



## Research article

# How complex international partnerships shape domestic research clusters: Difference-in-difference network formation and research re-orientation in the MIT Portugal Program



Mackenzie D. Hird<sup>a</sup>, Sebastian M. Pfotenhauer<sup>a,b,c,\*</sup>

<sup>a</sup> Institute for Data, Systems, and Society, Massachusetts Institute of Technology, 77 Massachusetts Ave, E40-206, Cambridge, MA, 02139, United States

<sup>b</sup> Munich Center for Technology in Society (MCTS), Technische Universität München, Arcisstr. 21, 80333 Munich, Germany

<sup>c</sup> TUM School of Management, Technische Universität München, Arcisstr. 21, 80333 Munich, Germany

## ARTICLE INFO

## Article history:

Received 21 December 2015

Received in revised form 1 June 2016

Accepted 30 October 2016

Available online 9 December 2016

## Keywords:

Research policy

Innovation policy

Social networks

Collaboration

Program evaluation

International partnerships

Maps of science

Difference in difference

System architecture

## ABSTRACT

This paper proposes a novel mixed-method approach to study the impact of complex international capacity-building partnerships as an emerging policy tool at the crossroads of four major research policy trends – university-centrism, collaboration, internationalization, and growing structural complexity. We combine bibliometric network analysis with difference-in-difference program evaluation, statistical matching techniques, and system architecture analysis to evaluate complex research partnerships more adequately ‘in their own terms.’ We apply our method to one national “flagship” policy initiative – the MIT Portugal Program – where we compare program participants to a carefully assembled peer group of non-participant Portuguese researchers to assess the impact of MIT-Portugal with regard to idiosyncratic, more structurally oriented, and arguably less conventional program goals. As part of this methodological approach, we propose difference-in-differences Content Overlay Maps (“maps of science”) as a means to evaluate how program participants change their research focus over time relative to their national peers. These findings are complemented by an analysis of the collaborative network of participants and their institutions, as well as more traditional forms of impact assessment. Our findings indicate that complex international capacity-building partnerships can have a significant impact on the ‘hosting’ country in terms of cluster formation and research re-orientation. Moreover, they suggest that our mixed-method approach provides a valuable tool for evaluating complex capacity-building initiatives in ways that do justice to their one-of-a-kind architectures and goals. Future research should aim to study more closely the relationship between different program architectures and program impacts, and combine our largely quantitative approach with ongoing qualitative and interpretive policy analysis.

© 2016 Published by Elsevier B.V.

## 1. Introduction

Over the last few decades, policy initiatives in research and innovation have notably grown in complexity. Governments around the globe are pursuing research and innovation policy increasingly as systemic phenomena, which frequently implies addressing multiple parallel objectives and using a variety of simultaneous mechanisms in a way that reflects the specific needs and social, political, and institutional contexts of their countries (Archibugi and Lundvall, 2001a; Edquist, 2006; Jasanoff, 2005; Jasanoff and Kim, 2015; Kuhlmann et al., 2012; Lundvall, 1992; Nelson, 1993;

OECD, 2010a; Pfotenhauer et al., 2016; Pfotenhauer, 2016; Smits and Kuhlmann, 2004; van Geenhuizen et al., 2005). This surge in complexity is further exacerbated by an increasing policy focus on collaboration in science, innovation, and higher education, oftentimes with a distinctly international orientation (Altbach et al., 2010; Altbach and Knight, 2007; Bhandari and Blumenthal, 2011; Bozeman et al., 2012; Carayannis and Laget, 2004; Freeman, 2014a; Georgiou, 1998a; Hunt and Gauthier-Loiselle, 2008a; Lee and Bozeman, 2005a,b; Luukkonen et al., 1992a; Shrum et al., 2007; Stromquist and Monkman, 2000; Subramanyam, 1983a,b; Wagner, 2005a; Wuchty et al., 2007a). In many of these ‘complex’ research and innovation policy initiatives, universities play a central role. This prominence of universities, particularly for questions of capacity-building, is owed at least in part to their assumed ability to simultaneously address questions of human resource formation, research, and innovation – a role sometimes

\* Corresponding author.

E-mail addresses: [mhird@mit.edu](mailto:mhird@mit.edu) (M.D. Hird), [sebastian.pfotenhauer@tum.de](mailto:sebastian.pfotenhauer@tum.de) (S.M. Pfotenhauer).

characterized as the “knowledge triangle” see also (Conceição and Heitor, 2001a,b; Cowan and Zinovyeva, 2013; EC, 2012; Etkowitz, 2008, 2003; Mansfield and Lee, 1996; Mansfield, 1991; Mowery, 2004; Pfotenhauer et al., 2013; Pfotenhauer, 2017; Salter and Martin, 2001; Shane, 2004; Slaughter and Leslie, 1997; Thorp and Goldstein, 2010; Winckler, 2010; Youtie and Shapira, 2008).

Given this consistent growth in research policy complexity and associated policy considerations around growing collaboration, internationalization, and differentiation, how do we assess the performance of policy initiatives in a way that does justice to their increasingly unique designs and purposes? Policy researchers and policy makers alike have recognized that one-size-fits-all approaches to evaluating research and innovation policy are ill-suited given the differences among national and regional innovation systems, and their largely idiosyncratic goals (Magro and Wilson, 2013; Rasmussen, 2008; Shrum et al., 2007; Tödting and Tripl, 2005). There is a growing awareness that complex initiatives need to be evaluated ‘on their own terms,’ and that the measurement of innovation and research performance has to follow suit in terms of observed complexity (OECD, 2010b). In this context, both qualitative case-study research and quantitative social network analysis (particularly bibliometric analysis) have offered compelling ways forward – the former with a focus on “structural and cultural aspects;” the latter through the construction of complex “indicators of trends and processes” (Shrum et al., 2007). In this paper, our primary focus rests with the latter, although we draw heavily and directly on complementary qualitative research (Pfotenhauer et al., 2016; Pfotenhauer, 2016; Pfotenhauer and Jasanoff, forthcoming-a, forthcoming-b).

Social network analysis is a tool for mapping topographic information such as position or strength of connection among a set of people within a social network. For example, bibliometric analysis can be used to trace the production, circulation, impact, and epistemic differentiation of knowledge by assessing available network information (e.g. citations or aggregate measures such as impact factors) for research publications, patents, or other bibliometric sources (Bozeman et al., 2012; Katz and Martin, 1997a,b; Melin, 2000; Ponomarev and Boardman, 2010). Yet, bibliometric network analysis poses several conceptual, empirical, and interpretative challenges, including questions of data availability (“looking where the light is”), its utility for program design and evaluation, or its normative implications (the “indicator game”). What is more, experience shows that many program evaluators tend to apply bibliometric analysis in under-complex, and at times negligent, ways. The need for a more sophisticated, judicious, and humble use of network analysis has been prominently articulated in the recent Leiden Manifesto (Hicks et al., 2015), which picks up key concerns about how to evaluate initiatives, organizations, and people ‘in their own terms’ and in conjunction with a range of other indicators, preferably including some qualitative methods. The Leiden Manifesto builds on decades’ worth of experience on the use and misuse of bibliometric analysis, and the potential and risks associated with it.

In this paper, we attempt to make a modest step in this proposed direction of a more sophisticated use of quantitative analysis through close attention to context, intent, and design. We propose a mixed-method approach that uses network analysis in combination with difference-in-difference evaluation, statistical matching techniques, and insights from system architecture thinking to study the characteristic impact of a complex international, collaborative policy initiative in a way that aims to do justice to its program architecture, idiosyncratic program goals, and local (peer) environment. We thereby aim to propose an avenue for evaluation that recognizes, to the extent possible, the ‘one-off’ character of the initiative – much in line with the principles and aspirations of the Leiden Manifesto, as well as wider calls to focus on the broader impacts, not

mere outcomes, of research collaboration (Bozeman et al., 2012). While our goal is to eventually develop a comparative evaluation methodology that can be applied across a range of complex research policy initiatives and with particular attention to qualitative and interpretive analysis, our objective for this paper is primarily to introduce, and demonstrate the utility of, our mixed-method quantitative approach using one concrete example. We study the impact of one complex international research policy initiative – the *MIT-Portugal Program*, or *MPP* for short – with a particular view towards less conventional, more structurally oriented policy goals alongside standard measures of research productivity, quality, and visibility. These structural goals include the empowerment of young researchers; the stimulation of intra-Portuguese institutional collaboration to create critical-mass distributed research clusters; and purposeful redirection of research activities in some of the leading groups of the country towards designated priority areas.

We begin by briefly introducing the MIT-Portugal Program and derive a set of five program goals that we wish to probe, including some that are specific to the program architecture. We then introduce some of the theoretical background that we draw on for our analysis and explain how we will use it in our research design. In particular, we focus on bibliometric social network analysis (which allows us to map collaboration dynamics among researchers and institutions), difference-in-differences (DID) analysis (which contrasts the performance of MPP program participants with that of similar peer group in Portugal), and two statistical matching techniques – Propensity Score Matching and Coarse Exact Matching (which allows us to augment the difference-in-difference approach to reduce potential selection biases in the sampling). DID approaches in conjunction with matching techniques have hitherto rarely been used in a bibliometric program evaluation context. Subsequently, we provide an overview of our data acquisition and preparation process before proceeding to the construction of the researcher networks. In the analysis section, we will then assess the efficacy of MPP with regard to the five goals – and with it the efficacy of our method. For the last part of the analysis, we further introduce Content Overlay Maps to trace for the first time changes in research focus of participating researchers in a DID configuration. We conclude by charting a comparative research agenda for understanding differences and similarities among different types of complex research policy initiatives.

## 2. The MIT-Portugal Program (MPP): case study synopsis and determination of evaluation goals

Launched in 2006, MPP is a large-scale consortium-type partnership involving five Portuguese universities, MIT, various national research laboratories, and industry partners “to strengthen the country’s knowledge base and international competitiveness through a strategic investment,” intended as a “high-profile effort to demonstrate that an investment in science, technology and higher education can have a positive, lasting impact on the economy by addressing key societal issues through quality education and research” (MIT, 2005). Sponsored by the Portuguese government, MPP facilitates a range of collaborative research activities, ranging from small one-time seed projects to large scale, multi-year “test-beds.” For example, in the ‘Green Island’ project based on the Azores, researchers perform work on sustainable energy systems and integrated electro-mobility solutions under the conditions of a limited-sized island “living laboratory” with a view towards testing technology at scale. Generally, MPP’s research and education programs are focused around four priority areas – Sustainable Energy Systems, Bioengineering Systems, Transportation Systems, and Engineering Design and Advanced Manufacturing – determined by the Government of Portugal in an assessment and

match-making process leading up to the collaboration. These priority areas were deliberately chosen either because the Portuguese institutions already had considerable strengths in these domains (e.g. transportation, sustainable energy) or because the government saw it as an opportunity to build new capacity in research domain with little precedence in Portugal (e.g. stem cell bioengineering). In both cases, they were chosen as areas for partnership with the goal to shape and strengthen future research directions at Portugal's leading institutions through an exchange with MIT (Pfotenhauer et al., 2013).

MPP's conception was further driven by a range of institutional and national reform goals. For example, MPP was explicitly designed to stimulate intra-Portuguese collaboration to overcome traditional patterns of institutional insulation and competition, and foster the formation of critical mass in the afore-mentioned priority areas. While Portugal has always had pockets of research excellence, these pockets did hitherto not suffice to propel the country to global research impact and visibility (Heitor and Bravo, 2010; Horta, 2008). In response to this perceived challenge, five universities were enlisted in a countrywide research and education consortium structure, where program internal-research calls were specifically designed such that always at least two different Portuguese universities collaborated on any given project. Here, MIT's role was primarily conceived as that of an incentive for these Portuguese universities to collaborate – a “kind of glue,” as one senior government official put it (Pfotenhauer et al., 2013) – which was deemed more important than the mere collaboration with MIT itself. The formation of such national clusters was furthermore supported by the creation of 7 new national graduate programs, housed simultaneously at multiple Portuguese universities and awarding degrees jointly by several member institutions. This program began with a number of semi-open scouting workshops, with leading researchers from the Portuguese universities and MIT being invited. The first partnerships and the definition of the focus area were based on discussions at these workshops where there was deemed to be a good fit in terms of research interests and previous expertise between the Portuguese and MIT researchers. This initial consortium was then opened through “open research calls” for project proposals inside the four focus areas. Over the course of the program, many students and faculty from Portugal spend extended research periods at MIT, creating a close-knit visitors community, while at the same time researchers and students from MIT spend time in Portugal for teaching or research collaborations.

MPP is what one could call a national ‘flagship’ policy initiative. In the trajectory of the country's university system, it marks a historical turning from a period of quasi-egalitarian expansion of the higher education system to one of excellence-formation and international competitiveness with greater focus on research. Interestingly, MPP's architecture and goals are a direct result of the socio-political and historical context: After the half-century of dictatorship until 1974, a newly democratic country inherited an elitist and politicized university system paired with the loss of its economic basis and a relatively low trust in governmental leadership and public institutions (Amaral and Magalhães, 2005; OECD, 2007). Partly in compensation of this heritage, the subsequent decades saw a strong emphasis on public access and equality, where research funds were distributed more or less uniformly throughout the system – at the expense of critical research mass and institutional excellence. The consortium structure of MPP and the formation of distributed national research clusters reflect both this equitist tradition that rendered a partnership between MIT and a single Portuguese university politically unfeasible, as well as fact that no single university had the cumulative research capacity to serve as an equal partner in a collaboration with MIT. This is consistent with research that emphasizes the structural barriers to, and lack of pay-off for, asymmetric partnerships between partners of

highly unequal “levels of development” (Toivanen and Ponomariov, 2011; Ynalvez and Shrum, 2011). In fact, MPP can be interpreted as a part of a wider “catching up” narrative that in Portugal was imagined to require external assistance (Pfotenhauer and Jasanoff, forthcoming-a).

