



Baština Akademije nauka i umjetnosti Bosne i Hercegovine

Sistem nauke-faktor poticaja ili ograničavanja razvoja: Science system-a Factor of Stimulation or Limitation in Development

Trifković, Miloš

2021-10-27

Academy of Sciences and Arts of Bosnia and Herzegovina

<https://bastina.anubih.ba/items/fe55fd34-af42-4399-8913-79b68cd89f42>

Sistem nauke

Preuzeto s Baštine Akademije nauka i umjetnosti Bosne i Hercegovine

<https://bastina.anubih.ba/>

Reforming Science Evaluation and Funding when Starting from a Low Performance Level

Key note address

*Štěpán Jurajda**

The Challenge

Many post-communist countries recently enjoyed rapid economic convergence towards their Western European counterparts, but there has been surprisingly little convergence in scientific research performance. In 2015 only about 3% of the European Research Council's (ERC) Starting Grants originated from the EU's new members. As demonstrated by, e.g., Jurajda et al. (2016), the post-soviet countries of Europe are still lagging far behind in terms of high-quality academic output, in particular in social and medical sciences, which are critical for the quality of local governance on many policy levels. Post-communist countries' publication output is also heavily focused on quantity as opposed to quality. In short, bibliometric evidence implies that several fields of science in these countries are consistently unable to catch up to EU-15 (world-level) scientific standards.

Are there science policies that could be adopted to close this persistent gap? Several countries have recently experimented with science evaluation and funding systems that rely on quantitative measures such as (quality-weighted) publication or citation counts. Much of the experience from these reforms has been negative in that quantitative indicators often lead to opportunistic behavior aimed at maximizing indicators, not the quality of scientific output. The typical lesson from these experiments is to rely on peer review of scientific output. However, such peer review often must rely on local scientists and is therefore fundamentally problematic in fields of science where a given country lags far behind international standards. More broadly, the

* CERGE-EI, a joint workplace of Charles University in Prague and the Economics Institute of the Czech Academy of Sciences. E-mail: stepan.jurajda@cerge-ei.cz

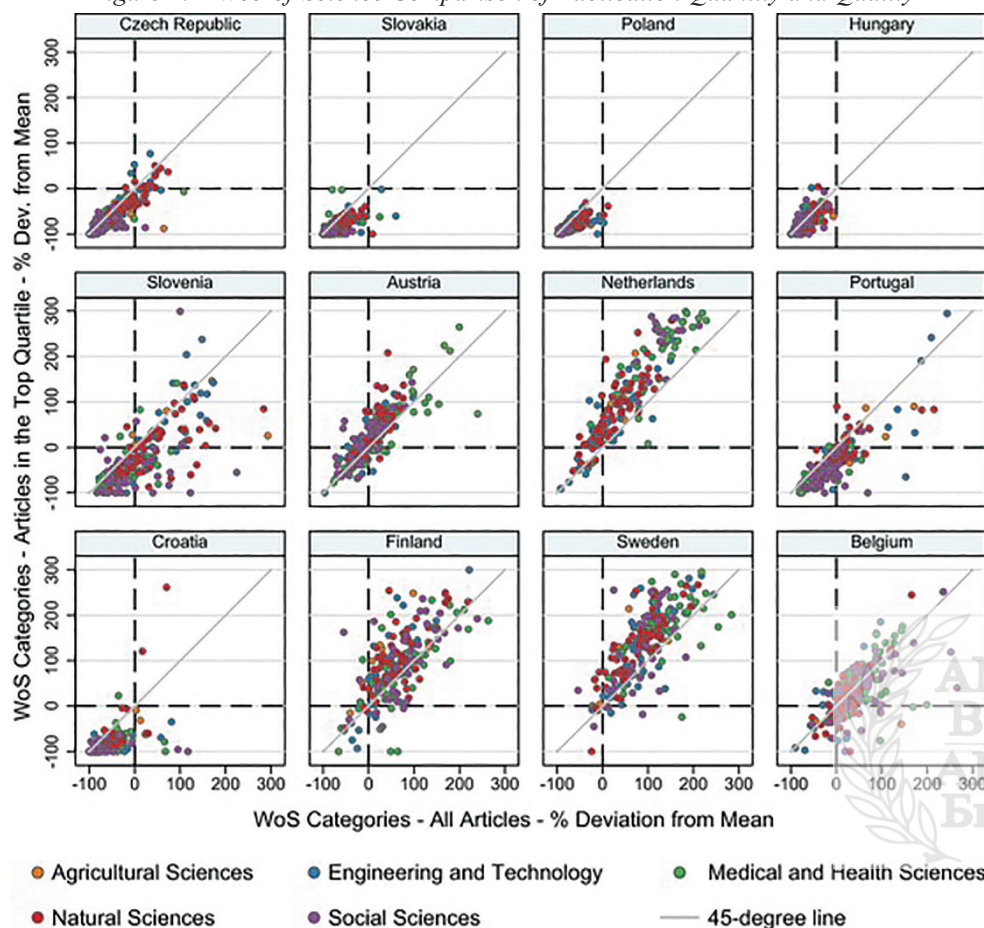
difficult question is how one can reform science funding and evaluation to help improve performance when entire national fields of science are lagging far behind international standards?

