access to) would be comprised of data on both awarded and rejected grants applications (while only the former are available), pool of applicant and non applicant researchers (only the first are available) and pool of potential partners (only the ones collaborating with teams of researchers participating in the program are available).

The study finds that the characteristics of partnering firms matter on the research output (number and quality of the publication). To better understand how firms and teams of academics decide to

collaborate on a research project, a matching model is developed with both "horizontal" (i.e. affinity: preference for specific types or specific fields of research) and "vertical" (i.e. ability: capacity to produce high quality research) characteristics was developed. Positive assortive matching is found, i.e. scientists need to match with similarly minded firms. On which characteristics affect the probability of obtaining a grant, they find persistance in grants' award but also considerable dynamics in team formation.

OECD's Knowledge Triangle Project: **Caroline Paunov**, OECD

The topic of this presentation was the present the OECD's plans on studying the impact of public reasearch on the innovation ecosystem. The "knowledge triangle project" was sketched in which ideally research, education and innovation mutually affect each other. The main questions that need to be answered in this respect are:

- What is the impact of public research to innovation and how does it contributes to the innovation ecosystem?
- What is the impact of the policies implemented to enhance the contribution of public research to innovation?
- How to best evaluate the impact of these policies?

The objective of this project is to provide a quantitative picture of the contributions and impacts of different mechanisms through (i) data-driven analyses of public research and its impact and (ii) by identifying "best practices" used in national evaluation and findings. The policy priorities of this analysis should be:

- Identifying the most effective impact channels;
- Implications of commercialization for public research;
- Identifying the area of research to prioritize with a look at the needs of innovative industries;
- Building absortive capacities for firms to use reasearch;
- Identifying national and international research priorities.

The suggested approach was to measure the quality, quantity and characteristics of public research output that might be of relevance for industry from the patents and publications generated by a research project, and to study the impact of public research on innovation by matching firm- and industry-level data on innovation performance and other characteristics. The policy analysis should be based on the "state of the art" of national impact assessment exercises and on developing policy indicators for quantitative impact assessment based on the OECD Science, Technology and

Innovation Outlook. Finally, the website of the Innovation Policy Platform (https://www.innovationpolicyplatform.org) – a collaborative and interactive web-platform (co-developed by the OECD with the World Bank) that provides access to data on impact assessment of innovation policies – was briefly presented.

12.30 - 13.00 "How to combine reviews in panel evaluations"

Marco Ottaviani, Bocconi University; **Peter N. Sørensen**, University of Copenhagen

This study looked at the microengineering of the grants award process and management (i.e. how the grants get awarded). Since the effectiveness of science funding rests on the quality of the review process, the mechanism through which projects get evaluated (and hence grants get assigned) is of paramount importance. Typically, a panel of experts – who get assigned to evaluate a given project on the basis of their expertise – report on their reviews of the project independently in the form of a score and a first (rough) ranking is obtained by aggregating the individual scores in a simple arithmetic average. The ranking is then refined through panel discussion(s). The author presented a model which tries to determine the optimal mechanism for the aggregation of reviews and some empirical applications of the model to evaluate the assignments of experts to specific projects and the design of the panel in terms of experts' incentives to review the project properly.

In the model, upon reviewing the project, each expert, also on the basis of his knowledge, forms his own idea about the quality of the project (formally, he receives privately a conditionally independent signal of the quality which diverges from the "true" quality by a zero-mean error with idiosyncratic precision) and submits a review with which he is trying to forecast the quality of the project. Assuming the reviewers generate the score/review they submit by linearly combinating their signal (the idea they made about the quality of the project influenced by their knowledge) and the ex-ante expected quality of the project (the prior) plus a idiosyncratic effect (e.g. a personal bias), the study tries to assess how the aggregation of reviews generated by simple heuristic methods (e.g. arithmetic average), compares with the benchmark of a Bayesian observer, which estimates the quality of the project conditional on the submitted review and which represents the best way to estimate the quality of the project given the submitted reviews. The study shows that, even in the case in which the individual reviewers are Bayesian (i.e. they show no idiosyncratic bias and give a weigth to their signals which is proportional to the precision of the signal –given by the ratio of the precision of their signal over the sum of the precision of the signal and the precision of the prior) and are all equally informed (i.e. the signal they receive has the same precision), the arithmetic average tends to put too much weight on the prior (i.e. the ex-ante expected quality of the project). This is an important finding that could explain the tendency to conservatism in most evaluation exercises.

Finally, the author showed some empirical application of the model. The first regarded the Bocconi job market committee and the joint structural estimation of the quality of the candidates (he emphasized the scores assigned were in general taking into account not only the quality of the candidate but also the likelihood the candidate would be accepting the offer) and the precision of the reviewer. They show that the ranking of candidates changes after controlling for idiosyncratic fixed effects, meaning that some personal bias was affecting the reviews. The second example was about reviews on restaurants on Yelp and emphasized the challenge for the identification of the parameters when there are many experts with few reviews and few overlaps across reviewers. The author would like to study the ERC review panels as it is a nice example with a wealth of information to study the impact of overlaps across reviewers that allows the identification of parameters and an illustration of the review matrix with the estimated precision of the reviewers and quality of the projects.