The above discussion illustrates for the case of Portugal how the goals, rationales, and feasible mechanisms of complex, large-scale research policy initiatives tend to be highly system-specific, resulting in one-of-a-kind architectures and idiosyncratic types of impact that one would like to achieve (and measure). Based on an in-depth analysis of the MPP system architecture (Pfotenhauer et al., 2016), we select a subset of five MPP goals as defined by the government and institutional stakeholders: (1) Conduct high-impact research and increase visibility; (2) Empower in particular a new generation of researchers; (3) Strengthen intra-Portugal research collaboration to achieve critical research mass and overcome patterns of inter-institutional isolation; (4) Strengthen collaboration with an MIT; (5) Shape research directions according to national strategic objectives. While the first and the fourth goal are what one could call “generic” program goals, the remaining three goals are more structural and closely tied to program architecture and socio-political factors in Portugal. We will discuss the operationalization of these program goals in the following section. For more information on MPP, see (Pfotenhauer et al., 2016, 2013; Pfotenhauer and Jasanoff, forthcoming-a, forthcoming-b).

In its scale, complexity, and system-specificity, MPP stands in for a broader class of recent flagship research policy initiatives which governments have mobilized to strengthen national or regional competitiveness, build capacity and competitive advantages in scientific or technological priority areas, foster the transition to knowledge-based economies. As mentioned at the outset and illustrated for MPP, these initiatives tend to sit in the intersection of several current policy trends: increasing university-centrism, increasing collaboration, increasing internationalization, and increasing complexity. For example, Portugal itself has launched four more large-scale international partnerships of varying design besides MPP, including collaborative initiatives with Carnegie Mellon University, the University of Texas at Austin, Harvard Medical School, and the German Fraunhofer Society. Singapore's Campus for Research Excellence and Technological Enterprise (CREATE) hosts research operations of as many of 10 international partners, including ETH Zurich, Cambridge University, University of California, Berkeley, Technion, and MIT. In the Middle East, several dedicated innovation cities are under construction (including Abu Dhabi's Masdar City or Qatar's Education City), which draw heavily on foreign partners and have universities at their heart. Elsewhere, we call this new generation of research policy initiatives in the intersection of these four trends Complex International Science, Technology, and Innovation Partnerships (CISTIPS) (Pfotenhauer et al., 2016). Here, we just note that the methodological tools and approaches developed in this paper speak to broader evaluation challenges to assess the increasingly complex, collaborative flagship initiatives ‘on their own terms.’

We wish to note that both authors of this study have been (partly) affiliated with MPP, brought in as “embedded” researchers by the program leadership specifically for the purpose of studying the impact of the program under conditions of structural complexity. Embedded social science research has become increasingly common for large-scale (or even small-scale) research initiatives. While one has to be cognizant of the potential risk for bias, it allowed both authors to complement quantitative analysis with interviews, periods of participant observation, and direct access to relevant actors and documents.

### 3. Theoretical background

#### 3.1. Collaboration trends and bibliometric analysis in research policy

International linkages and collaboration have become a key focus of research policy and national capacity-building efforts in science, technology, and innovation. Indeed, we are witnessing a veritable surge in collaborations across fields and institutions (Bozeman et al., 2013; Georghiou, 1998b; Katz and Martin, 1997a,b; Luukkonen et al., 1992b; Melin, 2000; Vinkler, 1993; Wagner, 2005b; Wuchty et al., 2007b). As exemplified by the case of MIT-Portugal, part of the appeal to policy-makers – such as the interest in targeted capability building or external facilitation of institutional reform – builds on well-known rationales from (inter)national development, which has focused on questions of technology transfer, capability building, and knowledge circulation as key ingredients for “catching up” to the innovation frontier (Amsden, 2001; Bozeman, 2000; Cohen and Levinthal, 1990; Fransman et al., 1984; Keller, 1996; Kim, 1997; Lee and Lim, 2001; Reddy and Zhao, 1990; Wei, 1995). More recently, international linkages and collaborations gained further importance in the context of an emerging globalized learning economy, in which rapid knowledge diffusion and updating, access to key nodes in knowledge networks, institutional diversity, and global interconnectedness are increasingly replacing classical growth factors such as the accumulation of capital and labor (Archibugi and Lundvall, 2001b; Conceição and Heitor, 2001b; Freeman, 2001; Llerena and Matt, 2005). Research collaborations have further been linked to positive effects in terms of scientific as well as broader economic productivity (Dietz and Bozeman, 2005a,b; Lee and Bozeman, 2005a,b; Subramanyam, 1983a,b; Wuchty et al., 2007b). What is more, internationalization and collaboration have also been tied to international labor mobility, especially with regard to university students and researchers (Bound et al., 2009; Freeman, 2014b, 2005; Hunt and Gauthier-Loiselle, 2008b). For example, is an open secret that the US has benefitted from the attraction and retention of highly talented students from around the globe (Finn, 2012).

Bibliometric analysis has both made this trend visible and served as an instrument to study it. Early work (Martin and Irvine, 1983; Moed et al., 1985) demonstrated how bibliometric data such as scientific articles, conference papers, and patents could serve as “monitoring devices.” However, even in these first applications of bibliometric data, researchers were cognizant for the need of multiple metrics and careful analysis, going so far as to recommend that such analysis should not be absolute, but instead comparative over time or between programs. Georghiou and Roessner (2000) found that evaluations of technology and research policies that solely focused on the “production function” of these programs were oftentimes not useful for policymakers or program managers because of uncontrollable contextual factors, externalities, spillovers and time lags. Over time, research attention has shifted increasingly towards including networking and other social aspects as explanatory components of research behavior and policy success. Sometimes described as the “contextual complement of human capital” (Burt, 2000), social capital is now generally accepted as a key ingredient for the exchange of information. However, social capital has been notoriously difficult to measure. Bibliometric analysis has generally taken the approach of measuring social capital through the nature and strength of collaborative ties established between group members in publications; other forms of social capital such as trust and broader visibility are much harder to capture (Shrum et al., 2001).

Reagans and Zuckerman (2001) first demonstrated the value of collaborative network structure to the output of science when they recast the diversity-performance debate (where the diver-

sity of research teams either increases ideation, and thus increases productivity, or increases collaborations costs, and thus decreases productivity) as two separate but complementary structural principles of a group’s collaborative network. Indeed, collaborative structure is critical both to the generation of ideas, where a low density of network connections may lead to unseen options and creativity, and the circulation, adoption, and research on those ideas, where high density of connections leads to acceptance and further study of an idea (Burt, 2004). As a result, research collaboration has been shown to have positive effects on scientific, as well as broader economic, productivity (Dietz and Bozeman, 2005a,b; Lee and Bozeman, 2005a,b; Subramanyam, 1983a,b; Wuchty et al., 2007a,b). A growing base of work on social network analysis has analyzed how the network structure of social capital relates to how knowledge is produced (Hansen, 1999; Reagans and McEvily, 2003; Reagans and Zuckerman, 2001), how knowledge is diffused to others (Centola and Macy, 2007; Granovetter, 1978), how individuals and ideas are influential (Bakshy et al., 2011; Kempe et al., 2003), and how groups and teams can coordinate and cooperate (Kearns et al., 2014; Lazer and Friedman, 2007). Murray (2004) found that scientists draw deeply on their collaborative network in the course of producing new knowledge. Besides their immediate environments in local laboratories, universities, or firms, scientists also rely on a broader network of social connections that may cut across societies in the form of an “invisible college,” both inside and outside their academic focus area, to generate and spread ideas (Price and Beaver, 1966; Wagner, 2008). In this way, structural differences of researchers’ social networks or their position within larger networks can have important implications for their future productivity (Casper and Murray, 2005). In short, we now know that the relationship between collaboration and any desired outcome is not a simple linear, but tends to be complex and contingent on many social factors, including network structure. Social network analysis is hence both a powerful tool for science policy researchers to understand how policies unfold, but also a challenge in bringing the tools to bear in a sufficiently complex way to capture salient idiosyncrasies.

In science, many different types of collaboration exist, ranging from close collaboration on conducting experiments to informal conversations and exchange of ideas. Yet, most studies of scientific collaboration focus only on publication co-authorship as a metric of collaboration (Bozeman et al., 2012). The reason, as Katz and Martin (1997a,b) point out, is that co-authorship data is easily available, verifiable and measured. There has been work where scientists report their research collaborations in surveys rather than using co-authorship data (see, for example, Melin, 2000 or more recently Bozeman and Gaughan, 2011) and while these types of collaboration surveys have a number of advantages, one should also be cognizant of the potential drawbacks of self-reported collaboration data. Further, owing to the prominence of co-authorship data, most studies of scientific collaboration have focused on individual-level research collaborations (Bozeman et al., 2012). Much less attention has been paid to institutional-level research collaborations in purely academic and research settings, although notable exceptions exist (though primarily in qualitative case studies) (Bailetti et al., 1994; Balmer, 1996; Cetina, 1999; Chompalov et al., 2002; Chompalov and Shrum, 1999; Crawford et al., 2010; Galison, 1994; Shrum et al., 2007; Zabusky, 1995). Those papers that do focus on institutions are primarily concerned with university-industry links or government-industry-university interactions (Carayannis et al., 2000; Eom and Lee, 2010; Etkowitz, 2002; Hagedoorn et al., 2000; Powell et al., 1996; Poyago-Theotoky et al., 2002), institutionalization and management of multidisciplinary, mission-oriented research collaborations (Bammer, 2008; Corley et al., 2006), or collaborative settings in cross-border provision of higher educa-

tion (Altbach and Knight, 2007; Knight, 2004; Verbik and Merkley, 2006).

Research into the nature of scientific production and collaboration has been accompanied by a growing interest in mapping research and development activity across sectors, motivating methodological developments such as heat maps of “hot research areas,” content overlay maps, or economic diversity maps based on export goods (Hidalgo et al., 2007; Rafols et al., 2010; Saka et al., 2010). On the other hand, scientific and technological fields have proven both notoriously difficult to classify and hardwired into institutional structures and publication outlets. Some emerging fields or disciplines have become structured and codified in their own right (such as biochemistry) while others remain at the fuzzy intersection between traditional disciplines (such as energy research). All this contributes to increasing measurement complexity of present-day research initiatives in terms of goals and mechanisms to be considered. More sophisticated content overlay maps may shed light on interesting facets of interdisciplinarity (for example, Grauwil et al., 2012); yet their uptake and application remains limited, with only one example of use at the program level (Garner et al., 2012).

### 3.2. Difference-in-differences approach

In our research design, we apply a difference-in-differences (DID) approach to map the impact of the program in comparison to a peer group of researchers in Portugal. Widely used in econometric and quantitative sociological studies (Abadie, 2005; Ashenfelter and Card, 1984; Card and Krueger, 1994), DID is based on a quasi-experimental research design that compares the mean value of a dependent variable over time between a treatment and control group (the former being subject to an intervention, the latter not) to reveal potential causal relationships and the determine effect size of the intervention. Through this cross-group comparison, DID generally eliminates the problems of selection bias associated with single-population time series or static comparisons. However, DID relies upon one critical assumption: besides questions of data availability, the treatment and control groups are assumed to be sufficiently similar to be comparable. That is, they have similar characteristics before the treatment and would continue to display the same characteristics if either both or none had been given the treatment. DID also presumes that the composition of individuals of both groups remains unchanged over time.

### 3.3. Matching techniques: accounting for potential biases

DID is still susceptible to certain biases. For example, given that DID compares mean outcomes of populations at different points in time without account for the variance in outcomes, it is possible large changes in a small sub-sample create a sample-wide effect and can mask that the majority of the sample might have changed very little (or not at all). In our case of bibliometric analysis, this would mean that changes in the output of the one or two most productive researchers will strongly affect the results and suggest a program-wide success where the true impact of the program has been extremely localized and limited. Further, relying on means and standard deviations does not account for the specific distributions within each sample population, which could in theory lead to comparison of very differently composed samples and introduce an invisible bias. Finally, because selection into the treatment group was not random, there may be significant selection bias: while our treatment and control groups appear similar before the treatment in terms of qualities we can measure, those selected may simply be better across non-observed metrics, which will lead to greater success in the future.

To remove this potential selection bias and increase our confidence in the DID analysis, one can use covariate matching techniques where instead of comparing the mean outcomes (as above) one matches up *sub-samples* of the treated and control population (either smaller groups or individuals) in terms of their pre-treatment covariates and then calculate the total effect of the treatment as the *average treatment effect between these matched sub-samples*. This matching process is primarily focused on reducing the difference between the covariate distributions for the treatment and control groups, sometimes called “imbalance.” Matching techniques add a more fine-grained comparison between the two samples analyzed via DID, ensuring similarity in terms of make-up and thus improving comparability (Rubin, 1980). On the downside, matching tends to discard a (potentially large) number of data points from the treatment group when there is no control group match present for them, which limits the overall sample size and at times the representativeness of the now-reduced treatment cohort. In practice, one has to consider these trade-offs carefully on a case-by-case basis.