This is a difficult question in part because field-wide low performance levels typically combine with low funding levels and with internal democratic processes in academia to prevent reforms and rapid change. This note provides a stylized description of recent reforms of science funding and evaluation in the Czech Republic. It aims to illustrate these difficulties and offer some suggestions on how to break reform gridlocks by offering two-tier funding schemes. The discussion is in part based on descriptive evidence from the Czech Republic and on the relevant scientific literature covering other EU countries, but most of the views and recommendations offered here are those of a reformer involved in the process of proposing, adopting and implementing a reform of science funding and evaluation, i.e., they are not particularly scientific.

Persistently Low Relative Performance in Social Sciences

There is growing evidence that post-communist countries have not yet been able to close the gap in scientific performance vis-à-vis the EU-15 countries. Figure 1 (taken from Jurajda et al., 2017) presents a 2010-2014 bibliometric comparison based on the Web of Science (WoS) of publication performance across 226 scientific disciplines for six post-communist EU member states relative to six EU-15 countries of a comparable size. The graph contrasts each country to the average performance of the entire sample in terms of per capita article counts on the horizontal axis and in terms of per capita article counts in the most influential journals, as judged based on a journal belonging to the best quarter of journals according to the Article Influence Score (AIS), a journal citation impact indicator. Post-communist countries, with the exception of a few scientific disciplines mainly in Slovenia, still lag far behind their EU counterparts. It is also clear from these statistics that research in post-communist countries tends to focus relatively more on quantity rather than quality. The relative publication performance of post-communist countries in the WoS is particularly weak in social and medical sciences and relatively strongest in natural sciences and engineering.

Figure 1: A Web-of-Science Comparison of Publication Quantity and Quality



Source: Jurajda et al. (2017).

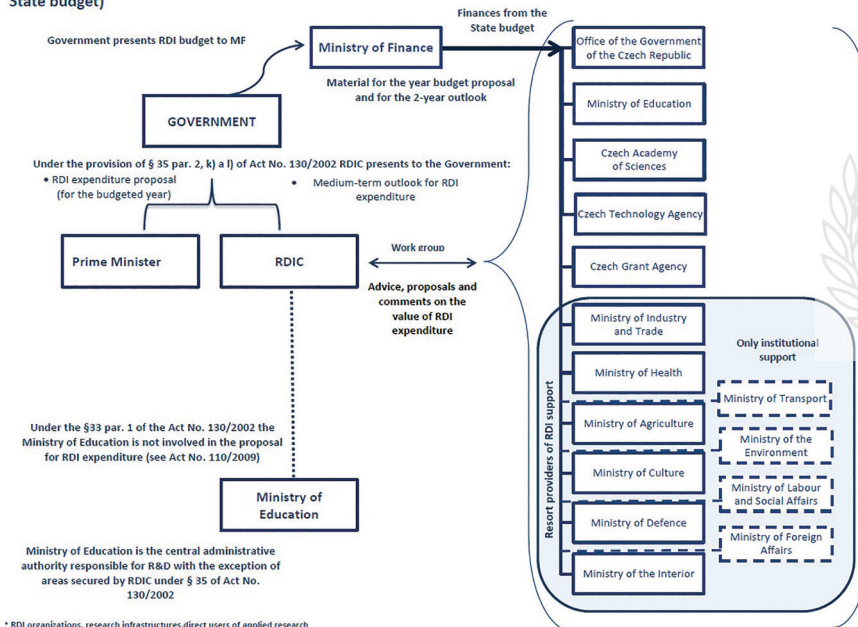
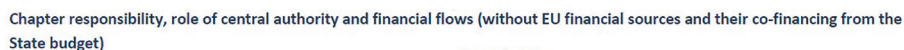
Science Funding in the Czech Republic: Basic Facts