14.00 - 14.30 "How can we know science funding works? Turning the scientific method on ourselves"

Pierre Azoulay, MIT (through videoconferencing)

The presentation started by analyzing current critiques of the public funding of medical research in the US by the National Institute of Health (NIH). There have been a series of complaints and debates about the effectiveness of these programs and whether they were producing important research. In particular, there is a growing concern that public funding agencies are failing to fund innovative research. However, the speaker pointed out that any meaningful criticisims of the system should clearly identify a counterfactual to which the performances of the system can be compared to. He further stressed that NIH should not just focus on breakthrough research, but should also incentivize "normal" incremental or follow-on discoveries – which are still extremely important for society, and that public funding policies should not ignore the rest of the ecosystem (there are many private funders in the US and public funding should not necessarily mimic what private funders are doing). He then pointed out some of the features of the current grant system that might not incentivize scientific risk-taking. In particular, he noticed that the current system is not very forgiving of failures, as it foresees limited timelines (3 to 5 years maximum) to show results and that there is limited flexibility for a researcher to adjust his agenda as new scientific opportuniy arise. The short-term horizon and limited flexibility probably give incentive to researchers to "play it safe".

He then presented a comparison of the NIH funding program with a private funding system provided by the Howard Hughes Medical Institute (HHMI). The main difference between the two programs in his opinion is in the time-horizon given to researchers in order to produce results. While both programs have basically a 5-year horizon, the first review in the HHMI funding program is rather lax and does not revolve on showing results but rather on whether funds have been used for something

that could have not been done without the grant, which in practice implies that the time-horizon to show results is around 10 years. A comparison of outcomes of comparable scientists in the NIH and HHMI funding programs show that, after 2-to-3 years, HHMI funding increases both tails of the distribution – research funded by HHMI tends to fail more often, but tends also to generate more successful (top-5 percentile) outcomes – and increases the use of different and novel keywords (a proxy for more innovative research), which is in line with what one would expect from a program incentivizing more risk-taking.

Finally, he emphasized some issues with the current system – in particular how can one separate the effect of resources from the incentive mechanism, since also the control group is well funded – and provided some thoughts on how a potential (better) new grant mechanism could look like. He also proposed the use of randomized control trials to evaluate the effectiveness of different funding policies.

14.30 - 15.00 "Analysing funding impact at Caltech; a US Starmetrics pilot"

Paula Stephan, Georgia State; **Jacques Mairesse**, ENSAE; **Michele Pezzoni**, Bocconi University; **Julia Lane**, American Institutes for Research

The presentation was based on a Starmetrics pilot conducted in the US at Caltech. Public funding places government agencies and universities under pressure to show relationship between research inputs and outputs. However, there is a lack of both methodology and data to support an analysis of the relationship between research output and research input, in particular with funding. In this study the authors do an analysis at the grant level to investigate differences in returns to public funding. They need to link the outcomes of the research to the grant. While there are many outcomes (publications, patents, PhD dissertations, PhD placements in academia and industry, etc.), this study focuses only on publications of research. There are several methods to identify this link, but they focus on just two, one based on a pure chronological order, i.e. the publication followed the grant, and direct links through student coauthorship and acknowledgment of funding received.

They found that accounting for attribution is important for predicting the quantity of output related to a specific flow of funds; but it does not seem to make a difference for measuring elasticity . However, the results are for a highly selective university and for federal grants and are hence preliminary.

15.30 - 17.30 Panel Discussion with key stakeholders on the why and how (not) of a European network for assessing the impact of science funding

Chair: **Reinhilde Veugelers**, KU Leuven & Bruegel Interventions: **Luc Soete**, U.Maastricht & RISE; **Julia Lane**, US Starmetrics; **Kurt Deketelaere**, LERU; **Jean Pierre**

Bourguignon, ERC; **Roman Arjona-Gracia**, DG RTD, EC; **Fréderique Sachwald**, Ministère de l'Education nationale, de l'Enseignement Supérieur et de la Recherche

This panel focused on discussing the current challenges and bottlenecks affecting the creation of a European network for assessing the impact of science funding.

It was consensual among the speakers that impact assessment in Europe is especially difficult due to Europe's research fragmentation. Models that function in one country do not necessarily work in others, thus funding policies cannot ignore these idiosyncrasies. This also raises the need to compare the impacts of science funding and investments at European versus national level, as well as the interactions between both.

It was also mentioned that there is still a lack of awareness for the need of financing science in Europe, and, due to the crisis, public funding has become more controversial. Consequently it is more crucial than ever to have more real evidences on the socioeconomic effects of science funding, as these can be used as tools to lobby politicians about the importance of financing science. These evidences cannot be limited only to the number of publications and patents, but need to go beyond these simple metrics, in such a way it can be shown whether or not science is empowering society.

Metrics are therefore critical for this analysis. It was explained that many of the existing measurements are outdated and/or inappropriate. Hence, it is essential to carefully think about what one really wants to measure, because metrics are needed to define success and understand the reasons for failure. The choice of metrics need to be done carefully taking into account that this choice will also influence the behavior of agents, as agents will adjust in accordance to these.

Finally, the panel unanimously endorsed that in order to fulfil these objectives one needs more data. These data are often already available at universities, statistical offices and increasingly also in social media. However, it is not always accessible to the researcher to be matched. Therefore, efforts need to be done in order to make these sources of data available.