Here, we use two different matching techniques with different degrees of restrictiveness *in addition to* standard (full-sample) DID, which will allow us to interpret results across three different methodologies and draw more robust conclusions. The two matching techniques we apply correspond broadly to the two main types of matching found in the literature. The first one, Propensity Score Matching (PSM), is a so-called Equal Percent Bias Reduction (EPBR) method, first coined by Rubin (1976). EPBR focuses on reducing bias across all covariates simultaneously, which is particularly important in cases where we do not know anything about the relative importance of covariates (or we don’t want to make any assumptions about their relative importance). PSM in particular – a widely used and accepted method for drawing causal estimates of effect size (Dehejia and Wahba, 2002) – calculates a single (composite) propensity score for researchers to be assigned to the treatment group based on their observable pre-treatment covariates, and then matches up researchers in the treatment and control group that have similar propensity scores. The individual propensity score is the coefficient of a probit estimate of being selected into the treatment group based on five observable pre-treatment covariates (cf. section Assembly of the dataset).

While PSM decreases the imbalance between our treatment and control groups and theoretically lowers the selection bias, the use of a propensity score can in theory lead to comparing empirically dissimilar researchers who have similar propensity scores and are matched together. For instance, two researchers could be matched with very similar average number of publications, journal impact factor and citations, differing only in their scientific domain. However, because publication patterns can be so different between fields, when compared to their within-field peers we might see one as very productive for their field and another as not very productive. To avoid this error, we turn to our second matching procedure.

Coarse exact matching (CEM), secondly, is a so-called Monotonous Imbalance Bounding (MIB) technique, which have only been introduced fairly recently. MIB methods focus on not exceeding an *ex-ante limit* on the amount of imbalance between distributions of pre-treatment covariates of the treatment and control groups. For CEM in particular, we temporarily coarsen covariates into larger bins and then sort researchers into strata where they are exactly matched for all coarsened covariates (Iacus et al., 2011, 2009), as opposed to matching researchers based on one composite score. Subsequently, we eliminate any strata that do not contain at least one treatment and control (where no matching is possible), and then match treatment and control data in each stratum. CEM (and MIB more generally) discards even more data that might otherwise be matched with other matching methods.

#### 4. Assembly of the datasets

Our MPP assessment applies the afore-described DID approach in conjunction with bibliometric analysis. We started by assembling the treatment group by selecting all (Portugal-based) researchers who were officially part of the program. We then assembled a control group by identifying all researchers in the same departments at the same Portuguese universities who were *not* affiliated with the program, dropping any who were no longer actively publishing in 2014, using affiliation with the same academic departments at the same universities as proxies for similar levels of excellence. We then manually checked both treatment and control groups to confirm that each researcher was not included in any of the other major international partnerships with foreign universities that Portugal maintains. We then custom-tailored the data set by dropping control-group researchers until our treatment and control groups had similar distributions with regard to four observable pre-treatment covariates: average number of publications (as a proxy for productivity), average impact factor (a proxy for quality), average number of citations per article (a proxy for visibility), and the number of years a researcher has been publishing (a proxy for “seniority”). We also control for academic discipline by collecting our control group from the same academic departments as the control group, and grouping these researchers into the same four focus areas as MPP-participants: Sustainable Energy Systems, Bioengineering Systems, Transportation Systems, and Engineering Design and Advanced Manufacturing. For the purpose of this study, this warrants that the only major difference between the two groups was whether researchers participated in MPP or not, thus controlling for any systematic bias in outcomes (including “hand-picked” program membership) between different universities, fields/departments, and individuals. We of course acknowledge the limits of our proxies: “High impact research” could be defined in many different ways (with differences between scientific fields), many all of them qualitative in nature and not captured by publication data. Yet, in following the goals for the partnership as expressed by the government, and recognizing that most institutions and governments today, including that of Portugal, use a combination of these metrics to monitor research activity, we decided to utilize these proxies. Further, there could be some systematic bias that the qualitative selection into the program captures when choosing between two researchers in the same departments and with the same pre-treatment covariates. There are certainly differences between researchers that will not be captured by our quantitative techniques. However, further work on limiting selection bias will have to either incorporate qualitative information or develop novel quantitative metrics or predictors of future productivity of each researcher. While both of these improvements would strengthen this method, and would be valuable follow-up work, they are outside the scope of this paper. From this process, we obtain our dataset of a total of 354 scientists: 234 researchers affiliated with MPP (treatment) and 120 non-affiliated researchers (control).

We manually collected all of these individuals’ conference publications, articles and books (or book chapters) from both the Web of Science and Scopus databases. We then combined and cross-checked these two databases, removing duplicates that appear in both. In cases of name ambiguities, we manually consulted the CV or Google Scholar curated profiles for each person to select the right publications. When name disambiguation was not possible, we removed the individuals from the sample group (leading to 12 MPP-affiliated researchers being dropped from the sample). We did not restrict our data collection to any specific time period, and therefore collect data from each researcher’s entire career. However, since MPP was launched in 2006, we consider our treatment period ranging from 2007 to 2013, the latter being a data-imposed

cut-off as the program is still ongoing. We further observe that the ‘parallel trends assumptions’ holds true for the pre-program period, whereas a significant difference in slope can be detected after 2006.

We then assemble two alternative control groups from the manually compiled control group using the afore-described statistical matching techniques PSM and CEM. For PSM, the propensity score is calculated using a probit estimate of being selected into the treatment group based on five observable pre-treatment covariates – the four publication characteristics (productivity, quality, visibility, seniority) as well as topic area (categorized by the four broad topic areas defined by the program itself). We investigated other covariates (such as h-index or cumulative measures of output) but found them to be tightly correlated with other metrics and not useful for matching. In the PSM analysis, we use a very strict caliper of 0.20, in order to draw stronger causal connections. Each treatment researcher is compared to their nearest-neighbor control researcher in propensity-score space (Cochran and Rubin, 1973). A moderate fraction of treatment and control researchers is dropped in this case, and we now compare 213 treatment researchers (92% of original) to 74 control researchers (62%). The descriptive quartile-quartile plots for this matching can be seen in Fig. 1a and b, where we see large improvements in the imbalance over our unmatched data.

The use of CEM reduces the sample even further, to 57 treatment researchers and 61 control researchers. While this discarded data makes the CEM analysis less representative of the entire program under evaluation, there is extremely little bias between the treatment and control (seen in Fig. 1c) and we can make very strong causal claims. We also include the effect sizes between our unmatched DID treatment and control for two reasons. First, while we acknowledge that this sample could be significantly biased, an unmatched DID comparison is the only method that allows for an evaluation of every researcher who receives treatment. As we move from Unmatched DID to PSM to CEM each method draws stronger causal claims than the previous method at the expense of representativeness of the evaluated group. Second, in an effort to keep with the fourth principle laid out in the Leiden Manifesto, we believe this helps “keep data collection and analytical processes open and transparent” (Hicks et al., 2015). In this way we hope to head off one possible critique of matching methods, where one could selectively tune their matching algorithm to discard data points in order to achieve a “desired” result. Indeed, in Table 1, where we summarize the means and standard deviations of each pre-treatment covariate for all three methods, while we see no statistically significant differences at  $p < 0.05$  between our unmatched treatment and control groups, we do see that across all covariates, the mean of our unmatched treatment group is larger than that of the control group. We have not included the dummy variables for the four focus areas in Table 1 for lack of an intuitive comparative meaning.

Fig. 1 shows quantile-quantile plots comparing the distributions of our four matching covariates for the treatment and control groups. Quantile-quantile plots are a less intuitive method of graphically comparing two distributions than simply comparing histograms, but are more powerful and avoid issues around the bins sizes of histograms. To construct a quantile-quantile plot data points are placed where they lie in the quantiles of both data sets (where the quantile is the fraction of the dataset that takes a lower value than the given value). The axes show the range of values for each dataset, from minimum to maximum value. For example, a point that sits in the 0.4 quantile of one data set (where 40% of the data points are lower than it) might be at the 0.3 quantile of the other data set (where only 30% of the points are lower than it). If the datasets are similar, the points will fall along the 45° line, where the quantiles of all points are the same in both datasets. We find that our control and treatment groups become indeed more similar as we move from DID to PSM to CEM, as indicated by how closely

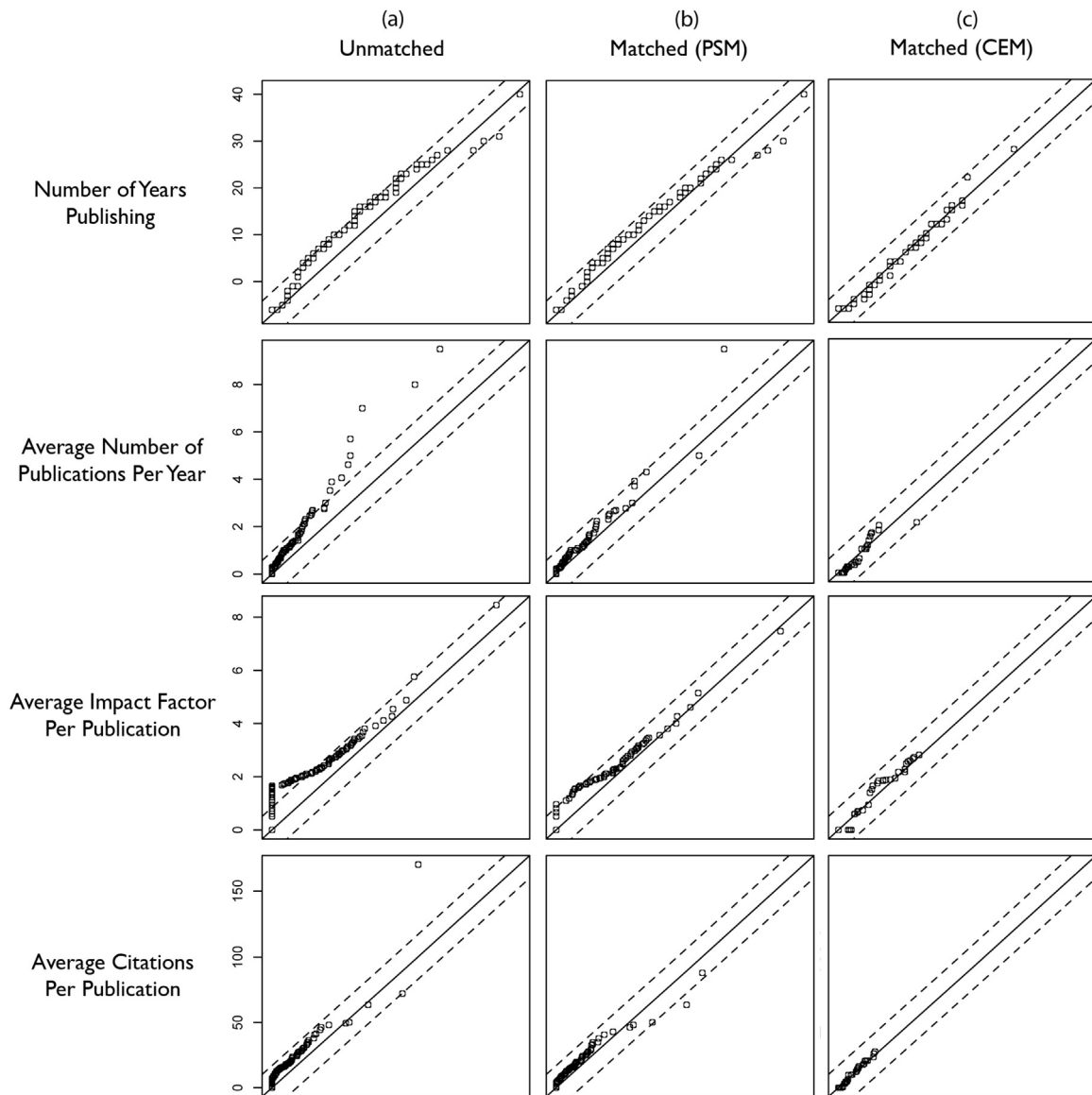


Fig. 1. Quantile-Quantile Plots of the a) unmatched data, b) matched data under PSM and c) matched data under CEM.

**Table 1**  
Mean and Standard Deviation of Pre-treatment Covariates under matching.

|                                  | Raw Treatment Mean | Raw Control Mean | Unmatched Mean Difference | Standard Deviation | PSM Mean Difference | Standard Deviation | CEM Mean Difference | Standard Deviation |
|----------------------------------|--------------------|------------------|---------------------------|--------------------|---------------------|--------------------|---------------------|--------------------|
| Productivity (articles per year) | 1.3513             | 0.7740           | 0.5773                    | 1.0855             | 0.0397              | 1.1592             | -0.0998             | 0.5907             |
| Quality (average impact factor)  | 2.1145             | 1.3910           | 0.7235                    | 1.4189             | -0.0531             | 1.4642             | -0.0326             | 1.0265             |
| Visibility (average citations)   | 18.7147            | 12.3136          | 6.4011                    | 17.3444            | 2.4406              | 15.7320            | -0.5096             | 7.9272             |
| Seniority (years publishing)     | 12.6793            | 9.0200           | 3.6593                    | 9.4964             | -1.0469             | 11.4003            | -0.3594             | 7.6416             |

points fall along the 45° line. At the same time, we see that the sample size drops and the number of outliers is drastically reduced. This validates our strategy of reporting the outcomes of all three methodologies alongside each to demonstrate both the robustness and transparency of our results. We will note that there are still two main selection biases that may have small effects on our treatment group. First, English Language proficiency may have been a large barrier to full participation and while there was no explicit test or requirement, it may have affected self-selection into the applicant

pool of MPP. Secondly, due to this self-selection into the MPP applicant pool, Treatment researchers may have been more interested in moving their research into one of the MPP focus areas, leading to a sample more likely to move research interests than our control group.