How is the administration of Czech science organized? There is no ministry of research, and the entire Research and Development and Innovation (R&D&I) sector is coordinated by a hybrid executive-strategic R&D&I Council (RDIC) headed by the Prime Minister, an advisory body to the government, which proposes budgets and policies, including the distribution of funding to so-called ‘providers’ of funding.¹ Within basic research, the key

¹ The country has a number of strategic plans and policy documents, where the National Economic Strategy embeds the 2030 Innovation Strategy, which subsumes the National R&D&I Policy, which in turn subsumes the National Research and Innovation Strategy for Smart Specialization (RIS3) Strategy.

providers of institutional funding (block grants) are the Ministry of Education (which provides funding to universities) and the Academy of Sciences (which covers its institutes). The institutional funding sector has been subject to two funding reforms, which are discussed below. Among key providers of competitive program funding are the Czech Science Foundation (“Grant Agency of the Czech Republic”), which covers basic science, and the Technological Agency, which covers applied research and innovation. There are competitive funding programs provided also through several ministries (of industry, health, culture, etc.). Finally, there is a set of applied research institutes at specific ministries, e.g. a geology institute at the Ministry of Environment.

Figure 2: The Public Administration of Czech Science



Source: The Czech RDIC

How are these funding channels evaluated? The program-based competitive funding schemes are evaluated qualitatively by their respective providers, with the evaluation reports being submitted to the RDIC, which approves of program extensions or not. Institutional funding (basic research funding of block grants of universities and research institutes) has undergone two major reforms over the last decade. These reforms have occurred on the background

of growing aggregate funding. The Czech Republic R&D expenditures (GERD) have grown substantially from 1.3% of GDP in 2010 to just under 2% in 2014, with a mild decline afterwards. This puts the Czech Republic's research funding intensity close to, but still below the EU-28 average. As we saw in Figure 1, the Czech Republic is lagging particularly far behind the EU-15 in Social Sciences. According to OECD statistics, the Czech Republic also features a particularly low share (of 6% in 2017) of GERD in Social Sciences and Humanities. The combination of low performance and funding is typical for social sciences in the region and may reflect the communist heritage.² In the Czech Republic, it is also a consequence of a recent funding and evaluation reform, which damaged social sciences for years to come.

Two Czech Reforms of Institutional Funding

Following a broadly criticized peer-review-based funding scheme operating in the 2000s, which relied on local low-quality panels for its evaluation, the Czech RDIC introduced a sweeping funding reform in 2008 based on quantitative bibliometric indicators. While other countries have experimented with formula-based funding schemes, this has been the first reform where all of institutional funding was based on formulas and not on any expert-administrator decision-making.

Since 2009, Czech academics began supplying all of their research output to a government's registry operated by the RDIC, which used the database of output to determine the distribution of about 1 billion Euro of research funding each year, roughly 50% of all research funding from the Czech government budget. The funding formula was based on the number of papers in the WoS published by a given research unit, with somewhat higher 'points' awarded to publications in journals featuring a higher Impact Factor (IF), and on the number of books and other types of publications, with each type of publication being assigned a fixed number of funding 'points'. The overall budget was then distributed in proportion to the distribution of these points to specific funding providers who used the same transparent scheme to distribute funding to universities, many of which used the same funding formulas to allocate funding to faculties and even individual departments. Crucially, the 'points' per type of publication were the same across all disciplines and fields of science.

² For a related discussion see Kovač et al. (2018).