We now construct an egocentric network for our treatment and control groups for each of our three groups (unmatched, PSM and CEM) cumulative over the entire time period of interest. Links are determined by co-publication in journals articles, conference pro-

**Table 2**  
Descriptive Network Statistics.

|                            | Treatment (MPP) | Control |
|----------------------------|-----------------|---------|
| Number of Authors (nodes)  | 21815           | 7710    |
| Clustering Coefficient     | 0.831           | 0.821   |
| Connected Components       | 51              | 20      |
| Network Diameter           | 10              | 9       |
| Network Radius             | 1               | 1       |
| Characteristic Path Length | 4.226           | 4.175   |
| Avg. Number of Neighbors   | 11.160          | 10.116  |
| Network Density            | 0.001           | 0.001   |
| Network Heterogeneity      | 2.187           | 1.767   |

ceedings and books. Accounting for co-authors of members of the treatment and control group, our treatment network contained 21815 unique authors with 89128 collaborative links between them, and our control network contained 7710 unique authors with 38999 collaborative links. While the size of these networks are quite different and reflect effects from both the pre-treatment period (where the collaboration networks are very similar) and the treatment period, one can observe a strong similarity between the descriptive statistics of network structure, summarized in Table 2. In particular, we see very similar clustering coefficients (the probability that a node's collaborators are also collaborators with each other) and average path lengths. Further, these high clustering and low average path lengths are indicative of the small world networks we typically see in this type of scientific co-authorship data (Newman, 2001; Watts and Strogatz, 1998).

## 5. Results

Here we report the results of our analysis, organized by the five MPP policy goals we are seeking to evaluate. We summarize our findings in Table 3. The first column of this table shows the baseline performance of our control group during the treatment period in each dependent variable, which will give the reader a sense of magnitude for the effect size of the program. We then report the results from each of our three analysis techniques, both the unmatched DID as well as the two different matching techniques. When observing these three techniques, recall that these matching techniques are meant to remove biases associated with our pre-treatment. Reading the table from the left to the right, the results move from (potentially) more biased to less biased as we move from unmatched difference-in-difference to PSM all the way to CEM, but at the cost of representativeness of the program as more data is dropped out of the sample. We present these results over the entire treatment period rather than on a year-to-year basis because the decreasing number of researchers as we change matching techniques increases the year-to-year volatility, making visual interpretation of the results difficult. However, to give readers a sense of the temporal nature of the effects, small effects begin in 2008 with large impacts from 2009 to 2011 with effects falling off 2012–2014. The standard error of the mean is included in parentheses.

### 5.1. Conducting high-impact research

We measure this goal using research output, average impact factor of journal publications, and average number of citations per journal article as proxies. In the first row of Table 3, we can see that MPP-affiliated researchers produced between 0.42 and 1.26 new articles per year above the control group, corresponding to a 13%, 31%, and 27% increase in their publication output for DID, PSM, and CEM, respectively. Each of these effect sizes was statistically significant, leading us to conclude that MPP does have an impact on the research output of affiliated researchers and that, if place

our confidence on the matching techniques, this impact is indeed sizeable.

One concern of such an increase in research output quantity is that it may come at the cost of quality. This concern, however, is not confirmed; on the contrary, in row 2 we see a simultaneous, statistically significant increase in impact factor of between 0.84 and 1.78, which for the matching techniques amounts to a relative increasing of 91% (PSM) and 108% (CEM), respectively.

Further, as shown in row 3, we see an increase in the number of citations MPP researchers receive compared the control group, again with a huge jump in effect size when moving from DID (2.42) to matching techniques (6.20 for PSM, 6.21 for CEM). For the latter, this amounts to a relative increase of approximately 124%. While this is not controlled for self-citations, it is clear that the program significantly increases the production of visible, high-impact research by affiliated researchers. We note that the observed increase in impact factor and citation may be partly artificially inflated by the inclusion of “large-name” MIT professors as co-authors, which may potentially influence journal acceptance rates (Lee et al., 2013; Ross et al., 2006) or disproportionately raise visibility. Yet, this outcome was partly intended by the Portuguese government and does not detract from the size and significance of the effect on the publication statistics of the participating institutions and units. An interesting follow-up question is therefore whether this relative increase of MPP-affiliated research over their peers can be maintained over time beyond the duration of the program, which we hope to address in future research.

### 5.2. Empower a new generation of scientists

To tackle the second goal of MPP to support the up-and-coming scientific generation of the country, we focus on two sub-groups of our sample: those with a publishing record of less than four years when the program began (which we take to be “junior faculty”), and those with more 15 years worth of publications (i.e. the “senior faculty”). This analysis complements previous investigations into the effects of student and junior faculty training through the introduction of new doctoral educational programs through MPP or dedicated “enable-the-teachers” programs (Pfotenhauer et al., 2013).

Rows 4 and 5 of Table 3 summarize the effects on these two researcher sub-groups. It is important to note that there are large standard errors associated with the PSM and CEM groups for both young and old faculty. The matching process and further splitting of the sample into junior, senior, and “other” leads to small sample sizes where results can no longer be inferred with statistical significance: for our junior faculty we have 29 treatment researchers and 16 control researcher for PSM and 23 treatment researchers and 21 control researchers for CEM; for our senior faculty we have 116 treatment researchers and 28 control researcher for PSM and 13 treatment researchers and 13 control researchers for CEM. This illustrates one of the weaknesses in using PSM and CEM. Moreover, it is important to remember that our control and treatment groups have the same distributions of pretreatment covariates as a whole, not within these subgroups. Accordingly, there are certainly qualitative differences between the unmatched subgroups and PSM subgroups respectively. We do report these findings here for our measures of productivity only because, while inconclusive, we think they are indicative of an actual effect that we could capture through large sample sizes.

Notwithstanding these statistical problems, we can see that while the effect sizes are roughly equal in junior and senior faculty in absolute terms (0.472 and 0.767 additional papers per year, respectively), the relative impact on publication output is quite dramatic: While this represents a reasonable 15–50% increase in publication output over the control group for senior faculty, for



**Table 3**Effect size of MIT Portugal Program on dependent variables under different matching techniques. Asterisks denote *p*-values of less than 0.05.

|  | Unmatched DID                         |                    | Propensity Score Matching (PSM)       |                    | Coarse Exact Matching (CEM)           |                    |
|--|---------------------------------------|--------------------|---------------------------------------|--------------------|---------------------------------------|--------------------|
|  | Baseline of Control Group (2007–2013) | Effect Size        | Baseline of Control Group (2007–2013) | Effect Size        | Baseline of Control Group (2007–2013) | Effect Size        |
| Average articles per year                  | 3.317                                 | 0.425<br>(0.022)*  | 4.112                                 | 1.264<br>(0.272)*  | 2.582                                 | 0.701<br>(0.184)*  |
| Average impact factor per publication      | 1.646                                 | 0.841<br>(0.025)*  | 1.950                                 | 1.776<br>(0.302)*  | 1.420                                 | 1.541<br>(0.373)*  |
| Average citations per article              | 6.218                                 | 2.428<br>(0.134)*  | 7.399                                 | 6.150<br>(1.830)*  | 5.037                                 | 6.212<br>(1.930)*  |
| Average articles per year (Junior Faculty) | 1.103                                 | 0.472<br>(0.162)   | 1.465                                 | 1.002<br>(0.099)*  | 1.156                                 | 0.759<br>(0.199)   |
| Average articles per year (Senior Faculty) | 4.785                                 | 0.767<br>(0.356)   | 5.335                                 | 0.756<br>(0.547)   | 2.754                                 | 1.412<br>(1.012)   |
| Number of new Portuguese Coauthors         | 21.629                                | 11.363<br>(2.112)* | 23.323                                | 21.763<br>(5.007)* | 16.632                                | 17.136<br>(3.116)* |

junior faculty we observe a very large increase in publication output of between 40 and 70% over the control group. Looking more closely at the data, it appears that while senior researchers, who already have an established collaboration network, can take advantage of the program but at the expense of previously existing collaborative links (i.e. they are “maxed out” already when entering the program), which explains the limited relative increase. On the other hand, junior faculty can take full advantage of these new research collaborations without deleting old productive links because they do not have many previous collaborators. These findings are consistent with self-reported effects observed in interviews carried out in the context of (Pfothenhauer et al., 2013), which indicate that the impact of participating in the partnership is indeed considerably higher on junior researchers as they experience a “jump-starting” of their career. This is consistent with findings by Lee and Bozeman (2005a,b).

### 5.3. Encourage Portuguese universities to work closer together to build domestic research clusters with critical mass

We have operationalized this goal in two ways. First, we measure how many new domestic research partners each researcher acquired during the treatment period. We display the average number of new Portuguese coauthors over the 7-year analysis time-frame in row 6 of Table 3. Over that period, we see large increases in new domestic research partners of between 50% and 100% depending on the analysis technique. These findings are particularly striking when considering that as a small country, Portuguese researchers have one of the highest rates of collaboration in the world already (Glänzel, 2001). Note that at least some of these new coauthors are likely additional student researchers available because of an influx of MPP funding. However, the observed effect cannot be solely attributed to additional students from additional funding: since 2006, 2534 unique new Portuguese collaborators were added to the network and there have only been 371 doctoral students in the program.

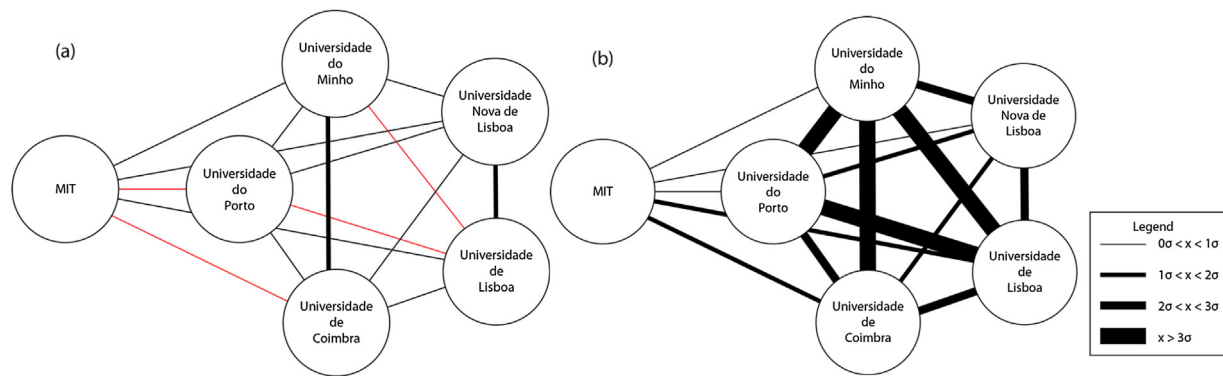
Second, we assess how institution-level collaboration patterns have changed over time before and after the launch of MPP as a result of individual-level collaboration. For this analysis, we used the data acquired through PSM matching, as it represents a middle-ground between restrictiveness (i.e. bias compensations) and representativeness (i.e. inclusion). Fig. 2 shows the differences in research connections by MPP participants relative to the control group and over two different periods of time: 2001–2006 (Fig. 2a) and 2007–2012 (Fig. 2b). For the understanding of this graph, it is important to recognize that it is *not* a depiction of the total number of connections, but indeed of differences in connectivity between

treatment and control. Here, the thickness of the links between universities is a function of the number of standard deviations that the treatment group is above (or below) the equivalent control group connection, and the color of the link represents the sign of that difference (with black representing more connections than the control group and red representing less connections than the control group).

We can see that prior to the onset of MPP (Fig. 2a), collaborations patterns are roughly equivalent between our treatment and control groups, with all but two links falling within one standard deviation (the exceptions being between the Universidade de Coimbra and the Universidade do Minho, and between the Universidade Nova de Lisboa and the Universidade of Lisboa). As expected, the network also contains several links where the treatment group-to-be is institutionally less strongly connected than in our control group (indicated by the red lines). In contrast, Fig. 2b shows the connections between universities from 2007 to 2012 (the first five years of MPP), where we see a tremendous increase in the number of intra-Portuguese connections for the MPP-affiliated group compared to the control group. Most university–university connections are now two or three standard deviations above their control group counterparts, indicating vastly more collaboration between the affiliated departments at these universities. Most interestingly, there is no observable correlation between the strength of connection before and during the program, where one might expect that all collaborative ties will grow but where there were stronger initial connections ties would grow more readily. In fact, the strongest tie – the link between Universidade de Lisboa and Universidade do Porto – was formerly negative.

### 5.4. Encourage Portuguese collaboration with MIT

Our fourth goal analyzed here is to increase Portuguese collaborations with MIT faculty. These collaborative links can also be seen in Fig. 2. The data shows that we equally see a relative increase in number of collaborations between Portuguese institutions and MIT vis-à-vis the control group, with roughly one standard deviation increase for all program-affiliated universities. This increase is moderate compared to the new collaborative ties established between Portuguese universities. Yet, it is precisely in line with the intention of the Portuguese government to use MPP primarily as a “glue” or convening authority among Portuguese universities to overcome traditional patterns of insularity and isolation, and support critical mass formation through increased national collaboration. While the change in tie strength with MIT is less striking than the domestic changes in tie strength, it also suggests that many of the domestic relationships formed as a result of the program



**Fig. 2.** Collaboration patterns between MPP affiliated universities for the five years previous to MPP (a) and five years during MPP (b). The weight of each line represents the number of standard deviations above (black) or below (red) the control group. (Please refer to the web version of this article for color figures.)

might be not dependent on the continued collaboration with MIT in order to be sustained in the future.