Within three years, this system has achieved massive reallocation of funding from low-output to high-output workplaces and from fields of science where even the best scientists in the world publish only few papers a year (mathematics, economics) to fields of science where a typical scientist publishes several papers a year (e.g., chemical engineering). Some natural science faculties increased their overall funding multiple-fold within a short time period, to a significant extent at the expense of social science and humanities.

The speed of reallocation and the arbitrary nature of cross-discipline redistribution of scientific funding were so dramatic that, by 2013, a modified version of the funding scheme was introduced, which fixed aggregate budget shares of major fields of science (for example, social sciences) and led to massive rat-race competition within each field-specific budget based on the production of high quantity of output. While higher-IF output did receive more ‘points’ than lower quality output, this gap in ‘point’ reward was dwarfed by the gap in costs of production. Given how easy it was to generate new WoS journals and to increase their IF,³ mass production of low quality output was the only way workplaces could keep their share of the field-specific budget. The rat race of mass production did not allow for funding increases for workplaces that focused on increasing the quality of their output. The situation was particularly dire in Social Sciences due to two factors: First, the overall budget of Social Sciences shrank due to the arbitrary reallocation of funding in the first three years of the reform, which intensified the rat race competition for funding. Second, many researchers in Social Sciences have never published in high-quality journals before the reform, which meant that mass production of low quality publications was the only way to compete for funding.⁴

The reform resulted in major reallocation towards higher-quality workplaces in natural sciences and in mass over-production of research output of low or zero scientific value in many fields - output that would not have been produced had it not been for the powerful transparent incentives set up by formulaic funding. In WoS Categories where such opportunistic behavior has been particularly strong, including several disciplines in technical fields and in social sciences, the Czech Republic now produces as many low-quality

³ Three practices were chiefly used to inflate journal IFs: First, WoS Proceedings corresponding to local conferences cited locally published journals heavily. Second, some journals increased self-citation rates. Third, “citation consortia” of semi-bogus journals (several economics journals in the Baltics, for example) were involved in the practice of citing each other heavily within an implicit citation consortium.

⁴ Some social science programs instructed their students to publish their master theses in near-predatory publishing houses such as Lambert Academic Publishing.

WoS articles (in the lowest 4th quartile of WoS AIS journal rankings) as would correspond to one fifth of such output of the entire EU-15. In one recent year, as a consequent of ‘point’ incentives, Czech economists produced more ISI Proceedings than all US economists; it is clear these have not systematically increased scientific knowledge in the field.

The composition of the RDIC changed around 2014 and the new Council introduced another major funding and evaluation reform in 2017. This second reform scrapped the formulaic funding scheme and replaced it with one that relies heavily on peer review, but disciplines peer review with bibliometric statistics. The new system only compares evaluated units within fields, not across fields, and it operates 5 modules (dimensions) of evaluation:

- M1 – Quality of Selected 10% of Results (peer review using external opponents)
- M2 – Research Performance (bibliometrics reviewed by discipline panels)
- M3 – Social Relevance (reported to ‘provider’)
- M4 – Viability (reported to ‘provider’)
- M5 – Strategy (reported to ‘provider’)

The first module focuses on peer-review evaluation of publications that fall outside of the WoS (where no impact measure is easily available) and on the evaluation of WoS output that may stand out in terms of its impact. The second module focuses on two statistics for each evaluated unit within a discipline: (a) its quality profile, where a high share of publications in lowest-ranked journals are deemed inappropriate, and (b) its share of high-quality output within a discipline relative to its share of funding (capacity) within a discipline. The capacity normalization step is yet to be fully implemented. Again, comparisons within M1 and M2 are performed within fields only. M1 and M2 performed annually at the national level (under the RDIC), then field-level grades are aggregated up to research-unit level (university, institute) by ‘providers’ who perform their own evaluation steps within the other three modules, which are operated in a 5-year cycle.

The WoS part of the bibliometric module M2 relies on the Article Influence Score to rank journals within 6 fields and 42 FORDs discipline categories (based on the OECD classification). WoS journals are segmented by AIS journals rankings into the 1st decile, and four quality quartiles. Panels write a discipline-wide report and then comment on the publication profiles of individual universities.

The funding reallocation achieved based on the new scheme thus far has been minimal, but the reform succeeded in that all providers now accept that

their funding decisions should reflect the ‘index funding’ principle, where research units achieving higher marks in evaluations should receive higher growth rate of their institutional funding. This is a major improvement over the previous scheme, which reallocated all funding each year (based on the rule of three) proportionally to evaluation ‘points’ corresponding to the funding formula. This resulted in large year-to-year increases as well as drops in funding, something that any scheme of science funding should avoid.

Lessons: Reform Obstacles and Traps

The above-provided subjective review of the Czech reform experience fits well into the existing scientific literature studying incentive provision in scientific knowledge production.

The printing press of quantitative publishing incentives

There is a growing literature on how scientists’ behavior is strongly affected by economic incentives (Stephan, 2012). Reformers aiming to provide incentives for high-quality science often argue that peer reviewing at journals and measures of journal quality (typically citation impact measures) provide an ideal external *quantitative* structure that can be used to allocate funding. This idea has led to several unfortunate unintended consequences in multiple countries.

Until the Czech reform of 2009 discussed above, Australia represented the only country that attempted to introduce simple bibliometric funding formulas to distribute major shares of research funding to universities. Butler (2003) reports on the effects of a simple funding formula used in Australia to fund university research, where no attempt was made to account for the quality of publication output. While such funding schemes were used since 1990, their budget shares were small (of the order of 5%). In 2001, however, the share of overall funding targeted to research and linked to a formula increased dramatically. Butler (2003) shows that with the introduction of the publication funding formulas, the number of publications in the bottom half of journals as ranked based on citation impact grew twice as fast as the number of articles in journals ranking in the top half, that such ‘inflation’ of low-quality research output occurred in almost all fields of science, and that it did not happen outside of the university sector in Australia, where the funding formula was not applied. While the Czech experience has thus far not been summarized in a

scientific evaluation of this type, the stylized description provided above is in line with the Australian experience.

Italy is another example of a country that introduced bibliometric-based evaluations with likely negative consequences in terms of strategic behavior of scientists. Seeber et al. (2019) show that Italian academics react to quantitative citation metrics used in promotions by quickly increasing self-citations, with the increase being strongest for social scientists. In line with this finding, Baccini et al. (2019) point out that the extent to which citations to Italian publications come from within Italy also increases after the introduction of citation-based metrics in academic evaluations and Bagues et al. (2019) find that around 5% of Italian researchers - a high share - have published in journals classified by the scholarly librarian Jeffrey Beall as 'predatory'.

Comparing apples and oranges

A particularly unfortunate feature of the Czech reform of 2008-2013 has been to award the same funding per article published in any field of science, despite the vast differences in publication practices across scientific disciplines. For similarly undesirable consequences of paying all researchers the same pay level across all disciplines, see Boyle (2008).

Inappropriate measures lead to unfortunate incentives

Whenever countries tie funding to the Impact Factor (IF) of journals where scientists publish, new journals appear that allow those scientists who cannot publish in high-quality outlets, to boost their 'accounting' output measures. Such journals can improve their IF by citations from locally organized conference proceedings, and they often do so. One way to avoid these printing-press incentives is to rely on the Article Influence Score (AIS) instead. The AIS is also an average per-article influence measure based on citations in the Web of Science (WoS), but it has several important advantages relative to the IF. It puts higher weight on citations from more prestigious journals, and it ignores citations to articles from the same journal, making it harder to manipulate. It is also important to avoid disadvantaging the humanities based on the fact that no such external quality proxy is available for books and non-WoS journals.

Another problem with restructuring funding allocation is that one wants to avoid unpredictable rapid changes in funding. A basic feature of the production

of scientific knowledge is that high-quality teams (departments, schools) take years to build, but can be disrupted rapidly by temporary absence of adequate funding. This dictates that funding schemes do not introduce large year-to-year fluctuations in funding. Similarly, while competitive (i.e., grant) funding provides incentives for quality growth, when the share of competitive funding relative to institutional (block-grant) funding is too high, this also creates too much uncertainty that curbs the ability of ambitious workplaces to maintain an upward quality trajectory. In a typical environment, this is not a binding issue since status quo is usually preserved through internal academic democratic processes, but the Czech reform of 2009/10 shows that reform efforts can go too far. This means that reallocation of funding towards higher quality workplaces or individuals ideally should come from aggregate increases of additional funding, not at the expense of their competitors within the system.