We furthermore note that approximately half of the MIT collaborators were not affiliated with the program at the beginning. Yet, interviews among program participants reveal that the informal connections made as a result of participating in the program lead to new introduction and collaboration opportunities, and that MIT researchers who begin collaborating with Portuguese faculty are often added to the program. There is evidence of further spillovers into the collaborative network of MIT researchers, and future work will focus on developing this evidence more deeply.

### 5.5. Shaping research directions

The fifth objective MPP seeks to address is to actively shape the research directions of participating researchers towards designated priority areas. To address this goal, we evaluate the evolution of publication activity of MPP-affiliated faculty in terms of their content – first purely as within-subjects estimates (i.e. independent of comparative DID developments among the peer group) to identify relevant dynamics, then with consideration of the peer by way of DID analysis.

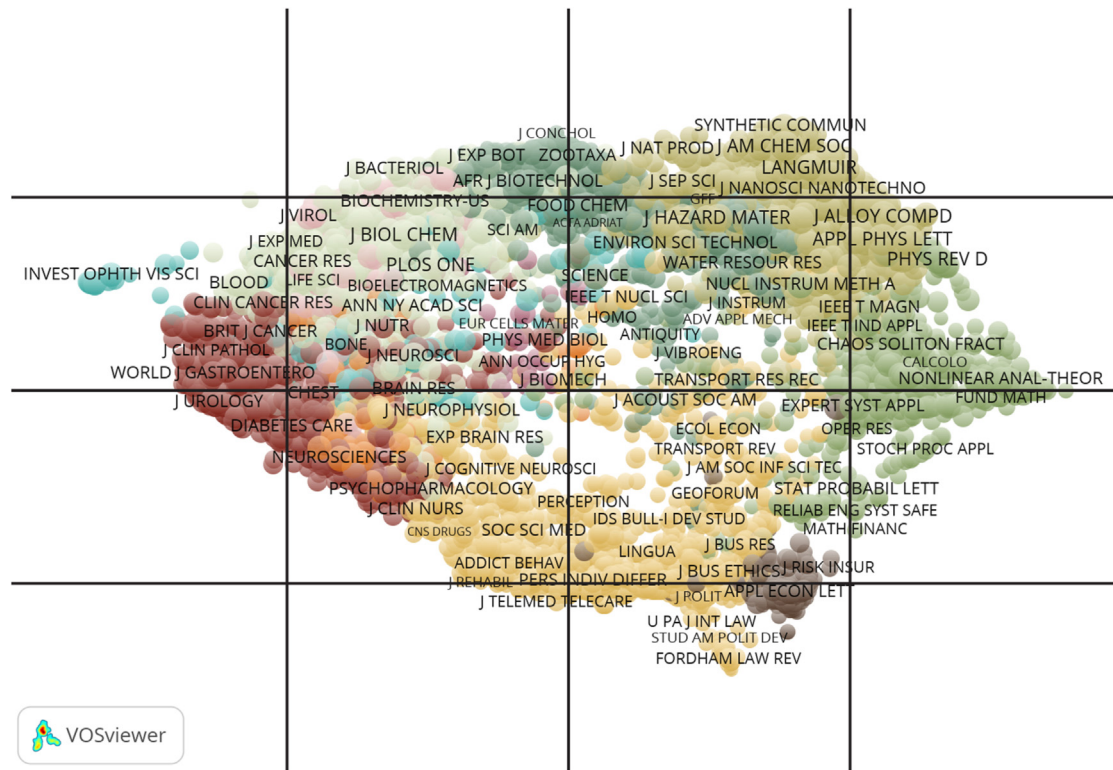
As mentioned above, MPP has four designated focus areas – Bioengineering Systems; Engineering Design and Advanced Manufacturing; Sustainable Energy Systems; Transportation Systems – and we analyze shifts in publication patterns within those areas separately. Here, we limit our discussion to two of these focus areas (Bioengineering Systems, Sustainable Energy Systems), both for reasons of data quality (i.e. limits on sample size for the other two) and because of the Portuguese government’s explicit attempt to shift research in particularly these domains. For example, prior to MPP, Portugal had no major research enterprise in stem cell research, and the program was designed explicitly to draw upon MIT expertise to jumpstart activities in this field. In contrast, in transportation systems Portugal has traditionally had research strength and less intention to shift research directions.

To evaluate changes in research focus over time, we use Content Overlay Maps, a technique used to visualize the research output of a country, university or other entity on a background “map of science.” Originally developed to better understand the nature of interdisciplinarity and evaluate how research topics shift over time (Rafols et al., 2010), “maps of science” are generated by calculating how close certain research fields are to one another through co-citations, shared co-references, or the number of common authors publishing within each field. These relationships are then flattened into a 2 dimensional space (or “map”). Here, we adopt a method based on the citing relationships between 8860 journals, as outlined in Leydesdorff and Rafols (2012). Each journal is placed based

on the map according to how often articles cite articles from other journals, measuring how much journals build upon previous work done. We chose to utilize the “citing” approach throughout our work, rather than “cited-by,” since there is no time dependence (citing maps are constant for each year; cited by maps will constantly be evolving over time as fields change and in particular draw lessons from other fields). Further, a “citing”-approach map can measure instant shifts in year-to-year without lag or dampening, which allows us to measure quick changes in research focus, which is of particular interest to us in this work. For these reasons, the developers of Content Overlay Maps prefer this approach and recommend its use over the “cited-by” approach (Leydesdorff and Rafols, 2012). We show this background map of science in Fig. 3 together with a superimposed coordinate system, where each node (journal) is color categorized by one of 19 clusters representing topic areas.

We now define the *center of research focus* for each researcher (analogous to a center-of-mass calculation), where we take the average of all journal articles published in a year, located at different positions based on a map of science and weighted by the number of publications by the author. We also define a *radius of research focus*, defined as the average distance away from this yearly research focus that a researcher published in that year. For this, we located each of the 19462 publications in our dataset by the weighted average of the articles it cites, rather than simply locating each publication by the position of the journal it appears in (which can have a narrow or large focus, depending on the journal). We choose this approach because citing patterns indicate the previous literature that an article builds upon (i.e. which research area the particular researcher feels familiar or comfortable with), which in our opinion serves as a good correlate for “focus area.” We also note that this approach dramatically decreased the variation in research focus and radius year-to-year for researchers compared to the simpler measure of journal location. We also find that our method provides robust and stable results even for those researchers who only published a few papers per year (where variation had been previously been highest under the simple one journal methodology) and that nearly every researcher maintained a research radius that was relatively constant. Where there were large changes in research focus or radius, we confirmed qualitatively that they were coupled with a true shift in research focus.

Note that we initially tried to use a more common Content Overlay Map technique based on the number of co-citations between the 225 Web of Science categories (formerly the 222 categories), where journals are assigned to these subject categories by citations patterns and expert judgment. We found very abrupt changes in both research focus and radius from year to year, particularly in the cases of those researchers who only published a few articles per year.



**Fig. 3.** Map of Science. The distance between journal pairs (nodes) is proportional to the number of co-referencing between them and colors are based on fit within 19 different clusters representing topic areas.

Interviews with a few of these researchers, along with qualitative reviews of their CVs, DID not confirm abrupt year-to-year changes that the method seemed to detect, which is why we discarded this coarser approach.

We now overlay the production of each affiliated researcher (i.e. the circular area covered by the *radius of research focus* centered on the *center of research focus*) onto the background “map of science” and visualize the cumulative density of research production for our two topic areas of interest over time, resulting in a number of time-resolved “heat maps” (Fig. 4). Note that there has been no smoothing of year-to-year variations, as we are trying to investigate the dynamics of change in research focus. It is also important to consider that while it may be tempting to use these metrics to quantitatively study changes in researcher focus radius, average distance between researchers, and so on, this would be improperly using this Map of Science. As (Rafols et al., 2010) note, these relationships are originally multidimensional and flattened to 2-dimensions, leading to many potential “correct” maps where the distance between each journals will change. Consequently, we are much more focused on observing changes in research focus (and the overall fields of those changes) rather than specifics of spreading and distance.

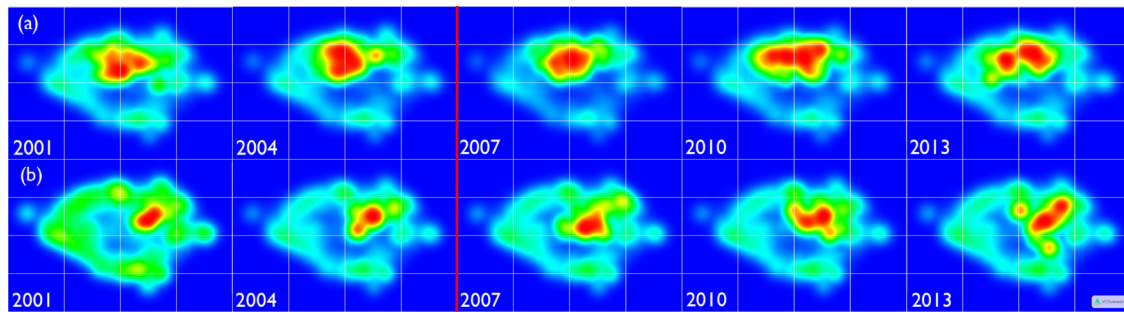
Fig. 4a shows the evolution of research focus of researchers affiliated with the Bioengineering Systems focus areas over the last 12 years in intervals of 3 years. We observe that from 2001 to 2007, we see little change in research focus, with researchers remaining centered around one major topic area, which we can broadly define as *basic biology* research. Subsequently, from 2007 to 2013, we observe see a diversification of research activities towards two pronounced, new research areas, accompanied by an abandoning of the old topic area. On the background map of science, we can identify these two new areas (roughly) as *biochemistry* (with some applied medical research) and *bioengineering/stem cell engineering* – noting, however, that such labels based on journal outlets should

be taken with a (relatively large) grain of salt. Yet, the observed dynamics in principle match the intended program effects.

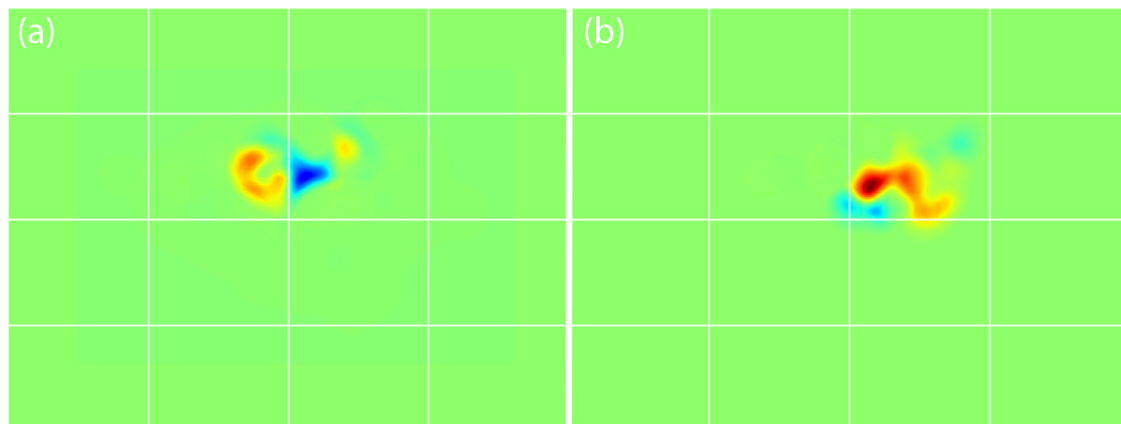
Fig. 4b tells a different story of change in research focus for the Sustainable Energy Systems faculty. Up until 2007, we see relatively little change in research patterns; however, starting in 2010 we see a spreading out of the covered research areas to include new domains such that by 2013 the former tightly clustered research space has spread out and spans almost an entire sector of the research map. However, unlike the Bioengineering Systems, not all faculty move away from their previous research interests, which is why the shift appears in the form of an expansion of the original center rather than two separate foci. Again, we can broadly categorize these developments according to publication outlets: Starting from a center around Energy Systems (which has been a focus since the beginning), we see a gradual expansion into Environmental Science and Transportation Systems (which are new areas of scientific engagement). Again, this points to a potential effect in agreement with stated program goals, as embodied by key projects around sustainable energy solutions and integrated energy-transportation systems such as in the Azores project.

These observed dynamics point out regions of interest where one can look for potential program impacts. To validate that these dynamics are indeed attributable to MPP, we now take an additional step and compare them to parallel developments in the control group, trying to distinguish MPP effects from broader trends in the field that might lead to similar shifts in research focus.

Fig. 5a & b show the difference-in-differences of publication activity on the map of science between treatment and control group when moving from 2007 to 2010 (the most important step for observed changes) for the Bioengineering Systems focus area and Sustainable Energy Systems focus area, respectively. We emphasize that, to the best of our knowledge, this is the first time that a DID approach has been used for Content Overlay Maps. Here, the red scale indicates different degrees of relative increase in research



**Fig. 4.** Heat map of research activity for two MPP focus areas in three-year steps: (a) Bioengineering Systems and (b) Sustainable Energy Systems faculty. The coordinate systems (grey lines) correspond to the background map of science (Fig. 3). The ‘treatment period’ (MPP launch) begins in 2007 (red line). (Please refer to the web version of this article for color figures.)



**Fig. 5.** Difference-in-Difference between 2010 and 2007 fields of publication for the (a) Bioengineering Systems and (b) Sustainable Energy Systems Focus Areas. Areas in red denote more publishing by the treatment group in 2010 than in 2007, normalized by changes in control group publishing over the same time period, while areas of blue represent less publishing in 2010 than 2007. (Please refer to the web version of this article for color figures.)

activity of the treatment group from 2010 to 2007 that was *not* also seen in the control group over that time, and the blue scale a relative decrease. That is, the graph thus shows a true DID-version of a content overlay map for two sample populations. In agreement with the above-discussed patterns, we find that the observed shifts represent indeed a trend that is seen in MPP-affiliated researchers only, who shift their research focus along the identified axes. For Bioengineering Systems (5a), we observe the emergence of two distinct new foci (red/yellow) and a simultaneous abandoning of the original area at the center (blue), demonstrating that there is no corresponding trend for the control group. Likewise, for Sustainable Energy Systems (5b), we observe a shift towards new publications in environmental science and/or transportation systems (red areas), yet without a relative decrease in the original research focus (almost no blue areas).

We conclude that MPP seems to have been successful in affecting the research focus of leading research groups in the country: For both Bioengineering Systems and Sustainable Energy Systems, the portfolio of research for program-affiliated faculty shifted recognizably, with new distinct areas of combined research emphasis emerging, while no such trend is observed for the control group. However, it is important to recall that while we have taken our control groups from the same fields and academic departments, we have not controlled explicitly for specific research focus and do not explicitly have a counterfactual situation to work with. Selection into, or application for, the MIT-Portugal program could have been motivated by an individual researcher’s existing desire to change research directions in the first place and the program participation only acted as a vehicle for this change. Accordingly, we have no way of disentangling the causal direction here: the observed

change in research focus could result from “pull” from funding and opportunities that come with MPP-affiliation; from “push” into new fields that drove self-selection into application to the program; or a combination thereof. However, we note that again that part of the motivation for the government to initiate the program was to jumpstart research in certain priority areas where Portugal did not exhibit significant research activity before (e.g. stem cell research). From an evaluation perspective and notwithstanding the causal direction, this goal was achieved.

## 6. Discussion and conclusions

In this paper, we investigated the impact of one large-scale national “flagship” research policy initiative – the MIT Portugal Program – that sits in the crossroads of four major current policy trends: university-centrism, collaboration, internationalization, and growing complexity. Our aim was to evaluate the partnership, to the extent possible, ‘on its own terms,’ tackling three structurally oriented, architecture- and context-specific program goals along two more conventional ones. To that end, we developed a custom-tailored mixed method approach that combines network analysis, system architecture analysis, difference-in-difference (DID) program evaluation and statistical matching techniques. We used publication data from the last 20 years to analyze changes in publication and collaboration behavior of participating and non-participating (peer) researchers before and after the partnerships commenced. To ensure comparability, we assembled the control group from the same departments at the same universities and compared them to our treatment (MPP) group, controlling for five criteria – comparable levels of seniority, comparable pre-MPP

productivity, comparable quality, comparable visibility and comparable research areas – and reducing potential selection biased through two matching mechanisms, Propensity Score Matching and Coarse Exact Matching.

This research has a number of broad policy implications. Overall, we find strong evidence that the MIT Portugal Program has indeed had a large impact and achieved many of its idiosyncratic program objectives. In terms of research productivity, we find that MPP-affiliated researchers produce between approximately 15% to 30% more articles than the control cohort while still increasing the quality (impact factor) and visibility (citations) of their publications. While this increase is an encouraging sign, one important question remains outstanding around the sustainability of this increase long term. One encouraging finding on the long-term impact is the relatively larger impact for junior faculty, for whom MPP participation amounts to a “jumpstarting” of their research career. Even if future long-term studies find that direct effects from collaboration with MIT fall off quickly after a partnership ends, increasing researcher visibility within the field early in their career could lead to significant feedback mechanisms, reinforcing early leads with funding, offers of co-authorship and so on. This underscores that the program was indeed successful in “training a new generation of scientific leaders” and adds to previous evidence around student performance (Pfotenhauer et al., 2013). While we see a larger effect on junior faculty from this partnership, we would caution policymakers in taking this as evidence advocating for programs targeted specifically at young researchers: the mechanism by which junior researchers increase their output could be solely from collaboration with MIT researchers, but could also be due in a large part to new collaborations domestically with other participants.

In particular, our findings around changes in collaborative network are very interesting for policymakers. We observed that MPP had a vast effect on domestic institutional collaboration patterns and intensity and, in particular, was a vehicle for addressing a long-standing concern of the Portuguese government about traditional patterns of institutional isolation, competition, and lack of collaboration. Interestingly, we see a much more moderate increase with collaboration activities between Portuguese institutions and MIT, which we take to indicate that all of these new intra-Portugal collaborative ties are not solely dependent on MIT to take place. In essence, this work has shown that the deliberate intervention into the “natural” progression of a collaboration network is possible and, perhaps most importantly, that such changes can persist over time. Particularly amongst new findings that show that collaborative network can be a strong predictor of future success (such as (Bertsimas et al., 2015) demonstrated around tenure decision), this work raises an important discussion around the ethics of nations, regions or even particular institutions deliberately intervening in the collaboration network of a researcher in order to increase their future output. Further case studies where similar programs have been concluded will be necessary to explore questions around the persistence of such collaborations without an external forcing function. Future case studies will also be able to examine if a large increase in collaborations could be idiosyncratic to this one case, where domestic researchers had a history of under-collaboration making it ripe for such an intervention, or if this is an outcome across other such international partnerships.

Our research demonstrates that large-scale international partnership like MPP can be utilized to shape the research focus of leading researchers and institutions. Influenced by collaborations with their MIT counterparts, Portuguese research teams at top universities were observed to migrate (on average) to new domains. Further study will have to be conducted to determine if these changes will be lasting or if this was simply a temporary shift on the part of researchers to “follow the money.” This work is limited by the self-selection into the program (by choosing to apply for the

program in one specific focus area) where treatment researchers may have been more biased towards shifting their research priorities in the direction of the program. Accordingly, there remains an outstanding question behind the mechanism responsible for changing these research topics, and if it is possible to change the research priorities of a nation, region or university with such a research policy tool without the prior interest of participating researchers. Finally, this case study has been focused on the impact on Portuguese faculty but there are a number of interesting questions surrounding the impact on participating MIT faculty, which we hope to address in future work.

Our study of MPP also breaks new methodological ground for policy analysis of complex collaborative research policy initiatives, in a number of ways. First, DID analysis is scarcely used in bibliometric studies. Here, we demonstrated the potential of such DID, when used together with both covariates matching techniques to properly assemble comparable control groups, to assess relative shifts in network formation and publication patterns. While matching techniques may seem like a “black box,” where the results are so opaque as to be untrusted, we believe that we have demonstrated that such an evaluation can be done authentically and transparently. Second, our research combined for the first time Content Overlay Maps (“maps of science”) with DID analysis, demonstrating how changes in research focus can be traced against the background of a peer group of researchers. We hope such maps can be used as an intuitive method for representing changes in research priorities more broadly amongst the policymaking community. Future work in this case could focus on multidimensional mapping of research focus, allowing a strong quantitative look at developing important metrics such as distance between researchers, areas of overlapping expertise and changes in research focus, each of which were limited by our reliance on our two-dimensional projection. Third, our approach treaded a fine line between inclusiveness and representativeness in treatment-control comparison on the one hand, and progressively restrictive eradication of remaining selection biases via statistical matching techniques on the other. We show that all three approaches lead to consistent results, leaving to the reader the choice which information she prefers. One interesting observation in this work is that unmatched DiD results nearly always demonstrated the smallest average effect, PSM the largest and CEM somewhere in between. At the same time, PSM also produced the largest average baseline outcomes of the control group, indicating that this effect was not simply a byproduct of the highest performing researchers being discarded under PSM. While we suspect that this result is idiosyncratic of our particular dataset of treatment and control groups, this pattern could also be a function of our utilization of a very small “caliper,” or maximum distance in propensity score away matched are allowed to happen. Future work on a number of different cases studies will flesh this relationship between average effect size and matching methodology. Finally, we derived program goals using a systems architecture perspective, which focuses the attention on structural goals embedded in program architecture and shaped by social, political, and historical context.

Future work will strive to explore this line of inquiry further by studying the relationship between different program architectures and network formation as well as institutional and cultural change using cross-case comparisons across a number of case studies of complex international research partnerships. For example, what influence does the consortium structure (including multiple Portuguese universities) of MPP have compared to, say, a classical bilateral partnership structure? How does the level of maturity of the Portuguese science and innovation system – its starting point and most recent evolution – exert influence on network structure? Such an understanding would help policymakers design future partnerships and will begin to illuminate the mechanisms

behind the impact of such programs. Any quantitative comparative study, however, will have to be complemented by dedicated qualitative analysis into the policy context that gives rise to, and limits, specific complex initiatives. This work has already begun, most notably in (Pfotenhauer et al., 2016, 2013; Pfotenhauer and Jasanoff, forthcoming-a, forthcoming-b; Pfotenhauer 2016), and we expect future work to combine both a qualitative and quantitative perspective. This would also provide a complement to studies of the impacts of different organizational practices on collaboration in university-industry contexts (Siegel et al., 2003), and to studies of different national knowledge cultures and their idiosyncratic forms of institutionalization (Jasanoff, 2005, 2004).

We believe that MPP is by no means a one-off exception (although its architecture is quite unique). Rather, it is representative of a new generation of structurally custom-tailored, large-scale policy initiatives with a focus on capacity building through international collaboration, anchored at leading national and international universities. In this environment, our study of MPP and the methodological developments underwriting it offer a way forward for program evaluation of complex policy initiatives, recognizing that evaluation is an equally complex task that needs to pay attention to idiosyncratic, oftentimes structurally oriented objectives of national flagship initiatives.

### Acknowledgments

This work was sponsored by the NSF Science of Science and Innovation Policy (SciSIP) Program under Collaborative Grant “Technology, Collaboration, Learning: Modeling Complex International Innovation Partnerships” (#1262263). MH gratefully acknowledges partial funding and support of the research agenda through the MIT-Portugal Program. The authors particularly thank Danielle Wood, Dava Newman, Dan Roos, Manuel Heitor, and Cesar Hidalgo for conceptual and methodological input.