Conclusions and Recommendations

The main lessons from the Czech reform can be summarized as follows: Growth in aggregate funding opens door to reforms, but without international-standards evaluations, progress is slow or none. Any evaluation exercises are to be performed strictly within disciplines. Formulaic ‘transparent’ incentives can backfire, making reliance on peer review critical, but this raises another key issue of who sets the composition of evaluation panels. International advisory boards can be helpful.⁵

One can formulate two recommendations based on these lessons: (i) Attempt to make as much of a commitment as possible to external non-formulaic evaluation benchmarks such as the number of awarded ERC grants by field. If local efforts do not make a difference on such external-benchmark front over the medium term, they are ultimately to be judged as not successful. (ii) Whenever there is any opportunity for public funding growth, divide funding into a main stable part, which is available to the existing scientific community with high level of predictability and continuity, and a structurally separate funding tier focused only on high-quality performance. This second tier can ideally rely in its funding allocation on evaluation and advisory

⁵ In the context of the Czech Republic, the Czech Science Foundation sub-contracts some of its evaluations to the ESF – Science Connect. The RDIC has an International Advisory Board, which includes top-level Czech scientists working abroad. All Czech universities and some institutes of the Czech Academy of Sciences now have an International Advisory Board. Several Nobel laureates are now involved in some of these boards.

boards composed of scientists based at world-leading research institutions, including those from the diaspora.

A highly ambitious project within this second tier would be to establish a large excellence center such as the Austrian IST (see <https://ist.ac.at/en/home/>)⁶. More realistic perhaps is the establishment of person-centered large-grant programs (similar to the ERC competition but at a lower level of expectations and funding, examples include the EXPRO grants of the Czech Science Foundation, the *Lumina Quaeruntur* grants of the Czech Academy of Sciences, or the Dioscuri centers of the MPG network) as well as workplace-based funding schemes. One example, where support for internationally competitive undertakings is desirable at the workplace level (as opposed to person level) is doctoral studies. There are good-practice examples to follow when designing a competition for the best doctoral program in the nation in every field and supporting that program (and its faculty) generously (e.g., the LC centers of the Czech Ministry of Education in the 2000s or the large German *Exzellenzinitiative*). If such efforts prove to be too ambitious, any national evaluation that provides higher growth rate of funding to best-evaluated institutions based on strictly within-field comparisons, is likely to gradually make a difference as well.

References

- Butler, L. (2003) “Modifying publication practices in response to funding formulas,” *Research Evaluation*, 12 (1), 39–46.
- Baccini, A., De Nicolao, G., and E. Petrovich (2019) “Citation gaming induced by bibliometric evaluation: A country-level comparative analysis,” *PLOS ONE*, 14(9).
- Bagues, M., Sylos-Labini, M., and N. Zinovyeva (2019) “A Walk on the Wild Side: ‘Predatory’ Journals and Information Asymmetries in Scientific Evaluations,” *Research Policy*, 48(2): 462–477.
- Boyle, G. (2008) “Pay Peanuts and Get Monkeys? Evidence from Academia,” *The B.E. Journal of Economic Analysis & Policy*: 8(1), Article 21.
- Jurajda, Š., Kozubek, S., München, D., and S. Škoda (2017) “Scientific publication performance in postcommunist countries: still lagging far behind,” *Scientometrics* 112, 315–328.
- Kovač, D., Scrbec, N., and B. Podobnik (2018) “Does it Pay Off to Research Economics—A Tale of Citation, Knowledge and Economic growth in Transition Countries,” *Physica A: Statistical Mechanics and its Applications*, 505(1): 293–305.

⁶ This may not require new buildings as it can correspond to a distributed center with its own governing board. More generally, while investment into buildings and laboratories (equipment) is often very important, the critical aspect of scientific progress is human capital and the ability to attract and keep promising researchers who then can help generate their equipment funding.

- Seeber, M., Cattaneo, M., Meoli, M., and P. Malighetti (2019) “Self-citations as Strategic Response to the Use of Metrics for Career Decisions,” *Research Policy*, 48(2): 478–491.
- Stephan, P. (2012) *How Economics Shapes Science*, Harvard University Press.