### References

- Abadie, A., 2005. *Semiparametric estimators*. *Rev. Econ. Stud.* 72, 1–19.
- Altbach, P.G., Knight, J., 2007. The internationalization of higher education: motivations and realities. *J. Stud. Int. Educ.* 11, 290–305, <http://dx.doi.org/10.1177/1028315307303542>.
- Altbach, P.G., Reisberg, L., Rumbley, L.E., 2010. *Trends in Global Higher Education. UNESCO*.
- Amaral, A., Magalhães, A., 2005. *Implementation of Higher Education Policies: A Portuguese Example*. Springer, Netherlands, Dordrecht.
- Amsden, A., 2001. *Rise of The Rest: Challenges to the West from Late-Industrializing Economies*. Oxford University Press, New York.
- Archibugi, D., Lundvall, B.-Å., 2001a. *The Globalizing Learning Economy: Major Socio-Economic Trends and European Innovation Policy*. Oxford University Press.
- Archibugi, D., Lundvall -Å, B. (Eds.), 2001b. *Oxford University Press, New York*.
- Ashenfelter, O.C., Card, D., 1984. Using the longitudinal structure of earnings to estimate the effect of training programs. *Rev. Econ. Stat.* 67, 648–660, <http://dx.doi.org/10.2307/1924810>.
- Bailetti, A.J., Callahan, J.R., DiPietro, P., 1994. A coordination structure approach to the management of projects. *IEEE Trans. Eng. Manag.* 41, 394–403, <http://dx.doi.org/10.1109/17.364565>.
- Bakshy, E., Hofman, J.M., Watts, D.J., Mason, W.A., 2011. *Everyone's an influencer: quantifying influence on twitter categories and subject descriptors*. *Proceedings of The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 65–74.
- Balmer, B., 1996. Managing mapping in the human genome project. *Soc. Stud. Sci.* 26, 531–573, <http://dx.doi.org/10.1177/030631296026003002>.
- Bammer, G., 2008. Enhancing research collaborations: three key management challenges. *Res. Policy* 37, 875–887, <http://dx.doi.org/10.1016/j.respol.2008.03.004>.
- Bertsimas, D., Brynjolfsson, E., Reichman, S., Silberholz, J., 2015. OR forum—tenure analytics: models for predicting research impact. *Oper. Res.* 63, 1246–1261, <http://dx.doi.org/10.1287/opre.2015.1447>.
- Bhandari, R., Blumenthal, P. (Eds.), 2011. *Palgrave Macmillan, New York*.
- Bound, J., Turner, S., Walsh, P., 2009. *Internationalization of U.S. doctorate education*. In: *Science and Engineering Careers in the United States*. NBER, pp. 59–97.
- Bozeman, B., Gaughan, M., 2011. How do men and women differ in. *Res. Policy* 40, 1393–1402, <http://dx.doi.org/10.1016/j.respol.2011.07.002>.
- Bozeman, B., Fay, D., Slade, C.P., 2012. Research collaboration in universities and academic entrepreneurship: the-state-of-the-art. *J. Technol. Transfer*, <http://dx.doi.org/10.1007/s10961-012-9281-8>.
- Bozeman, B., Fay, D., Slade, C.P., 2013. Research collaboration in universities and academic entrepreneurship: the-state-of-the-art. *J. Technol. Transfer* 38, 1–67, <http://dx.doi.org/10.1007/s10961-012-9281-8>.
- Bozeman, B., 2000. Technology transfer and public policy: a review of research and theory. *Res. Policy* 29, 627–655, [http://dx.doi.org/10.1016/S0048-7333\(99\)00093-1](http://dx.doi.org/10.1016/S0048-7333(99)00093-1).
- Burt, R.S., 2000. The network structure of social capital. *Res. Organ. Behav.* 22, 345–423, [http://dx.doi.org/10.1016/S0191-3085\(00\)22009-1](http://dx.doi.org/10.1016/S0191-3085(00)22009-1).
- Burt, R.S., 2004. Structural holes and good ideas. *Am. J. Sociol.* 110, 349–399.
- Carayannis, E.G., Laget, P., 2004. *Transatlantic innovation infrastructure networks: public-private, EU-US R&D partnerships*. *RD Manage.* 34, 17–31.
- Carayannis, E.G., Alexander, J., Ioannidis, A., 2000. Leveraging knowledge, learning, and innovation in forming strategic government-university-industry (GUI) R&D partnerships in the US, Germany, and France. *Technovation* 20 (9), 477–488.
- Card, D., Krueger, A.B., 1994. *Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania*. *Am. Econ. Rev.* 84, 772–793.
- Casper, S., Murray, F., 2005. Careers and clusters: analyzing the career network dynamic of biotechnology clusters. *J. Eng. Technol. Manage.* 22, 51–74, <http://dx.doi.org/10.1016/j.jengtecman.2004.11.009>.
- Centola, D., Macy, M., 2007. Complex contagions and the weakness of long ties. *Am. J. Sociol.* 113, 702–734.
- Cetina, K.K., 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press, Cambridge, Mass.
- Chompalov, I., Shrum, W., 1999. Institutional collaboration in science: a typology of technological practice. *Sci. Technol. Hum. Values* 24, 338–372, <http://dx.doi.org/10.1177/016224399902400302>.
- Chompalov, I., Genuth, J., Shrum, W., 2002. The organization of scientific collaborations. *Res. Policy* 31, 749–767, [http://dx.doi.org/10.1016/S0048-7333\(01\)00145-7](http://dx.doi.org/10.1016/S0048-7333(01)00145-7).
- Cochran, W.G., Rubin, D.B., 1973. *Controlling bias in observational studies: a review*. *Indian J. Stat. Ser. A* 35, 417–446.
- Cohen, W., Levinthal, D., 1990. Absorptive capacity: a new perspective on learning and innovation. *Adm. Sci. Q.* 35.
- Conceição, P., Heitor, M., 2001a. *Universities in the learning economy: balancing institutional integrity with organizational diversity*. In: Archibugi, D., Lundvall -Å, B. (Eds.), *The Globalizing Learning Economy*. Oxford University Press, New York, pp. 83–96.
- Conceição, P., Heitor, M., 2001b. *Universities in the learning economy: balancing institutional integrity with organizational diversity*. In: Archibugi, D., Lundvall -Å, B. (Eds.), *The Globalizing Learning Economy*. Oxford University Press, New York, pp. 83–96.
- Corley, E. a., Boardman, P.C., Bozeman, B., 2006. Design and the management of multi-institutional research collaborations: theoretical implications from two case studies. *Res. Policy* 35, 975–993, <http://dx.doi.org/10.1016/j.respol.2006.05.003>.
- Cowan, R., Zinovyeva, N., 2013. University effects on regional innovation. *Res. Policy* 42, 788–800, <http://dx.doi.org/10.1016/j.respol.2012.10.001>.
- Crawford, E., Shinn, T., Sörlin, S. (Eds.), 2010. 1993 edition. Springer.
- Dehejia, R.H., Wahba, S., 2002. *Propensity score-matching methods for nonexperimental causal studies*. *Rev. Econ. Stat.* 84, 151–161.
- Dietz, J.S., Bozeman, B., 2005a. Academic careers, patents, and productivity: industry experience as scientific and technical human capital. *Res. Policy* 34, 349–367, <http://dx.doi.org/10.1016/j.respol.2005.01.008>.
- Dietz, J.S., Bozeman, B., 2005b. Academic careers, patents, and productivity: industry experience as scientific and technical human capital. *Res. Policy* 34, 349–367, <http://dx.doi.org/10.1016/j.respol.2005.01.008>.
- EC, 2012. *The European Research Area in The Knowledge Triangle* [WWW Document], URL [http://ec.europa.eu/research/era/understanding/what/era\\_in\\_the\\_knowledge\\_triangle\\_en.htm](http://ec.europa.eu/research/era/understanding/what/era_in_the_knowledge_triangle_en.htm) (accessed 2.7.13).
- Edquist, C., 2006. *Systems of Innovation: Perspectives and Challenges*. In: *The Oxford Handbook of Innovation*. Oxford University Press, <http://dx.doi.org/10.1093/oxfordhb/9780199286805.003.0007>.
- Eom B.-Y., Lee K., 2010. Determinants of industry-academy linkages and, their impact on firm performance: The case of Korea as a latecomer in knowledge industrialization. *Res. Policy, Special Section on Government as Entrepreneur* 39, 625–639. doi: <http://dx.doi.org/10.1016/j.respol.2010.01.015>.
- Etzkowitz, H., 2002. *Incubation of incubators: innovation as a triple helix of university-industry-government networks*. *Sci. Public Policy* 29 (2), 115–128.
- Etzkowitz, H., 2003. *Research groups as quasi-firms: the invention of the entrepreneurial university*. *Res. Policy* 32, 109–121.
- Etzkowitz, H., 2008. *The Triple Helix: Industry, University, and Government in Innovation*. Routledge, New York.
- Finn, M.G., 2012. *Stay Rates of Foreign Doctorate Recipients from U.S. Universities 2009*. National Science Foundation Division of Science Resource Statistics.
- Fransman, M., King, K., Bell, M. (Eds.), 1984. *Technological Capability in the Third World*. Macmillan, pp. 187–209.
- Freeman, C., 2001. *The learning economy and international inequality*. In: Archibugi, D., Lundvall, B.-Å. (Eds.), *The Globalizing Learning Economy*. Oxford University Press, New York.
- Freeman, R.B., 2005. *Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership?* (Working Paper No. 11457). National Bureau of Economic Research.

- Freeman, R.B., 2014a. *Immigration, International Collaboration, and Innovation: Science and Technology Policy in The Global Economy* (Working Paper No. 20521). National Bureau of Economic Research.
- Freeman, R.B., 2014b. *Immigration, International Collaboration, and Innovation: Science and Technology Policy in The Global Economy* (Working Paper No. 20521). National Bureau of Economic Research.
- Galison, P.L., 1994. *Big Science: The Growth of Large-Scale Research*, 1st edition. Stanford University Press Stanford, Calif.
- Garner, J.G., Porter, A.L., Newman, N.C., Crowl, T.a., 2012. Assessing research network and disciplinary engagement changes induced by an NSF program. *Res. Eval.* 21, 89–104, <http://dx.doi.org/10.1093/reseval/rvs004>.
- Georghiu, L., Roessner, D., 2000. Evaluating technology programs: tools and methods. *Res. Policy* 29, 657–678, [http://dx.doi.org/10.1016/S0048-7333\(99\)00094-3](http://dx.doi.org/10.1016/S0048-7333(99)00094-3).
- Georghiu, L., 1998a. Global cooperation in research. *Res. Policy* 27, 611–626, [http://dx.doi.org/10.1016/S0048-7333\(98\)00054-7](http://dx.doi.org/10.1016/S0048-7333(98)00054-7).
- Georghiu, L., 1998b. Global cooperation in research. *Res. Policy* 27, 611–626, [http://dx.doi.org/10.1016/S0048-7333\(98\)00054-7](http://dx.doi.org/10.1016/S0048-7333(98)00054-7).
- Glänzel, W., 2001. National characteristics in international scientific co-authorship relations. *Scientometrics* 51, 69–115, <http://dx.doi.org/10.1023/A:1010512628145>.
- Granovetter, M., 1978. *Threshold models of collective behavior*. *Am. J. Sociol.* 83, 1420–1443.
- Grauwin, S., Beslon, G., Fleury, É., Franceschelli, S., Robardet, C., Rouquier, J.-B., Jensen, P., 2012. Complex systems science: dreams of universality, interdisciplinarity reality. *J. Am. Soc. Inf. Sci. Technol.* 63, 1327–1338, <http://dx.doi.org/10.1002/asi.22644>.
- Hagedoorn, J., Link, A.N., Vonortas, N.S., 2000. Research partnerships. *Res. Policy* 29, 567–586, [http://dx.doi.org/10.1016/S0048-7333\(99\)00090-6](http://dx.doi.org/10.1016/S0048-7333(99)00090-6).
- Hansen, M.T., 1999. The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. *Adm. Sci. Q.* 44, 82, <http://dx.doi.org/10.2307/2667032>.
- Heitor, M., Bravo, M., 2010. Portugal at the crossroads of change, facing the shock of the new: people, knowledge and ideas fostering the social fabric to facilitate the concentration of knowledge integrated communities. *Technol. Forecast. Soc. Change* 77, 218–247, <http://dx.doi.org/10.1016/j.techfore.2009.10.006>.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I., 2015. The Leiden Manifesto for research metrics. Use these ten principles to guide research evaluation. *Nature* 520, 9–11, <http://dx.doi.org/10.1038/520429a>.
- Hidalgo, C.A., Klingler, B., Barabási, A.-L., Hausmann, R., 2007. The product space conditions the development of nations. *Science* 317, 482–487, <http://dx.doi.org/10.1126/science.1144581>.
- Horta, H., 2008. *On improving the university research base: the technical university of lisbon case in perspective*. *High. Educ. Policy* 21, 123–146.
- Hunt, J., Gauthier-Loiselle, M., 2008a. *How Much Does Immigration Boost Innovation?* (Working Paper No. 14312). National Bureau of Economic Research.
- Hunt, J., Gauthier-Loiselle, M., 2008b. *How Much Does Immigration Boost Innovation?* (Working Paper No. 14312). National Bureau of Economic Research.
- Iacus, S.M., King, G., Porro, G., 2009. *cem: Software for Coarsened Exact Matching*, pp. 30.
- Iacus, S.M., King, G., Porro, G., 2011. Causal inference without balance checking: coarsened exact matching. *Polit. Anal.* 20, 1–24, <http://dx.doi.org/10.1093/pan/mp1013>.
- Jasanoff, S. (Ed.), 2004. *States of Knowledge: the Co-production of Science and Social Order*. Routledge, London; New York.
- Jasanoff, S., Kim, S.-H., 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press, Chicago; London.
- Jasanoff, S., 2005. *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton University Press.
- Katz, J.S., Martin, B.R., 1997a. What is research collaboration? *Res. Policy* 26, 1–18, [http://dx.doi.org/10.1016/S0048-7333\(96\)00917-1](http://dx.doi.org/10.1016/S0048-7333(96)00917-1).
- Katz, J.S., Martin, B.R., 1997b. What is research collaboration? *Res. Policy* 26, 1–18, [http://dx.doi.org/10.1016/S0048-7333\(96\)00917-1](http://dx.doi.org/10.1016/S0048-7333(96)00917-1).
- Kearns, M., Suri, S., Montfort, N., 2014. An Experimental Study of the Coloring Problem on Human Subject Networks Author(s): Michael Kearns, Siddharth Suri and Nick Montfort Source: *Science*, 313 824–827.
- Keller, W., 1996. Absorptive capacity: on the creation and acquisition of technology in development. *J. Dev. Econ.* 49, 199–227, [http://dx.doi.org/10.1016/0304-3878\(95\)00060-7](http://dx.doi.org/10.1016/0304-3878(95)00060-7).
- Kempe, D., Kleinberg, J., Tardos, É., 2003. Maximizing the spread of influence through a social network. *Proc. Ninth ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.* – KDD 03 137. doi: 10.1145/956755.956769.
- Kim, L., 1997. *Imitation to Innovation: The Dynamics of Korea's Technological Learning*. Harvard Business School Press, Boston, MA.
- Knight, J., 2004. *Internationalization remodeled: definition, approaches, and rationales*. *J. Stud. Int. Educ.* 8 (Spring), 5–31.
- Kuhlmann, Stefan, Shapira, Philip, Smits, Ruud, 2012. *Introduction, a systemic perspective: the innovation policy dance*. In: Smits, R., Kuhlmann, S., Shapira, P. (Eds.), *The Theory and Practice of Innovation Policy: An International Research Handbook*. Edward Elgar Pub, Cheltenham, UK u.a., pp. 1–22.
- Lazer, D., Friedman, A., 2007. *The network structure of exploration and exploitation*. *Adm. Sci. Q.* 52, 667–694.
- Lee, S., Bozeman, B., 2005a. The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.* 35, 673–702, <http://dx.doi.org/10.1177/0306312705052359>.
- Lee, S., Bozeman, B., 2005b. The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.* 35, 673–702, <http://dx.doi.org/10.1177/0306312705052359>.
- Lee, K., Lim, C., 2001. Technological regimes, catching-up and leapfrogging: findings from the Korean industries. *Res. Policy* 30, 459–483, [http://dx.doi.org/10.1016/S0048-7333\(00\)00088-3](http://dx.doi.org/10.1016/S0048-7333(00)00088-3).
- Lee, C.J., Sugimoto, C.R., Zhang, G., Cronin, B., 2013. Bias in peer review. *J. Am. Soc. Inf. Sci. Technol.* 64, 2–17, <http://dx.doi.org/10.1002/asi.22784>.
- Leydesdorff, L., Rafols, I., 2012. Interactive overlays: a new method for generating global journal maps from web-of-Science data. *J. Informetr.* 6, 318–332, <http://dx.doi.org/10.1016/j.joi.2011.11.003>.
- Llerena, P., Matt, M., 2005. *Innovation Policy in a Knowledge-Based Economy: Theory and Practice*. Springer.
- Lundvall, B.-A., 1992. *National Systems of Innovation*.
- Luukkonen, T., Persson, O., Sivertsen, G., 1992a. Understanding patterns of international scientific collaboration. *Sci. Technol. Hum. Values* 17, 101–126, <http://dx.doi.org/10.1177/016224399201700106>.
- Luukkonen, T., Persson, O., Sivertsen, G., 1992b. Understanding patterns of international scientific collaboration. *Sci. Technol. Hum. Values* 17, 101–126, <http://dx.doi.org/10.1177/016224399201700106>.
- MIT, 2005. *Program Overview Brochure*, MIT Portugal Program, [http://www.mitportugal.org/index.php?option=com\\_docman&task=doc.download&gid=3&Itemid=383](http://www.mitportugal.org/index.php?option=com_docman&task=doc.download&gid=3&Itemid=383).
- Magro, E., Wilson, J.R., 2013. Complex innovation policy systems: towards an evaluation mix. *Res. Policy* 42, 1647–1656, <http://dx.doi.org/10.1016/j.respol.2013.06.005>.
- Mansfield, E., Lee, J.-Y., 1996. The modern university: contributor to industrial innovation and recipient of industrial R&D support. *Res. Policy* 25, 1047–1058, [http://dx.doi.org/10.1016/S0048-7333\(96\)00893-1](http://dx.doi.org/10.1016/S0048-7333(96)00893-1).
- Mansfield, E., 1991. Academic research and industrial innovation. *Res. Policy* 20, 1–12, [http://dx.doi.org/10.1016/0048-7333\(91\)90080-A](http://dx.doi.org/10.1016/0048-7333(91)90080-A).
- Martin, B.R., Irvine, J., 1983. Assessing basic research: some partial indicators of scientific progress in radio astronomy. *Res. Policy* 12, 61–90, [http://dx.doi.org/10.1016/0048-7333\(83\)90005-7](http://dx.doi.org/10.1016/0048-7333(83)90005-7).
- Melin, G., 2000. Pragmatism and self-organization: research collaboration on the individual level. *Res. Policy* 29, 31–40, [http://dx.doi.org/10.1016/S0048-7333\(99\)00031-1](http://dx.doi.org/10.1016/S0048-7333(99)00031-1).
- Moed, H.F., Burger, W.J.M., Frankfort, J.G., Van Raan, a. F.J., 1985. The use of bibliometric data for the measurement of university research performance. *Res. Policy* 14, 131–149, [http://dx.doi.org/10.1016/0048-7333\(85\)90012-5](http://dx.doi.org/10.1016/0048-7333(85)90012-5).
- Mowery, D.C., 2004. *Ivory Tower and Industrial Innovation: University-Industry Technology Transfer Before and After the Bayh-Dole Act in the United States*. Stanford University Press.
- Murray, F., 2004. The role of academic inventors in entrepreneurial firms: sharing the laboratory life. *Res. Policy* 33, 643–659, <http://dx.doi.org/10.1016/j.respol.2004.01.013>.
- Nelson, R. (Ed.), 1993. *A Comparative Analysis*. Oxford University Press, New York/Oxford.
- Newman, M.E.J., 2001. Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev. E* 64, 016131, <http://dx.doi.org/10.1103/PhysRevE.64.016131>.
- OECD, 2007. *Tertiary Education in Portugal*. OECD, Paris.
- OECD, 2010a. *The OECD Innovation Strategy*. OECD, Paris.
- OECD, 2010b. *Measuring Innovation: A New Perspective*. OECD, Paris.
- Pfothenauer, S.M., Jasanoff, S., forthcoming-a. Panacea or diagnosis? Imaginaries of innovation and the “MIT model” in three political cultures. *Soc. Stud. Sci.*
- Pfothenauer, S.M., Jasanoff, S., forthcoming-b. Traveling imaginaries: The “practice turn” in Innovation Policy and the global circulation of innovation models. In: Tyfield, D., Randalls, S., Lave, R., Randalls, S. (Eds.), *The Routledge Handbook of the Political Economy of Science*. Routledge.
- Pfothenauer, S.M., 2016. *Co-producing Emirati science and society at Masdar Institute of Science and Technology*. In: Siddiqi, A., Diaz-Anadon, L. (Eds.), *Science and Technology Development in the Gulf States: Economic Diversification Through Regional Collaboration*. Gerlach Press.
- Pfothenauer, S.M., 2017. Can innovative universities be made? Questioning the globalizing “MIT model” In: Wisnioski, M., Hintz, E., Stettler, M. (Eds.), *Can Innovators Be Made?*
- Pfothenauer, S.M., Jacobs, J.S., Pertuze, J.A., Newman, D.J., Roos, D.T., 2013. Seeding change through international university partnerships: the MIT-Portugal Program as a driver of internationalization networking, and innovation. *High. Educ. Policy* 26, 217–242, <http://dx.doi.org/10.1057/hep.2012.28>.
- Pfothenauer, S.M., Wood, D., Roos, D., Newman, D., 2016. Architecting complex international science, technology, and innovation partnerships (CISTIPs): A study of four global MIT collaborations. *Technol. Forecast. Soc. Change* 104, 38–56, <http://dx.doi.org/10.1016/j.techfore.2015.12.006>.
- Ponomariov, B.L., Boardman, P.C., 2010. Influencing scientists’ collaboration and productivity patterns through new institutions: university research centers and scientific and technical human capital. *Res. Policy* 39, 613–624, <http://dx.doi.org/10.1016/j.respol.2010.02.013>.
- Powell, W.W., Koput, K.W., Smith-Doerr, L., 1996. Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. *Adm. Sci. Q.* 41, 116, <http://dx.doi.org/10.2307/2393988>.

- Poyago-Theotoky, J., Beath, J., Siegel, D.S., 2002. Universities and fundamental research: reflections on the growth of university–industry partnerships. *Oxf. Rev. Econ. Policy* 18, 10–21, <http://dx.doi.org/10.1093/oxrep/18.1.10>.
- Price, D.J., Beaver, D.D., 1966. Collaboration in an invisible college. *Am. Psychol.* 21, 1011–1018, <http://dx.doi.org/10.1037/h0024051>.
- Rafols, I., Porter, A.L., Leydesdorff, L., 2010. Science overlay maps: a new tool for research policy and library management. *J. Am. Soc. Inf. Sci. Technol.* 61, 1871–1887, <http://dx.doi.org/10.1002/asi.21368>.
- Rasmussen, E., 2008. Government instruments to support the commercialization of university research: lessons from Canada. *Technovation* 28, 506–517, <http://dx.doi.org/10.1016/j.technovation.2007.12.002>.
- Reagans, R., McEvily, B., 2003. Network structure and knowledge transfer: the effects of cohesion and range. *Adm. Sci. Q.* 48, 240–267.
- Reagans, R., Zuckerman, E., 2001. Networks, diversity, and productivity: the social capital of corporate R&D teams. *Organ. Sci.* 12, 502–517.
- Reddy, N.M., Zhao, L., 1990. International technology transfer: a review. *Res. Policy* 19, 285–307, [http://dx.doi.org/10.1016/0048-7333\(90\)90015-X](http://dx.doi.org/10.1016/0048-7333(90)90015-X).
- Ross, J.S., Gross, C.P., Desai, M.M., Hong, Y., Grant, A.O., Daniels, S.R., Hachinski, V.C., Gibbons, R.J., Gardner, T.J., Krumholz, H.M., 2006. Effect of blinded peer review on abstract acceptance. *JAMA J. Am. Med. Assoc.* 295, 1675–1680, [http://dx.doi.org/10.1016/S0093-3619\(08\)70551-3](http://dx.doi.org/10.1016/S0093-3619(08)70551-3).
- Rubin, D., 1976. Multivariate matching methods that are equal percent bias reducing, I: some examples author (s): Donald B. Rubin. *Biometrics* 32, 109–120.
- Rubin, D., 1980. Bias reduction using mahalanobis-metric matching. *Biometrics* 36, 293–298.
- Saka, A., Igami, M., Kuwahara, T., 2010. Science Map 2008 – Study on Hot Research Areas by Bibliometric Method, National Institute of Science and Technology Policy (NISTEP) Report 139.
- Salter, A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. *Res. Policy* 30, 509–532, [http://dx.doi.org/10.1016/S0048-7333\(00\)00091-3](http://dx.doi.org/10.1016/S0048-7333(00)00091-3).
- Shane, S., 2004. *Academic Entrepreneurship: University Spinoffs and Wealth Creation*. E. Elgar, Cheltenham, UK; Northampton, MA.
- Shrum, W., Chompalov, I., Genuth, J., 2001. Trust, conflict and performance in scientific collaborations. *Soc. Stud. Sci.* 31, 681–730, <http://dx.doi.org/10.1177/030631201031005002>.
- Shrum, W., Genuth, J., Chompalov, I., 2007. *Structures of Scientific Collaboration*. The MIT Press, Cambridge, Mass.
- Siegel, D.S., Waldman, D., Link, A., 2003. Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: an exploratory study. *Res. Policy* 32 (1), 27–48, [http://dx.doi.org/10.1016/S0048-7333\(01\)00196](http://dx.doi.org/10.1016/S0048-7333(01)00196).
- Slaughter, S., Leslie, L.L., 1997. *Academic Capitalism: Politics Policies and the Entrepreneurial University*. John Hopkins University Press.
- Smits, R., Kuhlmann, S., 2004. The rise of systemic instruments in innovation policy. *Int. J. Foresight Innov. Policy* 1, 4–32.
- Stromquist, N.P., Monkman, K., 2000. *Globalization and Education*. Rowman & Littlefield Publishers, Oxford.
- Subramanyam, K., 1983a. Bibliometric studies of research collaboration: a review. *J. Inf. Sci.* 6, 33–38, <http://dx.doi.org/10.1177/016555158300600105>.
- Subramanyam, K., 1983b. Bibliometric studies of research collaboration: a review. *J. Inf. Sci.* 6, 33–38, <http://dx.doi.org/10.1177/016555158300600105>.
- Tödtling, F., Trippl, M., 2005. One size fits all? Towards a differentiated regional innovation policy approach. *Res. Policy* 34, 1203–1219, <http://dx.doi.org/10.1016/j.respol.2005.01.018>.
- Thorp, H.H., Goldstein, B., 2010. *Engines of Innovation: The Entrepreneurial University in the Twenty-First Century*. Univ of North Carolina Press.
- Toivanen, H., Ponomarev, B., 2011. African regional innovation systems: bibliometric analysis of research collaboration patterns 2005–2009. *Scientometrics* 88, 471–493, <http://dx.doi.org/10.1007/s11192-011-0390-1>.
- Verbik, L., Merkle, C., 2006. The international branch campus – models and trends. *Obs. Borderless High. Educ.* 1, 1–31.
- Vinkler, P., 1993. Research contribution, authorship and team cooperativeness. *Scientometrics* 26, 213–230, <http://dx.doi.org/10.1007/BF02016801>.
- Wagner, C.S., 2005a. Six case studies of international collaboration in science. *Scientometrics* 62, 3–26, <http://dx.doi.org/10.1007/s11192-005-0001-0>.
- Wagner, C.S., 2005b. Six case studies of international collaboration in science. *Scientometrics* 62, 3–26, <http://dx.doi.org/10.1007/s11192-005-0001-0>.
- Wagner, C., 2008. *The New Invisible College: Science for Development*. Brookings Institution Press, Washington, DC.
- Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of small-world networks. *Nature* 393, 440–442, <http://dx.doi.org/10.1038/30918>.
- Wei, L., 1995. International technology transfer and development of technological capabilities: a theoretical framework. *Technol. Soc.* 17, 103–120, [http://dx.doi.org/10.1016/0160-791X\(94\)00028-C](http://dx.doi.org/10.1016/0160-791X(94)00028-C).
- Winckler, G., 2010. Innovation strategies of european universities in the triangle of education, research, and innovation. In: Weber, L.E., Duderstadt, J.J. (Eds.), *University Research for Innovation*, Glion Colloquium. *Economica*, pp. 107–124.
- Wuchty, S., Jones, B.F., Uzzi, B., 2007a. The increasing dominance of teams in production of knowledge. *Science* 316 (5827), 1036–1039.
- Wuchty, S., Jones, B.F., Uzzi, B., 2007b. The increasing dominance of teams in production of knowledge. *Science* 316 (5827), 1036–1039.
- Ynalvez, M.A., Shrum, W.M., 2011. Professional networks, scientific collaboration, and publication productivity in resource-constrained research institutions in a developing country. *Res. Policy* 40, 204–216.
- Youtie, J., Shapira, P., 2008. Building an innovation hub: a case study of the transformation of university roles in regional technological and economic development. *Res. Policy* 37, 1188–1204, <http://dx.doi.org/10.1016/j.respol.2008.04.012>.
- Zabusky, S.E., 1995. *Launching Europe: An Ethnography of European Cooperation in Space Science*. Princeton University Press, Princeton, NJ.
- van Geenhuizen, M., Gibson, D., Heitor, M., 2005. *Regional Development and Conditions for Innovation in the Network Society*. Purdue University Press.